Estimating Order Imbalance Using Low Frequency Data

JinGi Ha and Jianfeng Hu *

May 22, 2016

ABSTRACT

Order flow estimation is challenging in markets dominated by high frequency trading. We propose a method to estimate order imbalance using daily CRSP data. The estimated low frequency order imbalance (LFOI) has close relations with aggregate order imbalance estimated using high frequency data (HFOI). The LFOI positively predicts the price changes on the following day and the subsequent price reversal does not fully eliminate the positive price impact, suggesting that LFOI captures both the transitory price pressure and permanent information flow in the trading process, as HFOI does. However, the predictive ability of LFOI is even stronger than HFOI in the cross section of stocks. Subsample analysis validates the robust price impact of LFOI and the predictability extends to the weekly frequency. We also find that LFOI increases significantly around earnings announcements and contains valuable information regarding the announcement returns. The evidence suggests that the proposed LFOI is a good proxy for individual stock's net order flow, and can therefore be used in empirical studies that employ order flow variables at daily or longer horizons.

JEL classification: D82; G14.

*JinGi Ha and Jianfeng Hu are at Singapore Management University. We would like to thank Fangjian Fu, Luis Goncalves-Pinto, Allaudeen Hameed, Dashan Huang, Sheng Huang, Roger Loh, Wenlan Qian, David Reeb, Johan Sulaeman, Yuehua Tang, Qing Tong, Joe Zhang, and the seminar participants at National University of Singapore, Singapore Management University, and Zhejiang University for comments. All remaining errors are ours. Please address correspondence to JinGi Ha (jingiha.2014@pbs.smu.edu.sg, +65 9189 0853) and Jianfeng Hu (jianfenghu@smu.edu.sg) at Lee Kong Chian School of Business, Singapore Management University, 50 Stamford Road, Singapore, 178899.

I. Introduction

This article proposes new methods of estimating individual stock's net order flow using daily data from CRSP. Order flow is at the center of the market microstructure literature because it directly enters the price formation process. In the inventory-based models such as Stoll (1978), Ho and Stoll (1981, 1983), and Cohen, Maier, Schwartz, and Whitecomb (1981), customer order flow sets the dealer away from her optimal portfolio and the dealer adjusts the price quotes subsequently to attract offsetting orders and return to the optimal position. In the information-based models such as Kyle (1985), Glosten and Milgrom (1985), and Easley and O'Hara (1987), the market maker expects orders from both informed and uninformed traders and adjusts both her belief about the terminal value and the price quote in response to the arrival of order flow. Empirically, order flow is only observable to the market makers and the exchanges but not to general market participants or econometricians. Other than a few studies using limited samples of exchange-flagged order flow, most academic researchers need to estimate order flow from public data at the tick level such as the ISSM and TAQ data. Two generic approaches have been used to sign trade directions in order to estimate order flow, the tick rule and the quote rule. The tick rule signs a trade to be buyer-initiated if the trade price is higher than the last different trade price (an up-tick) and seller-initiated if the trade price is lower than the last different trade price (a down-tick). The intuition of the tick rule is the price impact of order flow – impatient buyers tend to take the ask price and push the trade price up and impatient sellers do the opposite with the bid price. One obvious drawback of the tick rule is when transactions are infrequent, the previous trade price may not be a good benchmark to infer the price impact because the liquidity demander can face price quotes far away from the previous transaction price. The second generic approach, the quote rule, compares the transaction price to the quoted bid and ask prices and assigns a trade to be buyer-initiated (seller-initiated) if the trade price is above (below) the midpoint of the bid and ask prices. The midpoint is used instead of the exact bid and ask prices because in ISSM and TAQ data, transactions are not matched to the quotes used to execute

the transactions and the matching needs to be done based on the time stamp. The accuracy of the quote rule largely depends on the matching of trades and quotes and this matching can be difficult when the quotes move fast. Given the imperfection of both rules, Lee and Ready (1991) propose to apply the quote rule first and use the tick rule for those trades fallen on the midpoint of the quote. Their methodology has become the most popular trade signing algorithm since then. The original Lee and Ready method matches a transaction to the national best bid and offer (NBBO) five seconds ago to account for reporting lag in the data. This treatment is shown unnecessary in recent data by Madhavan, Porter, and Weaver (2005) and Chordia, Roll, and Subrahmanyam (2005) as a result of improvement in the market structure. Alternatively, Ellis, Michaely, and O'Hara (2000) propose to apply the tick rule first and then the quote rule for Nasdaq stocks. Concerning the impact from high frequency trading in today's market, Easley, Lopez de Prado, and O'Hara (2012, 2016) propose to group trades into time or volume bulks and apply the tick test to the volume weighted price for the bulks. All these methods employ high frequency data. As the trading records grow at an explosive rate[1], estimating order flow becomes more and more challenging and time-consuming. Computing order flow at the tick level is essential for empirical market microstructure studies. However, in many other applications, order flow is aggregated to a net order imbalance measured at a lower frequency such as daily or weekly. For example, Chordia, Roll, and Subrahmanyam (2002) and Chordia and Subrahmanyam (2004) show that the daily order flow of stocks predict the subsequent return both at the market level and in the cross section. The auction type of models starting from Kyle (1985) also use the net order flow in each auction only. Given the research demand of using low frequency order flow and the technical difficulty of computing it from high frequency data, our goal in this paper is to design an order signing method without using high frequency data. Such methods are useful for empirical studies that only need net order flow information at low frequencies.

We rely on the positive contemporaneous price impact from order flow to determine

the direction of net order flow. Specifically, if a stock has a positive (negative) return on day t, we assign the net order flow of this stock to be positive (negative) on day t. To identify the magnitude of the order imbalance, we use the trading volume measured by share turnover, total number of shares traded, and total dollar volume traded. Among the three volume candidates, the share turnover is a standardized variable using the number of shares outstanding. This standardization potentially accounts for the cross-sectional variation in firm size and is likely to generate a better proxy for relative order flow. We multiply the sign of the return and the trading volume to arrive at the low frequency order imbalance (hereafter, LFOI) estimates. Admittedly, the use of total volume introduces noise in measuring the net order flow because trades at the opposite direction offset each other in net order flow calculation but add up in the total volume. Empirically, the use of this noisy proxy is biased against finding desirable relations between LFOI and the true order imbalance or stock prices. Nevertheless, to mitigate the concern of overestimating the magnitude of the order flow, we also use the original return instead of the sign of return under the notion that the price impact of order imbalance is constant and there is a linear relation between order imbalance and the contemporaneous return. We therefore construct six LFOI measures from daily CRSP: signed turnover, signed trading volume, signed trading dollar volume, the interaction of returns and turnover, the interaction of returns and trading volume, and the interaction of returns and trading dollar volume. To compare the effectiveness of LFOI, we also calculate two high frequency order imbalances (hereafter, HFOI) to set the benchmark. The first HFOI is based on the Lee-Ready (1991) algorithm with five-second quote lag before 1998 and no lag after 1998. The second HFOI is based on the time bulk tick test of Easley, Lopez de Prado, and O'Hara (2012), in which we bulk all trades in the same second into one record and sign the whole bulk of volume based on the volume-weighted transaction price.

In the first empirical analysis, we examine the association of LFOI with HFOI using correlation tests and regressions of HFOI on LFOI. We find the average contemporaneous correlation between HFOI and LFOI is around 0.2 with strong statistical significance.

Specifically, signed turnover (hereafter, LFOI1) and the interaction of returns and turnover (hereafter, LFOI2) have the highest correlations with both HFOI measures. Also all of the LFOIs have statistically significant coefficients in the Fama-MacBech (1973) regression of HFOI on each LFOI even after controlling for stock returns and liquidity factors. In both univariate and multivariate analysis, LFOI1 and LFOI2 still have the highest t-values of estimated coefficients among all LFOI measures. We then explain HFOI using all LFOI measures in a horse race test. The results suggest that LFOI1 and LFOI2 are the best proxies for HFOI among our six LFOIs because they have the most significant coefficient estimates in the test. The results hold in different subsamples based on firm size, liquidity, exchange market, and period subsamples.

Next, we turn to three applications of low frequency order imbalance. The first application is cross-sectional return prediction. Both the inventory and information models suggest that net order flow can affect subsequent stock returns. The prediction of contemporaneous price impact is positive in both models. However, the inventory model predicts an ultimate price reversal because the fundamental value does not change, while the information model predicts a permanent price impact. We find that our proposed and selected LFOIs (LFOI1 and LFOI2) show positive and significant predictive power for future stock returns at daily and weekly frequency. Although a reversal exists beyond one day, the stock price does not fully revert to the level before the order flow occurs, suggesting that the LFOIs capture both liquidity effect and information content. Moreover the LFOIs have even larger t-values than HFOIs in the regressions. The prediction power of LFOIs for future returns is robust in several subsamples based on size, liquidity, exchange market, and time. The second application concerns the fundamental information flow around earnings announcements. We find LFOIs increase significantly around the earnings announcement date for both positive and negative earnings surprises, and are significantly informative about the abnormal announcement returns, suggesting that the easy-to-compute LFOI can be sufficient to capture the presence of informed traders for such event studies. Lastly we test return predictability of LFOI at the

market level. Our LFOIs generate weak and negative coefficient estimates, consistent with Chordia, Roll, and Subrahmanyam's (2002) findings on contrarian trading at the market level.

The main contribution of the study is to systematically test the effectiveness of low frequency order imbalance measures that can be easily computed and used in empirical finance studies on different topics. We show that the signed turnover and the interaction of return and turnover are good proxies for stock order imbalance and contain significant information about future stock returns. For researchers not concerned with high frequency dynamics of price formation, these low frequency proxies can well serve the purpose of detecting price pressure and private information flow at least at the daily frequency.

Our methods are in the nature of the tick test on time bulks by Easley, Lopez de Prado, and O'Hara (2012). However, their method still relies on the intraday tick data as the tick test is performed on volume-weighted transaction price. Compared to their method, our low frequency order signing algorithm uses the end-of-day prices only. As a result, our method computes daily stock order imbalance at a much faster speed with the cost of introducing price noise due to illiquidity and being silence on the high frequency order flow. Nonetheless, in the cross-sectional pricing test, the low frequency order imbalances we propose have even stronger return predictability than the order imbalance based on the bulk tick test.

The low frequency order imbalance we propose is also related to the order imbalance measure in Pastor and Stambough (2003), which is the signed dollar volume based on stock returns. We also employ this proxy and directly compare it to the other low frequency order imbalance measures we propose. The results show that the signed dollar volume has lower correlations to the high frequency order imbalance and has weaker return predictive ability than signed turnover and the interaction of return and turnover in the cross section. The underperformance could be because using turnover instead of dollar volume standardizes the order imbalance measure in the cross section and relates better to anther standardized variable, the stock return. Campbell, Grossman, and Wang (1993) also interact returns and

turnover to predict subsequent returns at the market level. But they do not interpret the interaction as order imbalance. Rather, the turnover is used as a conditional variable in the same way as volatility to study the market return reversal. Unlike Campbell, Grossman, and Wang (1993), our focus is to propose an order flow measure at the individual stock level.

The rest of the paper is organized as follows. Section II describes the sample selection and how to construct our empirical measures of order flow. Section III reports empirical test results. In the first part, we report how well low frequency order flow reflect high frequency order flow. The second part includes three applications of low frequency order imbalance on stock return prediction at the stock level, at the market level, and around earnings announcements. Section IV concludes.

II. Data and variable construction

A. Sample selection

We employ mainly two data sets, Trades and Automated Quotes (hereafter, TAQ) and Center of Research in Security Prices (hereafter, CRSP) in the study. From TAQ, we extract all trade and quote messages between 9 AM to 4 PM EST with positive trading price and trading volume in New York Stock Exchange (NYSE) market, American Stock Exchange (AMEX) market, and National Association of Securities Dealers Automated Quotations (NASDAQ) exchange market. From CRSP, we extract information on common stock characteristics in NYSE, AMEX, and NASDAQ including daily stock return, daily stock price, close bid and ask prices, shares outstanding, and daily trading volume. We exclude observations from CRSP if they have a price lower than five dollars or if their scaled bid-ask spread, defined as bid-ask spread scaled by the average of bid and ask prices, is outside the interval between zero and one half. The sample period of both data sets is from 1 January 1993 to 31 December 2013. We limit our analysis to common stocks only with CRSP code of 10 and 11.

B. Order imbalance measures

We first consider traditional order imbalance measures using high frequency data. We designate each transaction in TAQ dataset as either buyer-initiated or seller-initiated according to the Lee and Ready (1991) algorithm or Easley, Lopez de Prado, and O'Hara (2012) algorithm. The former algorithm is basically the combination of a quote test and a tick test. The transaction is classified as buyer-initiated (seller-initiated) if its trading price is close to the national best bid (ask) price of the prevailing quote. To circumvent the concern on fast moving quotes in the recent sample period, we follow Holden and Jacobsen's (2014) quote adjustment on the monthly TAQ data to construct the NBBO prices after 2001. In the case the trading price is the middle of bid and ask prices, the transaction is classified as buyer-initiated (seller-initiated) if price change prior to the transaction is positive (negative). The latter algorithm is similar to the Lee and Ready algorithm. Main difference is that it conducts a tick test within a certain volume bucket. After classifying buyer-initiated and seller-initiated trades, we construct two following variables, High Frequency Order Flows (hereafter, HFOI) for each stock-day.

- HFOI1: the number of buyer-initiated shares less the number of seller-initiated shares from Lee-Ready algorithm, scaled by the total number of trades for each stock-day.
- HFOI2: the number of buyer-initiated shares less the number of seller-initiated shares from Easley *et al.* algorithm, scaled by the total number of trades for each stock-day.

Next, we propose six proxies of low frequency order imbalances by using daily CRSP data. The direction of the order imbalance is determined by the sign of the stock return on the same day because the contemporaneous price impact from order imbalance should be positive. We use the stock trading turnover ratio, number of shares traded, total dollar volume, and the return adjusted volumes to measure the size of the order imbalance for each stock-day observation. Hence, the low frequency order imbalances are formally defined as follows.

- LFOI1: the sign of the daily return multiplied by the daily turnover ratio for each

stock-day.

- LFOI2: the daily return multiplied by the daily turnover ratio for each stock-day.
- LFOI3: the sign of the daily return multiplied by the daily trading volume for each stock-day.
- LFOI4: the daily return multiplied by the daily trading volume for each stock-day.
- LFOI5: the sign of the daily return multiplied by the daily trading dollar volume for each stock-day.
- LFOI6: the daily return multiplied by the daily trading dollar volume for each stock-day.

In addition, to construct liquidity measures for pricing tests, we calculate the turnover ratio (TURN), bid-ask spread (BASPRD), and Amihud (2002) illiquidity (AMIHUD). The detailed definitions are following.

- TURN: daily trading volume over the number of shares outstanding for each stock-day.
- BASPRD: the difference of bid and ask prices scaled by the average of bid and ask prices for each stock-day.
- AMIHUD: the absolute value of daily returns divided by stock prices and its trading volume for each stock-day.

In this study, we mainly use turnover ratio and bid-ask spread as proxies for liquidity. Amihud illiquidity is only used in correlation table to provide more information on order imbalances (hereafter, OIs). After variable construction, we conduct both time-serial and cross-sectional winsorization at 1 and 99 percent to mitigate the effect of outliers in our sample.

C. Summary statistics

[Place Table I about here]

Table I presents descriptive statistics for OIs and liquidity proxies, averaging cross-sectional statistics over time. The sample covers 5289 working days and the average number of stocks at each working days is around 3930. The average number of stocks in AMIHUD is lower than other variables because some stocks have no trading at some days. To avoid reporting big numbers, we adjust the unit of variables. LFOI3 and LFOI4 are divided by one hundred thousand, and LFOI5 and LFOI6 are divided by one million.

[Place Table II about here]

Table II documents the correlation between OIs, liquidity proxies, and current stock returns. HFOI1 is highly and positively correlated with HFOI2 because Easley *et al.* algorithm is essentially similar to Lee-Ready algorithm except volume bucket usage. LFOI1 and LFOI2 are also correlated with HFOI1 and HFOI2 in moderate amount, comparing with other LFOIs. The correlation coefficients of LFOI1 and LFOI2 may not be high enough to argue they are a good proxy for HFOIs. To get rid of such concern, we conduct further analysis including regression analysis in the next section. All the LFOIs have significant t-statistics in the regression of HFOIs on each LFOI.

LFOI1 and LFOI2, LFOI3 and LFOI4, and LFOI5 and LFOI6 are highly and positively correlated with each other by their definition. LFOI3 and LFOI5 and LFOI4 and LFOI6 are also highly correlated because stock prices are persistent. All the correlations of OIs with turnover rates are positive because high order imbalance implies many trades. In contrast, bid-ask spread is negatively correlated with all OIs because, according to the inventory story, market makers widen bid-ask spread to avoid inventory risk when high order imbalance occurs. Amihud illiquidity looks like it does not have any explanation power to OIs. Lastly, daily returns have reasonable value of correlation, around twenty percent. That means daily returns capture the information on HFOIs. Based on the fact, we contruct six different LFOIs as a proxy for HFOIs by using low frequency data, CRSP. LFOIs have higher correlation coefficient than HFOIs have. Such high correlation may come from the fact that LFOIs

utilize information on stock returns, rather than that LFOIs are good estimation for order imbalance. We will deal with this issue in multivariate regression analysis in the next section.

III. Result

A. *How well does low frequency order imbalance reflect high frequency order imbalance?*

We have seen that LFOIs have reasonable correlation coefficient with HFOIs in the last section. To confirm those correlations are still holding in regression analysis, we test Fama-MacBeth (1973) regression of HFOIs on LFOIs. The estimated coefficient of LFOIs should be significant and positive if LFOIs represent HFOIs. In addition, LFOIs are the combination of return characteristics and liquidity factors, so someone could argue that the correlation between LFOIs and HFOIs might come from returns or liquidity factors in LFOIs. To quench the alternative explanation, we conduct multivariate analysis as well, controlling current returns and liquidity factors. Moreover, in this section we also provide some evidence on which LFOI is the best proxy among six different LFOIs. By using horse race, regression analysis on all the LFOIs in one single regression model, we argue that LFOI1 and LFOI2 are our best proxy for HFOIs.

A.1. Univariate and multivariate analysis

[Place Table III about here]

Table III proves that all the LFOIs have significant and positive estimated coefficient in the following model, even after controlling return and liquidity factors:

$$\text{HFOI}_{i,t} = \alpha_t + \beta_t^0 \text{LFOI}_{i,t} + \sum_{k=1}^{3} \beta_t^k \text{FACTOR}_{i,t} + \epsilon_t \tag{1}$$

, where FACTOR include concurrent returns (RET), bid-ask spread (BASPRD), and turnover (TURN). Both Panel A and Panel B shows that LFOI2 is the best proxy for HFOI1 and HFOI2, respectively, in terms of t-statistics and adjusted R-square. In Panel A, the univariate (multivariate) regression model on LFOI2 has about 5.5% (8.5%) R-square which is the highest R-square among LFOIs. Panel B also reports similar results to Panel A. Note that each LFOI use different unit, so the amount of estimated coefficient are little meaningful to compare LFOIs.

The sign of estimated coefficient in stock returns, RET, is positive because price pressure should be positively correlated with contemporaneous stock returns. The coefficient of bid-ask spread is negative. This is because order imbalance is more likely to be less when market participants are reluctant to buy or sell stocks due to high transaction costs, e.g., bid-ask spread. The coefficient of turnover rate, TURN, has mixed sign, and we cannot tell which sign turnover should have. This is because, by definition, order imbalance is the difference of buy and sell orders while turnover is the sum of them. Panel A and Panel B basically provide similar results to each other, but t-statistics and adjusted R-square in Panel B is higher than those in Panel A.

A.2. Horse race

[Place Table IV about here]

Table IV provides some evidence that LFOI1 and LFOI2 defeat other LFOIs in the horse race among LFOIs of HFOI1 and HFOI2. Especially, LFOI2 still has the highest t-statistics, 59.89. Note again that the amount of estimated coefficient could mislead our interpretation because LFOIs have different units. For the reason, we focus on t-statistics. The other control variables show similar signs to Table III. LFOIs, however, may differently behave in different subsamples on size, liquidity, exchange market, and periods. We conduct subsample tests to confirm LFOI1 and LFOI2 are suitable proxies for HFOIs, comparing with other candidate

LFOIs.

[Place Table V about here]

Table V shows us the result of horse races in four different types of subsamples. In all kind of subsample tests, LFOI1 or LFOI2 always outperform other LFOIs. In particular, LFOI2 is defeating others in most subsamples. Therefore we conclude that LFOI2 is the best proxy and LFOI1 is the second best proxy for HFOIs based on evidence from Table III to Table V.

B. Application of low frequency order imbalance

In Section II.C, LFOIs have reasonable correlation with HFOIs, and in the section III.A.1, we prove that the correlation is not fully explained by stock return and liquidity characteristics. Moreover we extract two best proxies, LFOI1 and LFOI2, for HFOIs among six LFOIs in the section III.A.2. In this section, we will explore the behavior of LFOIs.

LFOI1 and LFOI2 should behave like HFOIs if they represent actual OI well. To confirm this argument, we apply our LFOI proxies to the test of return predictability, following Chordia and Subrahmanyam (2004). Additionally, we do the test of return predictability around earnings announcement dates. In addition, we apply LFOIs to the test of market-level return predictability, following Campbell, Grossman, and Wang (1993). All of our applications show that LFOI1 and LFOI2 act like HFOIs, and interestingly, their prediction power for future returns even outperform HFOIs in our applications except the market level test.

B.1. Return prediction

[Place Table VII about here]

Table VII documents the results of the return predictability test by using the following model:

$$\mathrm{RET}_{i,t} = \alpha_t + \sum_{i=1}^{5} \beta_{t-i}\mathrm{OI}_{i,t-i} + \mathrm{BASPRD}_{i,t-1} + \mathrm{TURN}_{i,t-1} + \sum_{i=1}^{5} \gamma_{t-i}\mathrm{RET}_{i,t-i} + \sum_{i=1}^{5} \theta_{t-i}\mathrm{RET}^2_{i,t-i} + \epsilon_t \tag{2}$$

, where OI stands for order imbalances, BASPRD is bid-ask spread, TURN is turnover rate, and RET is stock return. According to Chordia and Subrahmanyam (2004), the first lagged term of OIs should be positively correlated with current stock return because of positive autocorrelation in OIs. Panel A in Table VII shows us consistent test results with Chordia and Shubrahmanyam (2004). The first lagged terms for HFOIs and LFOIs have positive signs, and the other lagged terms are negative because of return reversal. Interestingly, the prediction power of LFOIs is better than that of HFOIs. Moreover, we report the result of the same return predictability test using mid-quote stock returns instead of raw stock returns to remove a concern on bid-ask bounce within a trading day. The result of mid-quote stock returns is almost same as that of raw stock returns, which implies that bid-ask bounce has little influence on our test results.

We also add more control variables, bid-ask spread (BASPRD), turnover rate (TURN), lagged returns (RET), and lagged squared returns (RET^2), which Chordia and Subrahmanyam (2004) do not include. We put those control variables to isolate the effect of lagged OIs on current stock returns. Bid-ask spread has a positive sign, which is consistent with Amihud and Mendelson (1986, 1989). This is because, according to the model in Amihud and Mendelson (1986), market participants expect higher returns when they put their money into stocks with wider bid-ask spread. In addition, all of the lagged returns are negative because of stock return reversal. The lagged squared returns represent volatility of returns, so it is natural that higher lagged squared returns lead higher current returns.

Panel B in Table VII presents the weekly-based test results of return predictability. The key difference of Panel B from Panel A is that HFOI1 lose its return prediction power. This difference might come from the limit of Lee-Ready (1991) algorithm which cannot

perfectly recognize actual transactions, especially in high-frequency trading environment (Odders-White (2000)). HFOI2 is also problematic because it has negative sign in the return prediction test of mid-quote returns. We cannot explain why the negative sign of HFOI2 occurs in weekly-based regression.

[Place Table ?? about here]

Stock price should be permanently changed by OIs if they deliver information on future returns. However Table VII indicates that return reversal significantly occurs and it may be equivalent to or bigger than price change by OIs. In that case, OI is interpreted at most as the source of price pressure, rather than the source of future price information. To defeat the alternative explanation, we regress cumulative returns from time t to time $t+4$ on the same control variables as Table VII. Table **??** shows that, even after controlling volume-return reversal from time $t+1$ to time $t+4$, the first lagged OIs are still positively significant. That implies that volume-return reversals are not stronger than price pressure from OIs and therefore OIs permanently change stock price.

We also conduct subsample tests with different criteria including size, liquidity, exchange market, and period. In all of the subsample tests, their behavior are similar to HFOIs, and their predictive power is stronger than HFOIs.

[Place Table VIII about here]

Table VIII presents the predictive power of order imbalances in size subsamples. Generally speaking, small size stocks are naturally illiquid and information asymmetric because of their low price and little attention from market participants. Therefore small size stocks should be vulnerable to price pressure from OIs. Table VIII shows the tendency; all of OIs have the strongest prediction power for current stock returns in a small size subsample. This result is consistent with Chordia and Subrahmanyam (2004). Although HFOI1 lose its predictive power in medium and large size subsamples, lagged HFOI2, LFOI1, and LFOI2 are able to predict current returns in medium and large size subsamples.

[Place Table IX about here]

Table IX reports return predictability of order imbalances in liquidity subsamples. By definition, liquidity is the amount of trading volume without any price change, and OI makes price pressure. Therefore liquid stocks do not easily react to OI, while illiquid stocks react more to OI than liquid stocks. In Table IX all type of OIs have higher t-statistics in an illiquid subsample than in a liquid subsample.

[Place Table X about here]

Table X documents the prediction power of order imbalances in different exchange markets. OIs have return prediction power either in NYSE and AMEX in Panel A or in NASDAQ in Panel B. Since NASDAQ is holding smaller size stocks comparing with NYSE and AMEX, the effect of order imbalance is stronger on NASDAQ than on NYSE and AMEX. Table X reports such tendency. However HFOI1 in Panel B of Table X has an insignificant estimated coefficient of the first lagged term because of outliers in NASDAQ. As Odders-White (2000) reports in his paper, Lee-Ready (1991) algorithm systematically misclassifies transactions, so it loses many actual transactions. For instance, there are only three non-zero traded stocks out of about 4,000 stocks in NASDAQ at 01 December 1993. Therefore we need to be careful of the interpretation of HFOI1 test results, particularly in Panel B.

[Place Table XI about here]

Table XI reports return predictability of OIs during three subperiod. Trading behavior has changed over time. Exchange markets are getting more and more efficient and trading frequency is getting faster. Therefore the period for volume-return reversal should be getting shorter and weaker over time. In early period from 1993 to 2000, the fifth lagged term of OIs have negative sign, which implies volume-return reversal occurs at least for five days. In late period from 2008 to 2013, however, HFOI2, LFOI1, and LFOI2 do not have any significant estimated coefficients after the first lagged term, and their lagged returns also

have smaller coefficient than those in early period. Interpretation of HFOI1 is tricky since their behavior is not consistent with other OIs. We should note that, even though Lee and Radhakrishna (2000) and Chakrabarty, Moulton, Shkilko (2012) document the accuracy of Lee-Ready (1991) algorithm, misclassification may occur more in late period than in early period because of a number of high frequency trades as shown by Odders-White (2000).

B.2. Return prediction around corporate events

[Place Table XIV about here]

Table XIV presents the test result of return predictability around earnings announcement dates. We collect earnings announcement dates from Institutional Brokers' Estimate System (I/B/E/S) dataset. To construct cumulative abnormal returns (hereafter, CAR), we regress raw returns on Fama French three factors including market excess returns (MKT), small-minus-big (SMB), and high-minus-low (HML). After that we compute cumulative abnormal returns around earnings announcement dates. In Table XIV, we use three different CARs, CAR_t, $CAR_{t-1,t+1}$, and $CAR_{t-2,t+2}$, to make sure that stock prices react to earnings announcement.

Table XIV provides similar test results to Table VII. That is, lagged LFOIs have similar effect on current returns to lagged HFOIs even around earnings announcement dates. Lagged order imbalances in Panel A, Panel B, and Panel C can predict current stock returns although HFOI1 is not significant. The insignificance of HFOI1 may come from the limit of Lee-Ready (1991) algorithm as Panel B in Table VII (Odders-White (2000)). Bid-ask spread is also significantly positive in all three Panels, which is consistent with Amihud and Mendelson (1986). Stock returns has negative signs due to stock return reversal. Squared stock returns, representing return volatility, are positive in all three Panels.

[Place Figure 2 about here]

Figure 2 documents the dynamic of HFOIs and LFOIs in good and bad earnings news.

We classify earnings news into good and bad news based on cumulative abnormal returns (CARs) from two days before an earnings announcement date to two days after an earnings announcement date. Abnormal returns are defined as residual terms in a regression model of raw returns on a Fama-French three factor model. If CAR is positive, then we consider the earnings news to be good, and vice versa.

As Figure 2 shown, LFOIs have similar behaviors to HFOIs near earnings announcement dates. Also LFOIs differently react to the content of earnings announcement. LFOIs have positive OIs in good earnings news, while they have negative OIs in bad earnings news. Moreover, LFOIs suddenly increase right before an earnings announcement date and start to decrease after an earnings announcement date.

B.3. Return prediction at market level

[Place Table XV about here]

Table XV documents test results of return predictability of market level OIs. We employ value-weighted market returns (hereafter, VWRETD) or S&P returns as a proxy for market returns. To construct market-level order imbalances, we take a value-weight average of each OI over whole markets, NYSE, AMEX, and NASDAQ. As a proxy for market-level bid-ask spread (BASPRD) and turnover rate (TURN), we also use value-weighted average of BASPRD and TURN over whole markets. Following Campbell, Grossman, and Wang (1993), we add weekday dummies to control weekday effect.

Table XV shows that return prediction power of OIs is not significant at market level. The first lagged terms of HFOI1 and LFOI1 are insignificant in either the regression model of VWRETD in Panel A or the model of S&P returns in Panel B. Additionally HFOI2 shows weakly positive sign in both Panel A and Panel B, but it is rejected at 1% significance level. Lastly, value-weighted LFOI2 is negative in the regression model of VWRETD in Panel A. Campbell, Grossman, and Wang (1993) also report similar results to the negative sign

of value-weighted LFOI2 in Panel A. They argue that lagged market return times lagged trading volume represents volume-related return reversal, so lagged market returns given the certain amount of trading volume are negatively correlated with current market returns.

Liquidity factors in market level analysis lose their predictive power for market returns. However, return reversal still exists even at market level, and lagged squared market returns also have positive and significant estimated coefficient in both Panel A and Panel B.

To sum up, LFOI1 or LFOI2 is a good proxy for HFOIs. It has a reasonable correlation with HFOIs. Moreover the correlation of LFOIs and HFOIs does not disappear even after controlling the characteristics of stock returns and liquidity factors. Also LFOI1 or LFOI2 shows similar behavior to HFOIs and stronger predictive power for future returns in three different applications.

IV. Conclusion

In this paper, we propose to estimate low frequency order flow using daily CRSP data. The empirical analysis shows that the correlation of LFOI and HFOI is reasonably high. We also find that the estimated coefficient in the regression of HFOI on LFOI is positive and significant in different subsamples. Out of all the LFOI measures we consider, the signed turnover ratio and the interaction of return and turnover are closest to the HFOI. Then we show that LFOI has return predictive power at daily and weekly frequency, and we find that LFOI can capture the fundamental information flow around earnings announcements. Finally we show that LFOI has a weak and negative coefficient in a return prediction test at the market level, consistent with the evidence in Chordia, Roll, and Subrahmanyam (2002).

Our proposed LFOI is practically useful. LFOI can be calculated in very short time while HFOI can take much longer and greater computing power to calculate due to the increasing size of data sets. Also the easy-to-compute order flow are even more informative in terms of return prediction power. Its predictive power for stock returns still holds in a variety of

subsamples including size, liquidity, exchange market, and period subsamples. The empirical results in this paper suggests that the proposed LFOI is a good proxy for the information in order flow and it can be applied in empirical studies that utilize order flow at low frequency.

REFERENCES

[1] Amihud, Y., and H. Mendelson, 1986, Asset pricing and the bid ask spread, Journal of Financial Economics 17, 223-249.

[2] Amihud, Y., and H. Mendelson, 1989, The effects of beta, bid-ask spread, residual risk, and size on stock returns, Journal of Finance 44, 479-486.

[3] Campbell, John Y.; Grossman, Sanford J.; Wang, Jiang, 1993, Trading volume and serial correlation in stock returns, The Quarterly Journal of Economics 108, 35.

[4] Chakrabarty, Bidisha, Pamela C. Moulton, and Andriy Shkilko, 2012, Short sales, long sales, and the lee-ready trade classification algorithm revisited, Journal of Financial Markets 15, 467-491.

[5] Chordia, T., R. Roll, and A. Subrahmanyam, 2002, Order imbalance, liquidity, and market returns, Journal of Financial Economics 65, 111-130.

[6] Chordia, T., R. Roll, and A. Subrahmanyam, 2005, Evidence on the speed of convergence to market efficiency, Journal of Financial Economics 76, 271-292.

[7] Chordia, T., and A. Subrahmanyam, 2004, Order imbalance and individual stock returns: Theory and evidence, Journal of Financial Economics 72, 485-518.

[8] Cohen, K. J., S. F. Maier, R. A. Schwartz, and D. K. Whitcomb, 1981, Transcation costs, order placement strategy, and existence of the bid-ask spread, Journal of Political Economy 89, 287-305.

[9] Easley, David, Marcos M. Lopez de Prado, and Maureen O'Hara, 2012, Flow toxicity and liquidity in a high-frequency world, Review of Financial Studies 25, 1457-1493.

[10] Easley, D., and M. Ohara, 1987, Price, trade size, and information in securities markets, Journal of Financial Economics 19, 69-90.

[11] Easley David; Lopez de Prado, Marcos; O'Hara Maureen, 2016, Discerning information from trade data, Journal of Financial and Economics.

[12] Ellis, K., R. Michaely, and M. O'Hara, 2000, The accuracy of trade classification rules: Evidence from nasdaq, Journal of Financial and Quantitative Analysis 35, 529-551.

[13] Fama, E. F., and J. D. Macbeth, 1973, Risk, return, and equilibirum: Empirical tests, Journal of Political Economy 81, 607-636.

[14] Glosten, L. R., and P. R. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, Journal of Financial Economics 14, 71-100.

[15] Ho, T., and H. R. Stoll, 1981, Optimal dealer pricing under transactions and return uncertainty, Journal of Financial Economics 9, 47-73.

[16] Ho, T. S. Y., and H. R. Stoll, 1983, The dynamics of dealer markets under competition, Journal of Finance 38, 1053-1074.

[17] Holden, Craig W., and Stacey Jacobsen, 2014, Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions, Journal of Finance 69, 1747-1785.

[18] Kyle, A. S., 1985, Continuous auctions and insider trading, Econometrica 53, 1315-1335.

[19] Lee, C. M. C., and M. J. Ready, 1991, Inferring trade direction from intraday data, Journal of Finance 46, 733-746.

[20] Llorente, G., R. Michaely, G. Saar, and J. Wang, 2002, Dynamic volume-return relation of individual stocks, Review of Financial Studies 15, 1005-1047.

[21] Madhavan, A., D. Porter, and D. Weaver, 2005, Should securities markets be transparent?, Journal of Financial Markets 8, 265-287.

[22] Newey, W. K., and K. D. West, 1987, A simple, positive semidefinite, heteroskedasticity and autocorrelation cosistent covariance-matrix, Econometrica 55, 703-708.

[23] Odders-White, Elizabeth R., 2000, On the occurrence and consequences of inaccurate trade classification, Journal of Financial Markets 27.

[24] Pastor, L., and R. F. Stambaugh, 2003, Liquidity risk and expected stock returns, Journal of Political Economy 111, 642-685.

[25] Radhakrishna, Lee and, 2000, Inferring investor behavior - evidence from torq data, Journal of Financial Market 3, 29.

[26] Roll, Richard, Eduardo Schwartz, and Avanidhar Subrahmanyam, 2014, Trading activity in the equity market and its contingent claims: An empirical investigation, Journal of Empirical Finance 28, 13-35.

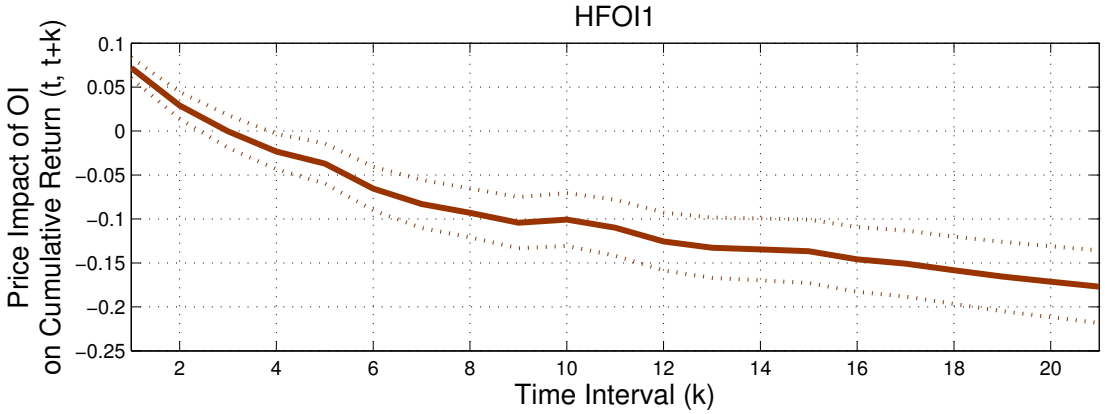
[27] Stoll, H. R., 1978, Supply of dealer services in securities markets, Journal of Finance 33, 1133-1151.

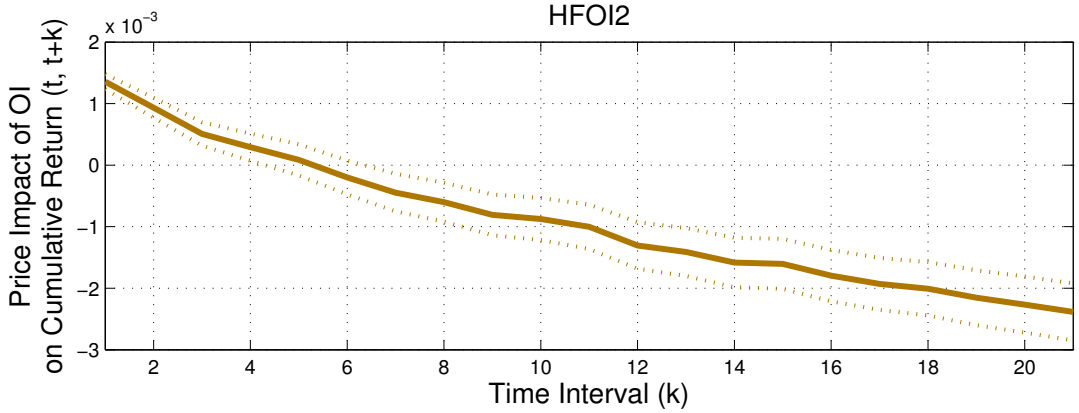
Notes

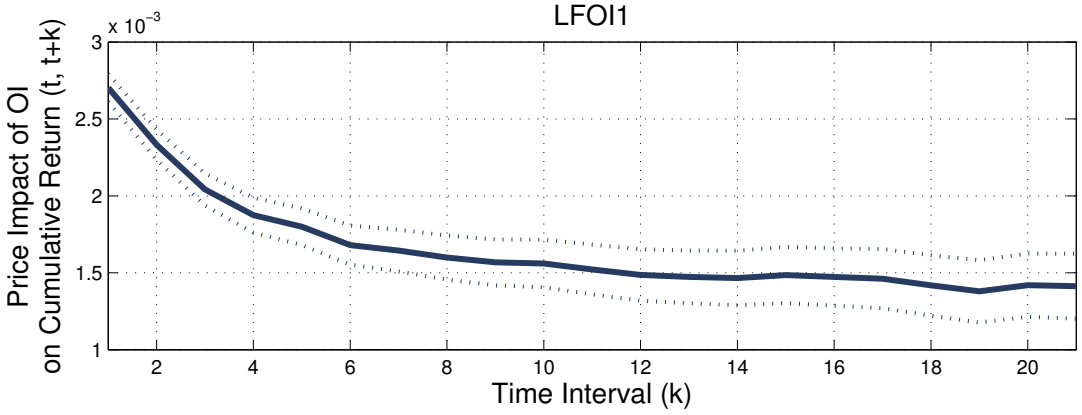
[1]For example, the total equity transactions data on October 30, 2014 is a compressed zip file of 271 MB in the NYSE TAQ database sample and the compressed file of total best bid and offer data on the same day reaches 6.6 GB. See *ftp://ftp.nyxdata.com*.

Figure 1.
Price Impact of OI on Future Cumulative Returns

This figure provides average beta in the Fama-MacBeth regression of cumulative raw returns on order imbalances. The sample period is from 01 January 1993 to 31 December 2013. The horizontal axis is for the interval (k) in cumulative returns from time t to time $t + k$. The vertical axis is for averged amount of beta in the regression of cumulative raw returns on order imbalance. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2011). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. The top figure is for the price impact of HFOI1, the second top figure is for HFOI2, the second bottom figure is for LFOI1, and the bottom figure is for LFOI2.


HFOI1
Price Impact of OI
on Cumulative Return (t, t+k)
0.1
0.05
0
−0.05
−0.1
−0.15
−0.2
−0.25
2
4
6
8
10
12
14
16
18
20
Time Interval (k)


HFOI2
x 10^−3
Price Impact of OI
on Cumulative Return (t, t+k)
2
1
0
−1
−2
−3
2
4
6
8
10
12
14
16
18
20
Time Interval (k)


LFOI1
x 10^−3
Price Impact of OI
on Cumulative Return (t, t+k)
3
2.5
2
1.5
1
2
4
6
8
10
12
14
16
18
20
Time Interval (k)

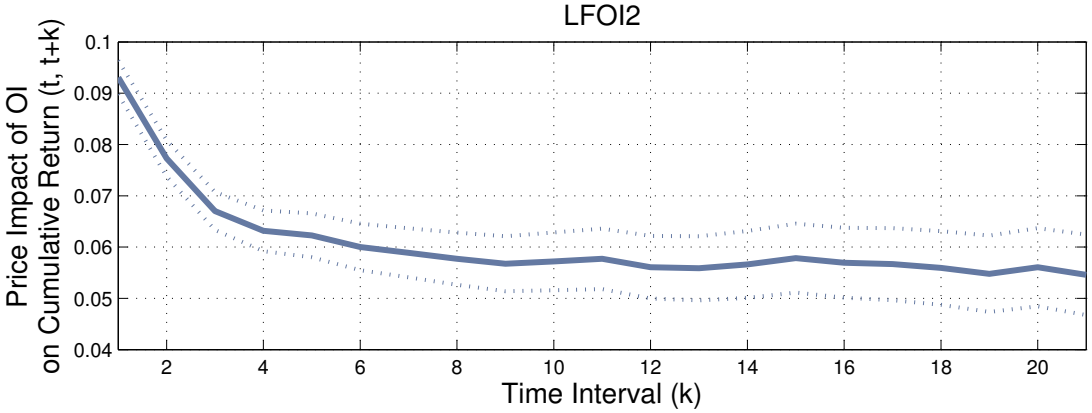
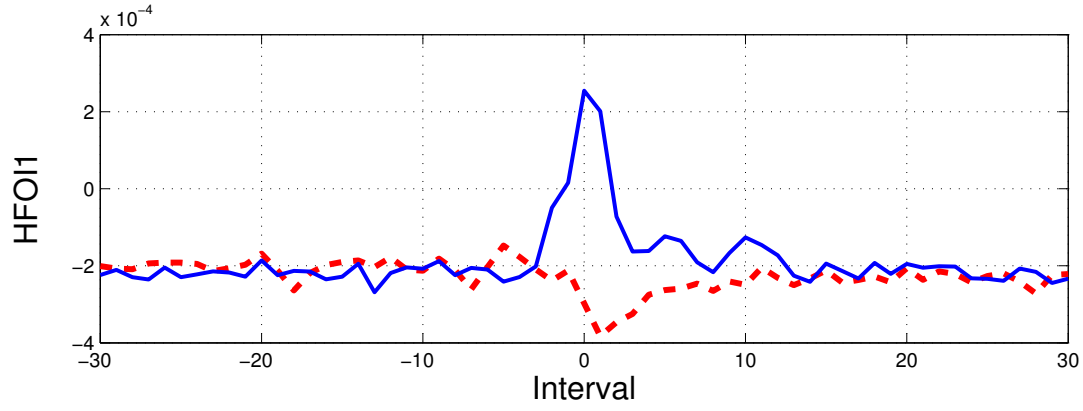
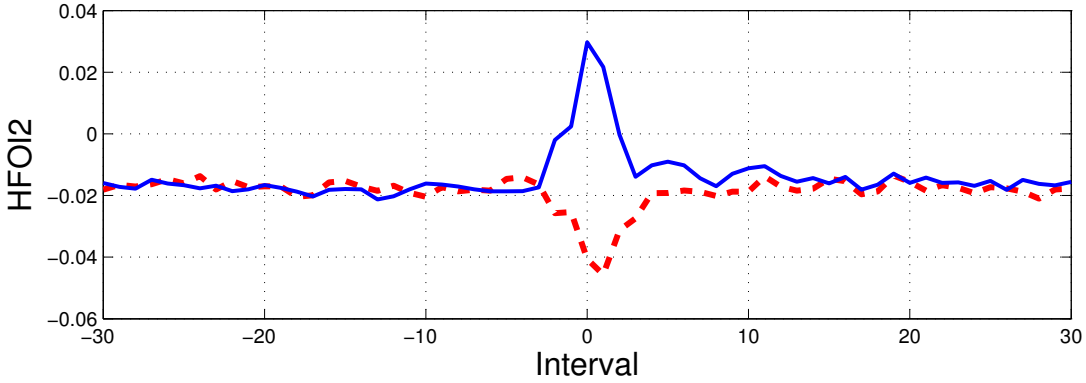
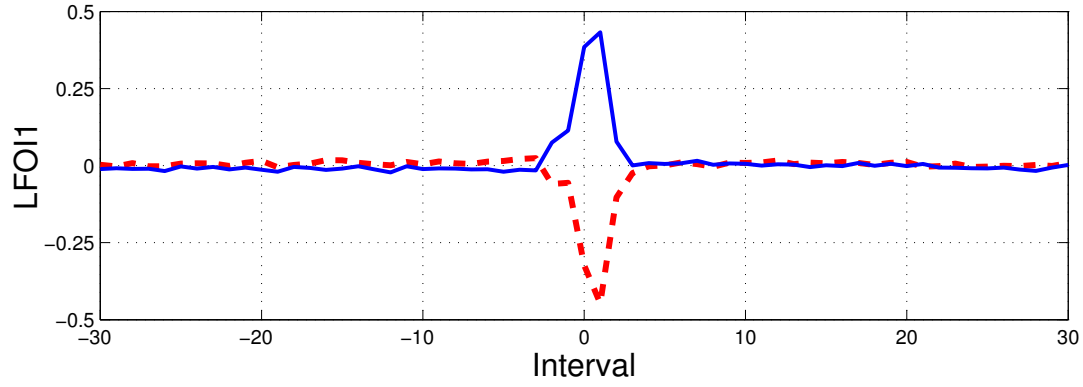

LFOI2
Price Impact of OI
on Cumulative Return (t, t+k)
0.1
0.09
0.08
0.07
0.06
0.05
0.04
2
4
6
8
10
12
14
16
18
20
Time Interval (k)

Figure 2.
The Dynamic of Order Imbalance around Earnings Announcement

This figure provides the time-serial average of abnormal cross-sectional order imbalances around earnings announcement dates. Abnormal cross-sectional order imbalances are cross-sectional order imbalances minus value-weighted market order imbalance. The sample period is from 01 January 1993 to 31 December 2013. We extract earnings announcement dates of each firms from the Institutional Brokers' Estimate System (I/B/E/S) dataset. The horizontal axis is for the interval of dates from earnings announcement dates. The vertical axis is for the average amount of order imbalances. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2013). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI4 is daily returns multiplied by daily turnover rate for each stock. We classify earnings contents into good and bad news based on cumulative abnormal returns from two days before earnings announcement date to two days after earnings announcement date. If the cumulative returns are positive, then we define the earnings news as good news, and vice versa. Abnormal returns are defined as residual terms in the regression model of raw returns on Fama-French three factor model. The top figure is for the behavior of HFOI1 near earnings announcement dates. The second top figure is for HFOI2. The second bottom figure is for LFOI1. The bottom figure is for LFOI4. In each figure, the blue solid (red dotted) line represents the dynamic of order imbalance for good (bad) earnings news.


x 10^−4
HFOI1
Interval


HFOI2
Interval


LFOI1
Interval

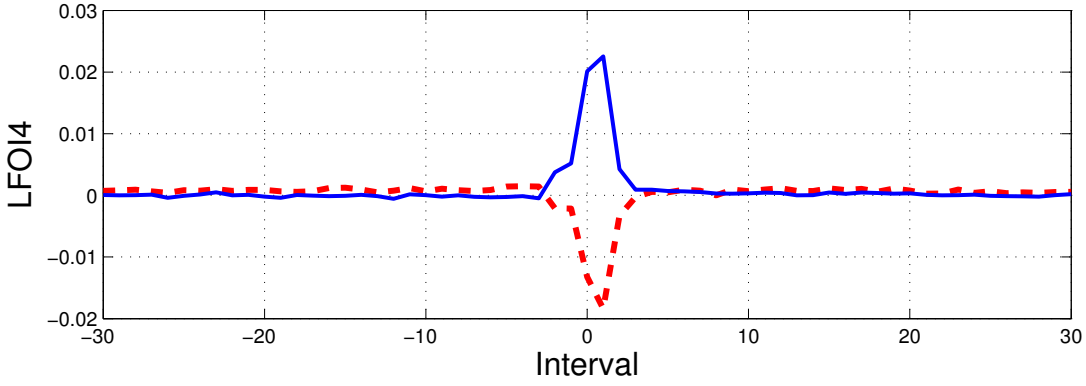

LFOI4
Interval

Table I
Summary statistics

This table documents descriptive statistics from 01 January 1993 to 31 December 2013 sample period. All the statistics are time-serial average of cross-sectional values over sample period. Number of Dates stands for the number of working dates during sample period. Average Number of Stocks is the average number of firms for each date. Moreover, this table reports mean, standard deviation, minimum, and maximum of each variable. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close prices, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2013, 2016). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. LFOI3 is the sign of daily returns multiplied by daily trading volume for each stock. LFOI4 is daily returns multiplied by daily trading volume for each stock. We scale LFOI3 and LFOI4 by dividing one hundred thousand. LFOI5 is the sign of daily returns multiplied by daily trading dollar volume for each stock. LFOI6 is daily returns multiplied by daily trading dollar volume for each stock. We scale LFOI5 and LFOI6 by dividing one million. RET is daily stock returns. TURN is daily turnover rate for each stock, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted bid-ask spread, the difference of bid and ask prices divided by the average of bid and ask prices. We only use bid-ask spread within the interval between 0 and 0.5. Also bid and ask prices are close price at each day. AMIHUD is the absolute value of daily returns divided by stock prices and its trading volume, following Amihud (2002).

	Number of Dates	Average Number of Stocks	Mean	Standard Deviation	Minimum	Maximum
HFOI1	5289	3929.77	0.00	0.00	-0.08	0.06
HFOI2	5289	3928.06	0.01	0.27	-6.66	5.08
LFOI1	5289	3937.24	0.02	0.77	-2.87	3.10
LFOI2	5289	3937.24	0.00	0.03	-0.12	0.16
LFOI3	5289	3937.24	0.01	1.22	-5.51	5.61
LFOI4	5289	3937.24	0.00	0.03	-0.14	0.15
LFOI5	5289	3937.24	0.06	4.28	-19.31	20.30
LFOI6	5289	3937.24	0.00	0.09	-0.40	0.47
RET	5289	3937.24	0.00	0.03	-0.27	0.45
TURN	5289	3938.59	0.56	0.70	0.00	4.38
BASPRD	5289	3938.60	0.02	0.02	0.00	0.10
AMIHUD	5289	3743.98	0.00	0.00	0.00	0.00

Table II
Correation table

This table reports a correlation table between OIs and other variables. The sample period is from 01 January 1993 to 31 December 2013. Correlation coefficients are the average of cross-sectional correlation over sample period. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2013). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. LFOI3 is the sign of daily returns multiplied by daily trading volume for each stock. LFOI4 is daily returns multiplied by daily trading volume for each stock. We scale LFOI3 and LFOI4 by dividing one hundred thousand. LFOI5 is the sign of daily returns multiplied by daily trading dollar volume for each stock. LFOI6 is daily returns multiplied by daily trading dollar volume for each stock. We scale LFOI5 and LFOI6 by dividing one million. RET is daily stock returns. TURN is daily turnover rate for each stock, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted bid-ask spread, the difference of bid and ask prices divided by the average of bid and ask prices. We only use bid-ask spread within the interval between 0 and 0.5. Also bid and ask prices are close price at each day. AMIHUD is the absolute value of daily returns divided by stock prices and its trading volume, following Amihud (2002).

	HFOI1	HFOI2	LFOI1	LFOI2	LFOI3	LFOI4	LFOI5	LFOI6
HFOI1	1.00							
HFOI2	0.55	1.00						
LFOI1	0.17	0.22	1.00					
LFOI2	0.19	0.24	0.82	1.00				
LFOI3	0.07	0.10	0.56	0.41	1.00			
LFOI4	0.12	0.15	0.59	0.64	0.83	1.00		
LFOI5	0.05	0.07	0.45	0.31	0.90	0.71	1.00	
LFOI6	0.09	0.12	0.51	0.51	0.81	0.89	0.84	1.00
RET	0.15	0.21	0.56	0.68	0.28	0.43	0.21	0.34
TURN	0.06	0.05	0.08	0.13	0.03	0.08	0.02	0.06
BASPRD	-0.03	-0.03	-0.01	-0.01	-0.01	-0.01	-0.01	-0.02
AMIHUD	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00

Table III
Univariate and Multivariate Regression Analysis of HFOIs on each LFOI

This table presents the Fama-MacBeth coefficient estimates from univariate and multivariate analyses. The sample period is from 01 January 1993 to 31 December 2013. The first row represents what kind of LFOI uses in a regression model. For example, the first and second column takes LFOI1 as a key control variable, LFOI. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2013). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. LFOI3 is the sign of daily returns multiplied by daily trading volume for each stock. LFOI4 is daily returns multiplied by daily trading volume for each stock. We scale LFOI3 and LFOI4 by dividing one hundred thousand. LFOI5 is the sign of daily returns multiplied by daily trading dollar volume for each stock. LFOI6 is daily returns multiplied by daily trading dollar volume for each stock. We scale LFOI5 and LFOI6 by dividing one million. TURN is daily turnover rate for each stock, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted bid-ask spread, the difference of bid and ask prices divided by the average of bid and ask prices. We only use bid-ask spread within the interval between 0 and 0.5. Also bid and ask prices are close price at each day. ALPHA is the intercept of regression models. ADJRSQ is adjusted R-square. OBS is the number of observation in each regression model. Panel A takes HFOI1 as a dependent variable, while Panel B takes HFOI2 as a dependent variable. In brackets, we report t-statistics of the average coefficient over sample period based on New-West (1987) standard errors. ***, **, and * indicate statistical significance at the 1, 5, and 10 per cent level, respectively.

Table. III (Continued)

Panel A. HFOI1, Lee-Ready OI scaled by shares outstanding (x100)												
	LFOI1	LFOI1	LFOI2	LFOI2	LFOI3	LFOI3	LFOI4	LFOI4	LFOI5	LFOI5	LFOI6	LFOI6
LFOI	0.094***	0.068***	2.605***	2.351***	0.076***	0.040***	4.474***	2.745***	0.014***	0.007***	1.177***	0.684***
	[75.10]	[65.49]	[83.00]	[69.82]	[52.02]	[39.65]	[52.85]	[41.31]	[45.93]	[32.33]	[47.60]	[36.42]
RET		0.773***		0.317***		1.326***		1.174***		1.371***		1.287***
		[34.57]		[12.82]		[61.37]		[54.57]		[63.15]		[59.89]
TURN		0.014***		0.004***		0.020***		0.015***		0.020***		0.016***
		[10.91]		[2.90]		[15.57]		[11.86]		[16.01]		[13.30]
BASPRD		-0.283***		-0.335***		-0.271***		-0.275***		-0.266***		-0.261***
		[-34.40]		[-38.54]		[-32.70]		[-33.02]		[-31.96]		[-31.41]
ALPHA	0.002***	-0.002***	0.000	0.002***	0.004***	-0.003***	0.002***	-0.002***	0.005***	-0.004***	0.003***	-0.003***
	[7.54]	[-3.84]	[1.44]	[4.82]	[11.73]	[-8.59]	[6.72]	[-5.94]	[12.09]	[-9.37]	[7.95]	[-8.06]
ADJRSQ	0.043***	0.079***	0.055***	0.085***	0.012***	0.067***	0.027***	0.072***	0.008***	0.066***	0.019***	0.070***
	[71.89]	[83.44]	[70.64]	[83.99]	[54.95]	[75.64]	[59.95]	[78.04]	[48.49]	[74.30]	[54.28]	[76.54]
OBS	20,777,518	20,777,518	20,777,518	20,777,518	20,777,518	20,777,518	20,777,518	20,777,518	20,777,518	20,777,518	20,777,518	20,777,518
Panel B. HFOI2, Bucket Volume OI scaled by shares outstanding (x100)												
	LFOI1	LFOI1	LFOI2	LFOI2	LFOI3	LFOI3	LFOI4	LFOI4	LFOI5	LFOI5	LFOI6	LFOI6
LFOI	9.825***	7.163***	280.6***	235.0***	8.431***	4.558***	502.3***	293.2***	1.613***	0.841***	134.8***	74.55***
	[95.44]	[77.96]	[106.62]	[85.71]	[52.26]	[42.70]	[52.38]	[42.54]	[48.74]	[38.27]	[48.02]	[38.70]
RET		104.9***		64.06***		164.5***		148.2***		170.4***		160.6***
		[66.74]		[35.96]		[108.98]		[97.64]		[112.38]		[106.11]
TURN		0.230***		-0.339***		0.883***		0.567***		0.930***		0.707***
		[3.25]		[-4.47]		[12.55]		[8.04]		[13.15]		[10.03]
BASPRD		-40.08***		-44.77***		;-39.32***		;-39.95***		-38.97***		-38.91***
		[-52.04]		[-57.39]		[-51.04]		[-51.69]		[-50.74]		[-50.60]
ALPHA	0.590***	0.883***	0.484***	1.119***	0.778***	0.698***	0.639***	0.783***	0.795***	0.673***	0.674***	0.720***
	[26.81]	[33.58]	[22.15]	[40.21]	[28.87]	[26.27]	[25.30]	[29.58]	[28.32]	[25.46]	[25.65]	[27.29]
ADJRSQ	0.055***	0.086***	0.065***	0.090***	0.013***	0.071***	0.028***	0.074***	0.007***	0.069***	0.018***	0.072***
	[113.17]	[114.25]	[118.28]	[116.89]	[86.87]	[97.47]	[93.89]	[101.30]	[77.79]	[95.25]	[83.94]	[98.21]
OBS	20,768,454	20,768,454	20,768,454	20,768,454	20,768,454	20,768,454	20,768,454	20,768,454	20,768,454	20,768,454	20,768,454	20,768,454

Table IV
Horse race between LFOIs

This table documents the Fama-MacBeth coefficient estimates of horse race regression. The sample period is from 01 January 1993 to 31 December 2012. The first (second) column takes HFOI1 (HFOI2) as a dependent variable. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2013). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. LFOI3 is the sign of daily returns multiplied by daily trading volume for each stock. LFOI4 is daily returns multiplied by daily trading volume for each stock. We scale LFOI3 and LFOI4 by dividing one hundred thousand. LFOI5 is the sign of daily returns multiplied by daily trading dollar volume for each stock. LFOI6 is daily returns multiplied by daily trading dollar volume for each stock. We scale LFOI5 and LFOI6 by dividing one million. TURN is daily turnover rate for each stock, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted bid-ask spread, the difference of bid and ask prices divided by the average of bid and ask prices. We only use bid-ask spread within the interval between 0 and 0.5. Also bid and ask prices are close price at each day. ALPHA is the intercept of regression models. ADJRSQ is adjusted R-square. OBS is the number of observation in each regression model. In brackets, we report t-statistics of the average coefficient over sample period based on New-West (1987) standard errors. ***, **, and * indicate statistical significance at the 1, 5, and 10 per cent level, respectively.

Table. IV (Continued)

	HFOI1 (x100)	HFOI2 (x100)
LFOI1	0.030***	4.112***
	[40.64]	[55.24]
LFOI2	1.840***	151.960***
	[59.89]	[63.71]
LFOI3	-0.000	0.367**
	[-0.19]	[2.06]
LFOI4	-0.384***	-35.601***
	[-4.05]	[-4.07]
LFOI5	-0.005***	-0.563***
	[-11.09]	[-12.68]
LFOI6	0.392***	37.235***
	[13.01]	[13.31]
RET	0.299***	63.030***
	[12.08]	[35.43]
TURN	0.002*	-0.338***
	[1.82]	[-4.46]
BASPRD	-0.296***	-40.575***
	[-35.14]	[-53.99]
ALPHA	0.001**	0.992***
	[2.16]	[36.65]
ADJRSQ	0.093***	0.097***
	[91.31]	[123.21]
OBS	20,777,518	20,768,454

Table V
Horse race between LFOIs within subsamples

This table reports the Fama-MacBeth estimate coefficient of horse race regression within different subsamles. The sample period is from 01 January 1993 to 31 December 2013. The first row of the table represent which HFOI is the dependent variable of a regression model. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2013). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. LFOI3 is the sign of daily returns multiplied by daily trading volume for each stock. LFOI4 is daily returns multiplied by daily trading volume for each stock. We scale LFOI3 and LFOI4 by dividing one hundred thousand. LFOI5 is the sign of daily returns multiplied by daily trading dollar volume for each stock. LFOI6 is daily returns multiplied by daily trading dollar volume for each stock. We scale LFOI5 and LFOI6 by dividing one million. TURN is daily turnover rate for each stock, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted bid-ask spread, the difference of bid and ask prices divided by the average of bid and ask prices. We only use bid-ask spread within the interval between 0 and 0.5. Also bid and ask prices are close price at each day. ALPHA is the intercept of regression models. ADJRSQ is adjusted R-square. OBS is the number of observation in each regression model. Panel A and Panel B separate the sample dataset into five subsamples based on market capitalization and bid-ask spread, respectively. We report test results of only three subsample regression in smallest, medium, and largest size and liquidity subsamples, respectively. Panel C separates into two subsamples based on exchange markets; NYSE and AMEX and NASDAQ. Panel D seperate into three subsamples based on subperiod; early period from 1993 to 2000, midst period from 2001 to 2006, and late period from 2007 to 2012. In brackets, we report t-statistics of the average coefficient over sample period based on New-West (1987) standard errors. ***, **, and * indicate statistical significance at the 1, 5, and 10 per cent level, respectively.

Table. V (Continued)

Panel A. Size

	Small		Medium		Large	
	HFOI1 (x100)	HFOI2 (x100)	HFOI1 (x100)	HFOI2 (x100)	HFOI1 (x100)	HFOI2 (x100)
LFOI1	0.134***	14.053***	0.007*	1.330***	0.014***	3.024***
	[21.14]	[22.16]	[1.75]	[4.09]	[15.50]	[40.13]
LFOI2	1.616***	155.028***	1.299***	106.612***	1.553***	104.776***
	[13.30]	[13.92]	[17.73]	[16.80]	[34.52]	[33.97]
LFOI3	-1.540***	-142.27***	-0.004	0.231	0.002**	0.210***
	[-10.63]	[-8.80]	[-0.24]	[0.14]	[2.05]	[2.60]
LFOI4	-3.603	-312.25	-1.857***	-147.11***	-0.141***	0.287
	[-1.34]	[-1.13]	[-4.56]	[-3.65]	[-2.98]	[0.06]
LFOI5	0.613***	34.535	0.131***	13.692***	-0.001**	-0.201***
	[2.90]	[1.51]	[6.36]	[7.41]	[-2.58]	[-10.25]
LFOI6	23.171***	2392.02***	2.578***	167.596***	0.048***	9.046***
	[6.44]	[6.81]	[10.53]	[6.89]	[3.32]	[6.74]
RET	0.189***	31.783***	-0.076**	42.839***	0.070*	78.984***
	[10.19]	[18.95]	[-2.20]	[17.32]	[1.73]	[30.00]
TURN	-0.112***	-9.370***	-0.023***	-0.645***	0.074***	4.137***
	[-41.03]	[-38.04]	[-15.57]	[-6.28]	[48.41]	[49.47]
BASPRD	-0.141***	-14.230***	0.047	-23.122***	-0.305**	59.581***
	[-16.69]	[-18.29]	[0.84]	[-6.18]	[-1.98]	[6.71]
ALPHA	0.006***	0.564***	0.002***	0.678***	-0.010***	0.458***
	[12.59]	[12.15]	[3.49]	[14.92]	[-19.32]	[13.49]
ADJRSQ	0.318***	0.301***	0.181***	0.178***	0.187***	0.188***
	[119.32]	[116.66]	[100.71]	[121.48]	[90.73]	[132.32]
OBS	4,134,495	4,125,901	4,160,764	4,160,736	4,162,225	4,162,223

Panel B. Liquidity (BID-ASK SPREAD)

	Illiquid		Less liquid		Liquid	
	HFOI1 (x100)	HFOI2 (x100)	HFOI1 (x100)	HFOI2 (x100)	HFOI1 (x100)	HFOI2 (x100)
LFOI1	0.089***	8.333***	0.031***	4.055***	0.019***	3.091***
	[26.44]	[23.87]	[22.94]	[29.10]	[17.66]	[35.45]
LFOI2	1.513***	162.810***	1.836***	154.085***	1.706***	120.082***
	[17.17]	[20.43]	[32.71]	[28.56]	[31.74]	[30.55]
LFOI3	-0.208***	-17.417***	0.009	1.572*	0.001	-0.219
	[-5.27]	[-3.32]	[1.16]	[1.90]	[0.33]	[-1.42]
LFOI4	0.200	-6.476	-1.228***	-121.68***	-0.436***	-40.211***
	[0.19]	[-0.06]	[-4.00]	[-3.95]	[-4.37]	[-5.09]
LFOI5	-0.015	-1.697	-0.008***	-0.892***	-0.002***	-0.110***
	[-0.54]	[-0.39]	[-2.97]	[-2.88]	[-3.47]	[-3.31]
LFOI6	4.592***	334.466***	0.762***	76.305***	0.172***	10.380***
	[5.53]	[3.54]	[6.31]	[6.12]	[6.72]	[4.71]
RET	0.245***	32.328***	0.241***	69.860***	0.061	78.067***
	[20.40]	[27.47]	[6.01]	[21.45]	[0.97]	[19.03]
TURN	-0.075***	-5.274***	-0.018***	-0.098	0.043***	1.308***
	[-29.29]	[-22.28]	[-12.36]	[-0.80]	[25.48]	[13.91]
BASPRD	-0.149***	-12.201***	-2.577***	-524.80***	-10.319***	-1089.9***
	[-18.63]	[-16.98]	[-5.47]	[-14.13]	[-10.05]	[-14.89]
ALPHA	0.006***	0.435***	0.013***	1.643***	0.008***	1.800***
	[12.52]	[9.07]	[11.98]	[17.07]	[8.92]	[29.58]
ADJRSQ	0.236***	0.224***	0.131***	0.144***	0.144***	0.138***
	[110.07]	[106.16]	[96.89]	[120.45]	[85.85]	[110.01]
OBS	4,133,359	4,124,415	4,161,783	4,161,765	4,160,270	4,160,267

Table. V (Continued)

Panel C. Exchange market

	NYSE and AMEX		NASDAQ	
	HFOI1 (x100)	HFOI2 (x100)	HFOI1 (x100)	HFOI2 (x100)
LFOI1	0.026***	4.369***	0.036***	3.780***
	[31.48]	[44.35]	[33.46]	[41.34]
LFOI4	2.023***	219.286***	1.764***	157.870***
	[50.28]	[52.31]	[45.77]	[51.25]
LFOI2	0.005***	0.244	-0.030***	-1.929***
	[3.72]	[1.64]	[-5.21]	[-3.93]
LFOI5	-0.503***	-72.062***	-0.489***	-57.997***
	[-7.24]	[-9.12]	[-2.61]	[-3.39]
LFOI3	-0.003***	-0.484***	-0.005***	-0.626***
	[-9.62]	[-13.92]	[-3.74]	[-4.96]
LFOI6	0.087***	30.316***	0.790***	45.654***
	[4.39]	[12.99]	[12.27]	[8.16]
RET	0.260***	65.052***	0.293***	59.936***
	[9.64]	[27.08]	[10.42]	[29.55]
TURN	0.032***	4.586***	-0.016***	-2.902***
	[30.47]	[40.17]	[-10.12]	[-37.10]
BASPRD	-0.217***	-63.605***	-0.229***	-32.164***
	[-14.96]	[-45.55]	[-21.96]	[-41.69]
ALPHA	-0.001***	0.710***	-0.000	0.930***
	[-3.12]	[18.72]	[-0.46]	[29.09]
ADJRSQ	0.143***	0.182***	0.105***	0.097***
	[94.73]	[120.18]	[88.79]	[105.98]
OBS	8,417,294	8,416,899	12,360,224	12,351,555

Panel D. Subperiod

	Early (1993-2000)		Midst (2001-2006)		Late (2007-2013)	
	HFOI1 (x100)	HFOI2 (x100)	HFOI1 (x100)	HFOI2 (x100)	HFOI1 (x100)	HFOI2 (x100)
LFOI1	0.063***	8.650***	0.015***	1.888***	0.004***	0.814***
	[44.03]	[67.45]	[15.29]	[23.69]	[8.67]	[19.62]
LFOI4	2.666***	211.515***	1.986***	130.623***	0.770***	101.933***
	[46.65]	[45.38]	[39.60]	[34.02]	[21.54]	[36.06]
LFOI2	0.000	1.199***	-0.002	-0.237***	-0.000	-0.069***
	[0.04]	[2.60]	[-1.40]	[-3.02]	[-0.55]	[-2.85]
LFOI5	-0.631***	-77.469***	-0.282***	-28.715***	-0.188***	6.508***
	[-2.60]	[-3.44]	[-4.40]	[-6.07]	[-8.79]	[3.51]
LFOI3	-0.013***	-1.398***	-0.001***	-0.100***	0.000***	-0.003
	[-11.05]	[-12.42]	[-4.17]	[-4.53]	[6.05]	[-0.48]
LFOI6	1.028***	88.768***	0.092***	15.292***	-0.080***	-3.077***
	[13.73]	[12.61]	[4.18]	[8.62]	[-13.29]	[-6.21]
RET	0.506***	65.242***	0.467***	85.760***	-0.082*	41.054***
	[13.98]	[40.73]	[10.11]	[21.86]	[-1.77]	[11.23]
TURN	-0.046***	-3.591***	0.074***	4.776***	-0.004***	-0.984***
	[-20.68]	[-29.38]	[37.32]	[43.89]	[-5.81]	[-16.14]
BASPRD	-0.326***	-37.222***	-0.352***	-49.814***	-0.215***	-36.516***
	[-46.35]	[-65.87]	[-18.12]	[-26.34]	[-12.54]	[-25.84]
ALPHA	0.008***	1.936***	-0.010***	0.412***	0.002***	0.406***
	[12.33]	[53.53]	[-13.14]	[8.28]	[2.93]	[8.79]
ADJRSQ	0.113***	0.099***	0.126***	0.111***	0.042***	0.082***
	[85.47]	[118.97]	[61.61]	[69.29]	[31.89]	[51.40]
OBS	10,028,176	10,028,120	5,390,639	5,390,570	5,358,703	5,349,764

Table VI
Relation between OI and Control Variable

This table shows the results of Fama-MacBeth regression of HFOIs and LFOIs on all of the control variables including lagged OI, liquidity factors, returns, and squared returns. The sample period is from 01 January 1993 to 31 December 2013. The first row represents which kind of OI is employed as a dependent variable in a regression model. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2013). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. TURN is daily turnover rate for each stock, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted bid-ask spread, the difference of bid and ask prices divided by the average of bid and ask prices. We only use bid-ask spread within the interval between 0 and 0.5. Also bid and ask prices are close price at each day. RET is daily stock returns. RET^2 is daily squared stock returns. ALPHA is the intercept of regression models. ADJRSQ is adjusted R-square. OBS is the number of observation in each regression model. Each subscript of variables indicates its time lag. In brackets, we report t-statistics of the average coefficient over sample period based on New-West (1987) standard errors. ***, **, and * indicate statistical significance at the 1, 5, and 10 per cent level, respectively.

Table. VI (Continued)

	HFOI1	HFOI2	LFOI1	LFOI2
OI_{t-1}	0.113***	0.090***	-0.009***	-0.025***
	[64.95]	[57.44]	[-10.67]	[-20.70]
OI_{t-2}	0.056***	0.054***	-0.007***	-0.015***
	[46.06]	[45.26]	[-8.84]	[-14.40]
OI_{t-3}	0.046***	0.045***	-0.003***	-0.006***
	[38.66]	[42.20]	[-3.56]	[-6.59]
OI_{t-4}	0.041***	0.042***	-0.001	-0.003***
	[34.51]	[40.92]	[-1.14]	[-3.00]
OI_{t-5}	0.040***	0.040***	0.000	-0.000
	[33.93]	[40.94]	[0.45]	[-0.42]
$BASPRD_t$	-0.001***	-0.096	0.289***	0.016***
	[-2.65]	[-1.47]	[6.25]	[11.61]
$TURN_t$	0.000**	0.022***	0.078***	0.005***
	[2.00]	[3.63]	[8.63]	[13.76]
RET_t	0.016***	1.992***	16.964***	0.709***
	[25.78]	[39.69]	[48.03]	[60.85]
RET_{t-1}	0.002***	0.257***	1.446***	0.074***
	[15.43]	[26.53]	[44.37]	[52.60]
RET_{t-2}	0.000	0.027***	0.448***	0.022***
	[1.49]	[4.67]	[21.94]	[25.60]
RET_{t-3}	-0.001	-0.061	0.169***	0.011***
	[-1.37]	[-1.48]	[3.03]	[13.60]
RET_{t-4}	-0.000***	-0.044***	0.143***	0.007***
	[-6.39]	[-11.22]	[8.54]	[8.37]
RET_{t-5}	-0.000***	-0.060***	0.098***	0.003***
	[-6.61]	[-15.31]	[6.38]	[5.08]
RET_t^2	-0.007***	-2.519***	-16.776***	-0.164***
	[-3.04]	[-14.96]	[-22.56]	[-10.89]
RET_{t-1}^2	-0.008***	-1.851***	-2.959***	-0.103***
	[-7.75]	[-22.69]	[-15.09]	[-17.80]
RET_{t-2}^2	-0.005***	-0.880***	-1.228***	-0.028***
	[-8.01]	[-14.69]	[-8.50]	[-6.87]
RET_{t-3}^2	-0.003***	-0.445***	-0.468***	-0.001
	[-5.16]	[-9.22]	[-3.38]	[-0.13]
RET_{t-4}^2	-0.002***	-0.412***	-0.549***	-0.009**
	[-4.50]	[-8.70]	[-4.40]	[-2.38]
RET_{t-5}^2	-0.003***	-0.357***	-0.139	0.008**
	[-4.70]	[-7.55]	[-1.24]	[2.47]
ALPHA	0.000	0.004	-0.002	-0.001***
	[0.14]	[1.25]	[-0.55]	[-8.64]
ADJRSQ	0.167***	0.158***	0.426***	0.607***
	[49.55]	[54.97]	[228.37]	[341.84]
OBS	19,752,776	19,728,024	19,956,999	19,956,999

Table VII
Return predictability

This table shows the results of Fama-MacBeth regression to measure return predictability of HFOIs and LFOIs. The sample period is from 01 January 1993 to 31 December 2013. The first row represents which kind of lagged OI uses in a regression model. The dependent variable is either raw returns or mid-quote return at time t. The second row represents which type of return is the dependent variable of a regression model. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2013). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. TURN is daily turnover rate for each stock, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted bid-ask spread, the difference of bid and ask prices divided by the average of bid and ask prices. We only use bid-ask spread within the interval between 0 and 0.5. Also bid and ask prices are close price at each day. RET is daily stock returns. RET^2 is daily squared stock returns. ALPHA is the intercept of regression models. ADJRSQ is adjusted R-square. OBS is the number of observation in each regression model. Panel A reports daily-based Fama-MacBeth regression for return predictability. Panel B reports weekly-based Fama-MacBeth regression. Each subscript of variables indicates its time lag. In brackets, we report t-statistics of the average coefficient over sample period based on New-West (1987) standard errors. ***, **, and * indicate statistical significance at the 1, 5, and 10 per cent level, respectively.

Table. VII (Continued)

Panel A. Daily-based regression

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
OI_{t-1}	12.747***	4.063***	0.172***	0.060***	0.332***	0.122***	11.942***	3.388***
	[11.27]	[8.17]	[19.89]	[12.38]	[24.87]	[31.18]	[25.92]	[33.09]
OI_{t-2}	-3.349***	-4.045***	-0.028***	-0.036***	-0.003	-0.028***	-0.208**	-1.328***
	[-7.22]	[-10.71]	[-5.10]	[-8.93]	[-0.87]	[-11.51]	[-2.06]	[-15.15]
OI_{t-3}	-2.873***	-1.534***	-0.039***	-0.024***	-0.023***	-0.008***	-0.883***	-0.437***
	[-6.52]	[-4.47]	[-8.06]	[-6.36]	[-8.35]	[-3.66]	[-9.22]	[-6.05]
OI_{t-4}	-2.858***	-1.558***	-0.029***	-0.015***	-0.023***	-0.003	-0.698***	-0.025
	[-6.83]	[-4.56]	[-6.10]	[-3.86]	[-8.86]	[-1.44]	[-8.02]	[-0.39]
OI_{t-5}	-2.508***	-1.343***	-0.031***	-0.017***	-0.016***	0.001	-0.535***	0.068
	[-6.02]	[-3.88]	[-6.42]	[-4.56]	[-6.32]	[0.39]	[-6.55]	[1.06]
$BASPRD_{t-1}$	2.635***	0.949***	2.751***	0.997***	2.313***	0.783***	2.291***	0.820***
	[18.09]	[7.58]	[18.14]	[7.84]	[16.31]	[6.33]	[16.17]	[6.64]
$TURN_{t-1}$	0.107***	0.062***	0.097***	0.056***	0.091***	0.068***	0.070***	0.075***
	[15.68]	[12.61]	[14.41]	[11.38]	[16.88]	[14.67]	[15.12]	[16.17]
RET_{t-1}	-7.237***	0.231	-7.258***	0.255	-10.405***	-1.390***	-12.547***	-1.965***
	[-28.66]	[1.38]	[-28.47]	[1.53]	[-34.10]	[-8.17]	[-37.30]	[-9.77]
RET_{t-2}	-1.870***	-1.495***	-1.891***	-1.495***	-2.142***	-1.265***	-2.252***	-0.928***
	[-19.72]	[-16.94]	[-19.90]	[-16.78]	[-20.54]	[-15.46]	[-19.85]	[-10.70]
RET_{t-3}	-0.880***	-1.111***	-0.864***	-1.095***	-0.725***	-1.014***	-0.594***	-0.897***
	[-9.29]	[-14.55]	[-9.07]	[-14.25]	[-8.01]	[-12.25]	[-6.27]	[-10.59]
RET_{t-4}	-0.610***	-0.859***	-0.595***	-0.848***	-0.312***	-0.882***	-0.222**	-0.908***
	[-7.17]	[-10.80]	[-6.98]	[-10.56]	[-3.47]	[-11.79]	[-2.37]	[-11.69]
RET_{t-5}	-0.443***	-0.611***	-0.430***	-0.596***	-0.204**	-0.683***	-0.104	-0.697***
	[-4.04]	[-7.53]	[-3.89]	[-7.28]	[-2.41]	[-8.99]	[-1.17]	[-8.85]
RET^2_{t-1}	22.020***	26.356***	22.255***	26.342***	19.862***	26.918***	17.551***	23.269***
	[11.88]	[23.02]	[12.03]	[23.11]	[6.85]	[23.58]	[6.09]	[21.16]
RET^2_{t-2}	2.692**	4.432***	2.799**	4.493***	2.472**	3.061***	3.261***	5.684***
	[2.36]	[5.26]	[2.46]	[5.33]	[2.52]	[3.68]	[3.33]	[6.80]
RET^2_{t-3}	3.549***	6.398***	3.613***	6.441***	0.998	4.690***	1.875	6.559***
	[6.21]	[7.82]	[6.36]	[7.85]	[0.50]	[4.39]	[0.93]	[6.09]
RET^2_{t-4}	3.924***	7.050***	3.924***	6.682***	7.821*	8.080***	8.486**	9.606***
	[8.29]	[8.90]	[8.34]	[8.49]	[1.87]	[4.72]	[2.02]	[5.60]
RET^2_{t-5}	3.573***	6.800***	3.595***	6.727***	3.291***	6.503***	3.930***	7.516***
	[6.89]	[8.20]	[6.91]	[8.13]	[6.49]	[7.92]	[7.81]	[9.21]
ALPHA	-0.075***	-0.034***	-0.073***	-0.032***	-0.060***	-0.024***	-0.058***	-0.027***
	[-16.75]	[-9.36]	[-16.37]	[-8.82]	[-13.70]	[-6.40]	[-13.56]	[-7.39]
ADJRSQ	0.034***	0.026***	0.033***	0.026***	0.037***	0.027***	0.039***	0.029***
	[48.77]	[48.38]	[47.54]	[48.00]	[50.92]	[48.07]	[52.73]	[47.71]
OBS	20,544,907	19,875,273	20,521,945	19,852,495	20,736,732	20,064,705	20,736,732	20,064,705

Table. VII (Continued)

Panel B. Weekly-based regression

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
OI_{t-1}	0.015	-0.004	0.000**	-0.000**	0.001***	0.000***	0.022***	0.014***
	[1.54]	[-0.40]	[2.18]	[-2.01]	[10.41]	[7.55]	[16.11]	[13.38]
OI_{t-2}	-0.023***	-0.013	-0.000***	-0.000***	-0.000**	0.000***	0.001	0.008***
	[-2.90]	[-1.55]	[-4.17]	[-3.52]	[-2.41]	[4.72]	[0.73]	[7.49]
OI_{t-3}	-0.019**	-0.017**	-0.000***	-0.000***	-0.000***	0.000***	-0.001	0.006***
	[-2.51]	[-2.54]	[-4.80]	[-4.35]	[-2.91]	[3.08]	[-1.23]	[5.88]
OI_{t-4}	-0.014**	-0.014**	-0.000***	-0.000***	-0.000***	0.000	-0.002*	0.002***
	[-2.22]	[-2.17]	[-4.08]	[-4.58]	[-4.04]	[1.12]	[-1.74]	[2.98]
$BASPRD_{t-1}$	0.069***	0.037***	0.068***	0.036***	0.066***	0.035***	0.067***	0.036***
	[8.16]	[4.99]	[8.16]	[4.93]	[7.93]	[4.77]	[7.97]	[4.82]
$TURN_{t-1}$	-0.000	0.000	-0.000	0.000	-0.000	0.000	-0.000***	-0.000
	[-1.27]	[1.29]	[-1.50]	[0.92]	[-0.43]	[1.56]	[-2.91]	[-0.91]
RET_{t-1}	-0.041***	-0.034***	-0.041***	-0.034***	-0.048***	-0.041***	-0.059***	-0.048***
	[-16.58]	[-15.45]	[-16.74]	[-15.50]	[-18.94]	[-18.77]	[-22.80]	[-21.43]
RET_{t-2}	-0.007***	-0.015***	-0.006***	-0.015***	-0.006**	-0.019***	-0.009***	-0.023***
	[-3.21]	[-7.38]	[-2.90]	[-7.26]	[-2.56]	[-9.08]	[-3.74]	[-10.29]
RET_{t-3}	-0.002	-0.010***	-0.001	-0.010***	-0.001	-0.013***	-0.002	-0.016***
	[-1.02]	[-5.76]	[-0.60]	[-5.40]	[-0.39]	[-7.32]	[-0.86]	[-8.54]
RET_{t-4}	0.001	-0.005***	0.002	-0.005***	0.003	-0.006***	0.002	-0.008***
	[0.64]	[-2.99]	[0.97]	[-2.65]	[1.46]	[-3.42]	[0.82]	[-4.08]
RET^2_{t-1}	0.080***	0.101***	0.078***	0.097***	0.079***	0.099***	0.081***	0.099***
	[12.37]	[9.43]	[11.90]	[9.02]	[12.45]	[9.29]	[13.00]	[9.29]
RET^2_{t-2}	0.053***	0.080***	0.051***	0.077***	0.050***	0.080***	0.053***	0.079***
	[7.42]	[7.04]	[7.28]	[6.72]	[7.18]	[7.24]	[7.59]	[6.91]
RET^2_{t-3}	0.047***	0.062***	0.046***	0.060***	0.045***	0.063***	0.048***	0.060***
	[5.90]	[6.51]	[5.66]	[6.30]	[5.86]	[6.63]	[6.23]	[6.25]
RET^2_{t-4}	0.033***	0.071***	0.031***	0.067***	0.032***	0.071***	0.033***	0.070***
	[6.07]	[6.85]	[5.78]	[6.48]	[5.78]	[6.75]	[6.10]	[6.78]
ALPHA	-0.001***	0.000	-0.001***	0.000	-0.000**	0.000	-0.000**	0.000
	[-2.86]	[0.49]	[-2.60]	[0.96]	[-2.14]	[1.29]	[-2.35]	[1.07]
ADJRSQ	0.025***	0.024***	0.025***	0.024***	0.025***	0.025***	0.027***	0.026***
	[23.88]	[22.82]	[23.73]	[22.88]	[24.11]	[22.71]	[24.41]	[23.12]
OBS	4,385,256	3,985,225	4,385,256	3,985,225	4,385,256	3,985,225	4,385,256	3,985,225

Table VIII
Return predictability in size subsamples

This table shows the results of Fama-MacBeth regression to measure return predictability of HFOIs and LFOIs within each size subsamples. We separate whole sample dataset into five subsamples based on market capitalization. In this tabale, we report test results of only three subsample regression. Panel A is for the smallest size firms. Panel B is for medium size firms. Panel C is for the largest size firms. The sample period is from 01 January 1993 to 31 December 2012. The first row represents which kind of lagged OI uses in a regression model. The dependent variable is either raw returns or mid-quote return at time t. The second row represents which type of return is the dependent variable of a regression model. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2011). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. TURN is daily turnover rate for each stock, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted bid-ask spread, the difference of bid and ask prices divided by the average of bid and ask prices. We only use bid-ask spread within the interval between 0 and 0.5. Also bid and ask prices are close price at each day. RET is daily stock returns. RET^2 is daily squared stock returns. ALPHA is the intercept of regression models. ADJRSQ is adjusted R-square. OBS is the number of observation in each regression model. Each subscript of variables indicates its time lag. In brackets, we report t-statistics of the average coefficient over sample period based on New-West (1987) standard errors. ***, **, and * indicate statistical significance at the 1, 5, and 10 per cent level, respectively.

Table. VIII (Continued)

Panel A. Small size

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
OI_{t-1}	17.278***	6.892***	0.222***	0.088***	0.517***	0.218***	16.938***	5.523***
	[12.21]	[10.76]	[22.05]	[14.26]	[31.23]	[37.01]	[32.82]	[38.03]
OI_{t-2}	-4.284***	-5.386***	-0.029***	-0.041***	0.026***	-0.025***	0.438***	-1.504***
	[-6.29]	[-10.25]	[-3.83]	[-7.81]	[5.03]	[-6.88]	[2.60]	[-12.32]
OI_{t-3}	-3.990***	-1.981***	-0.048***	-0.032***	-0.033***	-0.012***	-1.249***	-0.738***
	[-6.30]	[-4.03]	[-6.99]	[-6.44]	[-7.43]	[-3.73]	[-9.33]	[-7.43]
OI_{t-4}	-3.431***	-1.615***	-0.044***	-0.023***	-0.034***	-0.000	-1.052***	-0.035
	[-5.86]	[-3.59]	[-6.84]	[-4.74]	[-8.31]	[-0.16]	[-8.53]	[-0.40]
OI_{t-5}	-3.756***	-1.738***	-0.047***	-0.027***	-0.026***	0.003	-0.836***	0.048
	[-6.12]	[-3.52]	[-7.09]	[-5.26]	[-6.64]	[0.83]	[-7.14]	[0.53]
$BASPRD_{t-1}$	2.983***	0.845***	3.079***	0.882***	2.570***	0.628***	2.531***	0.639***
	[20.84]	[7.03]	[20.77]	[7.22]	[18.53]	[5.32]	[18.30]	[5.44]
$TURN_{t-1}$	0.139***	0.069***	0.127***	0.062***	0.116***	0.081***	0.093***	0.096***
	[14.25]	[10.43]	[13.51]	[9.55]	[15.53]	[13.08]	[13.32]	[15.01]
RET_{t-1}	-9.958***	0.009	-9.951***	0.066	-13.738***	-2.232***	-16.149***	-2.829***
	[-33.50]	[0.04]	[-33.31]	[0.33]	[-40.66]	[-10.43]	[-43.57]	[-11.48]
RET_{t-2}	-2.371***	-1.321***	-2.377***	-1.308***	-2.953***	-1.310***	-3.164***	-0.882***
	[-21.09]	[-17.14]	[-21.26]	[-16.84]	[-24.50]	[-16.39]	[-23.24]	[-9.70]
RET_{t-3}	-1.083***	-1.228***	-1.070***	-1.211***	-0.946***	-1.178***	-0.748***	-0.939***
	[-12.79]	[-16.80]	[-12.33]	[-16.48]	[-11.75]	[-15.67]	[-8.44]	[-11.89]
RET_{t-4}	-0.454***	-0.915***	-0.412***	-0.887***	-0.301***	-0.959***	-0.148	-0.952***
	[-5.38]	[-12.50]	[-4.85]	[-12.05]	[-3.64]	[-12.63]	[-1.62]	[-11.86]
RET_{t-5}	-0.333***	-0.686***	-0.303***	-0.657***	-0.223**	-0.768***	-0.094	-0.764***
	[-4.68]	[-9.87]	[-4.19]	[-9.34]	[-2.31]	[-10.37]	[-0.93]	[-9.89]
RET^2_{t-1}	27.037***	32.829***	27.445***	32.884***	29.455***	33.405***	26.147***	28.137***
	[24.48]	[25.97]	[24.71]	[25.94]	[25.92]	[25.31]	[24.88]	[22.43]
RET^2_{t-2}	4.050***	2.974***	4.281***	3.154***	4.172***	2.323**	5.173***	5.338***
	[6.11]	[3.04]	[6.46]	[3.19]	[6.28]	[2.36]	[7.79]	[5.39]
RET^2_{t-3}	3.781***	6.313***	3.800***	6.466***	3.819***	6.137***	4.805***	8.051***
	[6.12]	[7.01]	[6.17]	[7.14]	[6.24]	[6.84]	[8.02]	[8.90]
RET^2_{t-4}	4.352***	7.279***	4.397***	6.951***	4.068***	6.999***	5.146***	8.955***
	[7.26]	[8.10]	[7.19]	[7.79]	[6.94]	[7.79]	[8.73]	[9.82]
RET^2_{t-5}	3.739***	6.549***	3.641***	6.571***	3.585***	6.624***	4.487***	7.745***
	[5.93]	[7.27]	[5.70]	[7.23]	[5.88]	[7.42]	[7.40]	[8.61]
ALPHA	-0.116***	-0.043***	-0.114***	-0.042***	-0.095***	-0.029***	-0.092***	-0.031***
	[-19.24]	[-9.25]	[-19.05]	[-8.93]	[-16.25]	[-6.25]	[-16.01]	[-6.73]
ADJRSQ	0.042***	0.027***	0.041***	0.026***	0.045***	0.028***	0.049***	0.030***
	[48.28]	[58.21]	[47.01]	[57.51]	[51.88]	[59.35]	[54.17]	[58.85]
OBS	11,049,644	10,757,899	11,026,743	10,735,182	11,229,784	10,935,508	11,229,784	10,935,508

Table. VIII (Continued)

Panel B. Medium size

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
OI_{t-1}	0.269	-2.628	0.028*	-0.008	0.099***	0.081***	2.658***	1.796***
	[0.19]	[-1.53]	[1.78]	[-0.55]	[16.05]	[13.59]	[11.63]	[8.49]
OI_{t-2}	-2.954**	-4.762***	-0.016	-0.044*	-0.012**	-0.012**	-1.034***	-0.977***
	[-2.43]	[-3.04]	[-1.04]	[-1.95]	[-2.46]	[-2.37]	[-5.65]	[-5.14]
OI_{t-3}	-2.015*	-1.682	-0.015	-0.004	-0.003	-0.006	-0.092	-0.170
	[-1.89]	[-1.29]	[-1.04]	[-0.24]	[-0.74]	[-1.29]	[-0.56]	[-0.90]
OI_{t-4}	-4.915***	-5.107***	-0.002	0.010	0.002	0.001	0.031	-0.003
	[-4.00]	[-3.20]	[-0.13]	[0.64]	[0.37]	[0.30]	[0.19]	[-0.02]
OI_{t-5}	-0.729	-0.009	-0.022	-0.032*	0.009*	0.008*	0.049	0.128
	[-0.55]	[-0.01]	[-1.45]	[-1.80]	[1.88]	[1.70]	[0.31]	[0.79]
$BASPRD_{t-1}$	10.257***	10.320***	10.418***	10.351***	10.130***	10.283***	10.181***	10.216***
	[7.45]	[8.39]	[7.57]	[8.44]	[7.33]	[8.29]	[7.44]	[8.36]
$TURN_{t-1}$	0.061***	0.062***	0.050***	0.055***	0.064***	0.061***	0.056***	0.059***
	[8.59]	[8.97]	[7.13]	[7.75]	[9.52]	[9.52]	[8.16]	[7.82]
RET_{t-1}	-1.043***	0.565***	-1.130***	0.580***	-2.415***	-0.627***	-2.588***	-0.614**
	[-6.04]	[3.09]	[-6.67]	[3.07]	[-12.32]	[-3.18]	[-11.14]	[-2.48]
RET_{t-2}	-1.454***	-1.539***	-1.536***	-1.564***	-1.204***	-1.388***	-0.904***	-0.977***
	[-10.12]	[-9.92]	[-10.42]	[-10.16]	[-6.95]	[-7.60]	[-4.71]	[-4.91]
RET_{t-3}	-0.823***	-0.812***	-0.833***	-0.797***	-0.720***	-0.619***	-0.752***	-0.726***
	[-5.85]	[-5.86]	[-5.86]	[-5.61]	[-4.40]	[-3.47]	[-4.13]	[-3.72]
RET_{t-4}	-0.783***	-0.788***	-0.856***	-0.890***	-0.862***	-0.902***	-0.833***	-0.792***
	[-5.64]	[-5.54]	[-6.07]	[-6.52]	[-5.19]	[-5.48]	[-4.71]	[-4.21]
RET_{t-5}	-0.553***	-0.622***	-0.555***	-0.627***	-0.797***	-0.856***	-0.595***	-0.745***
	[-4.10]	[-4.56]	[-4.08]	[-4.67]	[-4.84]	[-4.93]	[-3.28]	[-4.37]
RET^2_{t-1}	11.606***	6.456*	13.102***	7.783**	10.569***	8.817*	9.389***	4.068
	[3.91]	[1.87]	[4.37]	[2.23]	[3.56]	[1.91]	[3.33]	[1.18]
RET^2_{t-2}	4.859*	4.129	4.102	5.542*	1.953	3.578	4.312*	5.672*
	[1.90]	[1.40]	[1.52]	[1.90]	[0.77]	[1.25]	[1.74]	[1.79]
RET^2_{t-3}	4.088	7.260**	3.853	4.789*	2.843	4.337	5.122*	4.411
	[1.59]	[2.15]	[1.59]	[1.90]	[1.09]	[1.53]	[1.85]	[1.60]
RET^2_{t-4}	5.654**	7.325***	5.413**	3.281	4.547*	3.116	4.767**	4.396
	[2.30]	[2.62]	[2.20]	[1.09]	[1.86]	[1.02]	[1.98]	[1.56]
RET^2_{t-5}	5.852**	8.924***	6.149***	8.504***	5.453**	7.022**	7.567***	9.376***
	[2.55]	[3.28]	[2.60]	[3.16]	[2.36]	[2.41]	[3.18]	[3.15]
ALPHA	-0.054***	-0.051***	-0.050***	-0.046***	-0.042***	-0.042***	-0.047***	-0.044***
	[-8.86]	[-8.67]	[-8.37]	[-8.07]	[-6.49]	[-6.78]	[-7.48]	[-7.40]
ADJRSQ	0.070***	0.070***	0.069***	0.069***	0.073***	0.073***	0.078***	0.078***
	[39.22]	[44.26]	[39.19]	[43.83]	[40.70]	[42.49]	[41.43]	[45.38]
OBS	2,368,127	2,277,680	2,368,124	2,277,677	2,370,046	2,279,466	2,370,046	2,279,466

Table. VIII (Continued)

Panel C. Big size

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
OI_{t-1}	2.504	4.853	0.102**	0.175***	0.061***	0.083***	1.972***	2.657***
	[0.73]	[1.49]	[2.37]	[2.59]	[7.65]	[8.82]	[4.21]	[5.71]
OI_{t-2}	-3.239	-4.647	-0.069	-0.107**	-0.013*	-0.019**	-0.391	-0.101
	[-1.05]	[-1.06]	[-1.64]	[-2.09]	[-1.94]	[-2.18]	[-1.25]	[-0.24]
OI_{t-3}	0.428	-0.817	-0.003	-0.032	-0.007	-0.010	0.571	-0.247
	[0.15]	[-0.17]	[-0.07]	[-0.69]	[-1.14]	[-1.38]	[1.51]	[-0.66]
OI_{t-4}	2.546	0.583	0.040	0.049	-0.005	-0.007	0.148	0.017
	[0.63]	[0.19]	[1.05]	[1.02]	[-0.74]	[-1.08]	[0.49]	[0.05]
OI_{t-5}	-3.386	-2.456	0.017	0.002	-0.002	-0.012*	0.261	-0.038
	[-0.98]	[-0.63]	[0.40]	[0.06]	[-0.42]	[-1.68]	[0.89]	[-0.12]
$BASPRD_{t-1}$	10.378***	10.823***	10.506***	11.364***	11.185***	12.376***	10.740***	11.966***
	[3.23]	[3.51]	[3.33]	[3.68]	[3.59]	[4.01]	[3.47]	[3.92]
$TURN_{t-1}$	0.063***	0.067***	0.052***	0.055***	0.061***	0.065***	0.058***	0.069***
	[5.04]	[4.48]	[4.21]	[3.44]	[5.11]	[4.79]	[4.92]	[5.14]
RET_{t-1}	-1.868***	-1.606***	-2.021***	-1.810***	-2.507***	-2.614***	-2.614***	-2.807***
	[-9.38]	[-7.68]	[-10.19]	[-8.35]	[-10.32]	[-10.20]	[-8.80]	[-10.05]
RET_{t-2}	-2.148***	-2.222***	-2.151***	-2.176***	-1.813***	-1.683***	-1.960***	-2.147***
	[-11.16]	[-11.04]	[-10.66]	[-10.41]	[-7.94]	[-5.58]	[-7.97]	[-7.70]
RET_{t-3}	-1.233***	-1.223***	-1.270***	-1.316***	-1.253***	-1.437***	-1.626***	-1.430***
	[-7.16]	[-6.95]	[-7.21]	[-7.34]	[-6.08]	[-6.93]	[-6.88]	[-5.67]
RET_{t-4}	-0.894***	-0.989***	-0.953***	-1.084***	-0.825***	-0.926***	-1.135***	-1.167***
	[-5.18]	[-5.27]	[-5.43]	[-5.42]	[-3.92]	[-4.14]	[-5.05]	[-4.81]
RET_{t-5}	-0.811***	-0.868***	-0.887***	-0.856***	-0.744***	-0.574***	-0.865***	-0.581**
	[-4.97]	[-4.96]	[-5.25]	[-4.57]	[-3.78]	[-2.74]	[-3.90]	[-2.28]
RET_{t-1}^2	-15.252***	-27.297***	-13.317***	-30.069***	-17.186***	-29.712***	-17.691***	-32.318***
	[-2.91]	[-4.04]	[-2.60]	[-4.16]	[-3.12]	[-4.96]	[-3.27]	[-5.11]
RET_{t-2}^2	2.229	5.760	-1.568	3.716	-4.606	-2.555	-1.442	-0.717
	[0.47]	[0.79]	[-0.33]	[0.67]	[-1.02]	[-0.45]	[-0.31]	[-0.12]
RET_{t-3}^2	6.234	1.488	5.087	3.547	0.286	-1.233	4.970	11.542
	[1.33]	[0.23]	[1.13]	[0.70]	[0.06]	[-0.25]	[1.03]	[1.55]
RET_{t-4}^2	1.184	3.445	-0.889	0.639	-0.784	0.418	-1.340	-2.596
	[0.28]	[0.65]	[-0.22]	[0.11]	[-0.18]	[0.06]	[-0.30]	[-0.42]
RET_{t-5}^2	-1.990	-7.350	-2.923	-4.962	-2.328	-5.943	-2.791	-6.781
	[-0.45]	[-1.47]	[-0.68]	[-1.01]	[-0.54]	[-1.25]	[-0.60]	[-1.12]
ALPHA	-0.035***	-0.036***	-0.034***	-0.035***	-0.031***	-0.035***	-0.035***	-0.037***
	[-5.50]	[-5.18]	[-4.95]	[-5.16]	[-4.50]	[-4.56]	[-5.54]	[-4.65]
ADJRSQ	0.104***	0.107***	0.102***	0.105***	0.107***	0.110***	0.116***	0.119***
	[51.48]	[50.25]	[52.68]	[49.82]	[54.90]	[52.45]	[56.64]	[53.84]
OBS	1,707,404	1,630,103	1,707,404	1,630,103	1,708,250	1,630,923	1,708,250	1,630,923

Table IX
Return predictability in liquidity subsamples

This table shows the results of Fama-MacBeth regression to measure return predictability of HFOIs and LFOIs within each size subsamples. We separate whole sample dataset into five subsamples based on bid-ask spread. In this tabale, we report test results of only three subsample regressions. Panel A is for stocks with the narrowest bid-ask spread. Panel B is for stocks with medium bid-ask spread. Panel C is for stocks with the widest bid-ask spread. The sample period is from 01 January 1993 to 31 December 2013. The first row represents which kind of lagged OI uses in a regression model. The dependent variable is either raw returns or mid-quote return at time t. The second row represents which type of return is the dependent variable of a regression model. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2011). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. TURN is daily turnover rate for each stock, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted bid-ask spread, the difference of bid and ask prices divided by the average of bid and ask prices. We only use bid-ask spread within the interval between 0 and 0.5. Also bid and ask prices are close price at each day. RET is daily stock returns. RET^2 is daily squared stock returns. ALPHA is the intercept of regression models. ADJRSQ is adjusted R-square. OBS is the number of observation in each regression model. Each subscript of variables indicates its time lag. In brackets, we report t-statistics of the average coefficient over sample period based on New-West (1987) standard errors. ***, **, and * indicate statistical significance at the 1, 5, and 10 per cent level, respectively.

Table. IX (Continued)

Panel A. Liquid stock (Narrow bid-ask spread)

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
OI_{t-1}	-0.119	-0.104	0.044***	0.028***	0.062***	0.065***	1.706***	1.416***
	[-0.13]	[-0.13]	[3.97]	[2.93]	[13.53]	[15.10]	[9.74]	[9.02]
OI_{t-2}	-2.747***	-2.467***	-0.029***	-0.030***	-0.031***	-0.030***	-1.298***	-1.171***
	[-3.68]	[-3.60]	[-2.99]	[-3.49]	[-8.07]	[-8.81]	[-8.90]	[-9.02]
OI_{t-3}	-2.004***	-1.945***	-0.024***	-0.021**	-0.013***	-0.009***	-0.550***	-0.386***
	[-2.60]	[-2.74]	[-2.62]	[-2.51]	[-3.62]	[-2.72]	[-4.05]	[-3.07]
OI_{t-4}	-1.623**	-2.262***	-0.014	-0.016*	-0.004	-0.002	-0.034	0.026
	[-2.15]	[-3.17]	[-1.48]	[-1.89]	[-1.17]	[-0.80]	[-0.27]	[0.21]
OI_{t-5}	-0.963	-0.737	0.002	0.001	-0.002	-0.002	-0.057	-0.019
	[-1.31]	[-1.06]	[0.21]	[0.15]	[-0.74]	[-0.67]	[-0.46]	[-0.17]
$BASPRD_{t-1}$	15.879***	14.059***	16.368***	14.312***	15.845***	13.988***	15.974***	14.088***
	[6.67]	[7.08]	[6.86]	[7.19]	[6.71]	[7.12]	[6.81]	[7.21]
$TURN_{t-1}$	0.064***	0.062***	0.055***	0.055***	0.064***	0.059***	0.066***	0.063***
	[10.29]	[11.09]	[8.72]	[9.64]	[10.62]	[11.19]	[11.52]	[12.12]
RET_{t-1}	0.337**	1.397***	0.308*	1.388***	-0.487***	0.423**	-0.749***	0.468**
	[2.06]	[7.94]	[1.89]	[7.92]	[-2.71]	[2.37]	[-3.53]	[2.18]
RET_{t-2}	-1.380***	-1.487***	-1.384***	-1.490***	-0.811***	-0.900***	-0.539***	-0.681***
	[-10.88]	[-11.65]	[-10.68]	[-11.44]	[-5.50]	[-6.31]	[-3.24]	[-4.23]
RET_{t-3}	-0.825***	-0.909***	-0.819***	-0.921***	-0.578***	-0.716***	-0.414***	-0.667***
	[-6.79]	[-7.56]	[-6.69]	[-7.58]	[-4.15]	[-5.36]	[-2.61]	[-4.47]
RET_{t-4}	-0.851***	-0.898***	-0.822***	-0.861***	-0.856***	-0.929***	-0.823***	-0.918***
	[-6.99]	[-7.52]	[-6.68]	[-7.12]	[-6.10]	[-6.89]	[-5.31]	[-6.12]
RET_{t-5}	-0.429***	-0.491***	-0.444***	-0.508***	-0.415***	-0.479***	-0.329**	-0.419***
	[-3.75]	[-4.24]	[-3.86]	[-4.36]	[-3.19]	[-3.75]	[-2.20]	[-2.88]
RET^2_{t-1}	5.482**	1.993	6.170**	2.447	5.977**	2.928	4.661*	0.786
	[2.20]	[0.72]	[2.50]	[0.90]	[2.31]	[1.04]	[1.87]	[0.29]
RET^2_{t-2}	5.299***	4.315**	4.616**	3.713*	3.203*	1.393	5.068***	4.763**
	[2.87]	[2.11]	[2.49]	[1.79]	[1.67]	[0.66]	[2.70]	[2.28]
RET^2_{t-3}	-0.090	1.563	-0.205	1.674	-0.940	0.101	0.661	2.404
	[-0.05]	[0.82]	[-0.11]	[0.87]	[-0.51]	[0.05]	[0.36]	[1.25]
RET^2_{t-4}	5.495***	6.240***	5.363***	5.772***	4.546**	4.577**	6.301***	7.168***
	[3.00]	[3.20]	[2.94]	[2.95]	[2.49]	[2.39]	[3.33]	[3.66]
RET^2_{t-5}	-0.336	0.870	-0.836	0.877	-0.744	0.714	0.130	1.696
	[-0.19]	[0.43]	[-0.46]	[0.44]	[-0.42]	[0.36]	[0.07]	[0.87]
ALPHA	-0.091***	-0.075***	-0.089***	-0.073***	-0.082***	-0.066***	-0.084***	-0.068***
	[-14.78]	[-14.30]	[-14.36]	[-13.70]	[-12.90]	[-11.91]	[-13.59]	[-12.74]
ADJRSQ	0.061***	0.060***	0.060***	0.059***	0.062***	0.061***	0.067***	0.066***
	[53.97]	[59.01]	[52.98]	[57.76]	[55.49]	[59.49]	[56.25]	[60.16]
OBS	4,139,737	3,995,098	4,139,719	3,995,080	4,145,669	4,000,911	4,145,669	4,000,911

Table. IX (Continued)

Panel B. Less liquid stock

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
OI_{t-1}	8.301***	4.304***	0.090***	0.039***	0.175***	0.110***	5.600***	2.081***
	[6.59]	[4.43]	[7.52]	[4.10]	[20.14]	[18.79]	[17.35]	[12.19]
OI_{t-2}	-3.379***	-4.132***	-0.021**	-0.029***	-0.033***	-0.031***	-1.877***	-1.813***
	[-3.59]	[-5.29]	[-2.06]	[-3.55]	[-7.40]	[-7.53]	[-11.23]	[-12.17]
OI_{t-3}	-1.750*	-1.233	-0.015	-0.017**	-0.016***	-0.010***	-0.951***	-0.733***
	[-1.95]	[-1.61]	[-1.51]	[-2.02]	[-3.80]	[-2.68]	[-6.54]	[-5.81]
OI_{t-4}	-1.791**	-0.300	-0.024**	-0.016*	-0.009**	-0.001	-0.281**	-0.018
	[-2.11]	[-0.39]	[-2.54]	[-1.92]	[-2.40]	[-0.34]	[-2.14]	[-0.15]
OI_{t-5}	-3.067***	-1.931**	-0.018**	-0.010	-0.002	0.001	-0.154	0.061
	[-3.76]	[-2.53]	[-2.03]	[-1.33]	[-0.43]	[0.26]	[-1.11]	[0.49]
$BASPRD_{t-1}$	17.538***	13.624***	17.558***	13.683***	17.027***	13.162***	17.254***	13.331***
	[14.53]	[14.90]	[14.54]	[14.97]	[13.99]	[14.40]	[14.15]	[14.59]
$TURN_{t-1}$	0.088***	0.076***	0.079***	0.069***	0.090***	0.087***	0.094***	0.102***
	[12.10]	[12.23]	[11.01]	[11.30]	[12.61]	[13.62]	[13.07]	[14.96]
RET_{t-1}	-2.784***	1.317***	-2.808***	1.330***	-4.406***	0.179	-5.190***	0.250
	[-15.99]	[7.95]	[-16.21]	[8.04]	[-20.10]	[1.08]	[-20.08]	[1.29]
RET_{t-2}	-1.195***	-1.158***	-1.215***	-1.154***	-0.881***	-0.819***	-0.393***	-0.209*
	[-10.71]	[-10.86]	[-10.96]	[-10.71]	[-7.13]	[-7.05]	[-2.88]	[-1.66]
RET_{t-3}	-0.891***	-0.965***	-0.885***	-0.951***	-0.704***	-0.817***	-0.345***	-0.488***
	[-8.63]	[-9.41]	[-8.59]	[-9.21]	[-6.26]	[-7.29]	[-2.75]	[-3.90]
RET_{t-4}	-0.753***	-0.860***	-0.740***	-0.839***	-0.656***	-0.821***	-0.644***	-0.852***
	[-7.45]	[-8.46]	[-7.32]	[-8.16]	[-5.68]	[-7.32]	[-4.95]	[-6.90]
RET_{t-5}	-0.578***	-0.594***	-0.571***	-0.582***	-0.634***	-0.708***	-0.523***	-0.678***
	[-5.70]	[-6.19]	[-5.50]	[-6.00]	[-5.23]	[-6.61]	[-3.83]	[-5.68]
RET^2_{t-1}	19.415***	17.012***	19.745***	17.024***	19.794***	16.359***	16.738***	13.714***
	[11.16]	[8.74]	[11.36]	[8.83]	[10.84]	[8.14]	[9.75]	[7.05]
RET^2_{t-2}	4.837***	2.513	4.889***	2.304	4.151***	0.816	6.463***	4.448***
	[3.55]	[1.54]	[3.56]	[1.40]	[2.98]	[0.49]	[4.60]	[2.70]
RET^2_{t-3}	3.899***	4.598***	4.072***	4.663***	2.673**	3.472**	4.328***	5.822***
	[3.32]	[3.16]	[3.45]	[3.18]	[2.28]	[2.38]	[3.61]	[3.96]
RET^2_{t-4}	2.939***	6.527***	2.827**	6.227***	2.840**	5.823***	3.941***	7.926***
	[2.64]	[4.50]	[2.52]	[4.26]	[2.54]	[4.03]	[3.48]	[5.43]
RET^2_{t-5}	1.441	3.501**	1.626	3.630**	1.405	3.589**	2.409**	4.627***
	[1.26]	[2.39]	[1.43]	[2.49]	[1.20]	[2.43]	[2.04]	[3.13]
ALPHA	-0.175***	-0.132***	-0.173***	-0.131***	-0.162***	-0.120***	-0.164***	-0.125***
	[-20.23]	[-18.21]	[-20.24]	[-18.14]	[-18.91]	[-16.45]	[-19.56]	[-17.39]
ADJRSQ	0.049***	0.045***	0.048***	0.044***	0.050***	0.046***	0.054***	0.049***
	[54.42]	[64.52]	[53.53]	[63.14]	[56.73]	[64.04]	[58.74]	[64.54]
OBS	4,130,988	3,998,755	4,130,840	3,998,608	4,148,286	4,015,812	4,148,286	4,015,812

Table. VIII (Continued)

Panel C. Illiquid stock (wide bid-ask spread)

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
OI_{t-1}	41.834***	11.281	0.415***	0.218***	0.875***	0.395***	27.371***	10.869***
	[8.92]	[1.35]	[22.62]	[15.36]	[38.58]	[38.33]	[45.74]	[35.10]
OI_{t-2}	-0.199	48.996	0.030*	-0.023*	0.153***	-0.001	4.088***	-0.906***
	[-0.05]	[0.80]	[1.92]	[-1.95]	[12.31]	[-0.09]	[11.38]	[-4.13]
OI_{t-3}	-16.515***	28.313	-0.054***	-0.041***	-0.017*	0.003	-0.924***	-0.625***
	[-3.73]	[0.52]	[-3.07]	[-3.60]	[-1.69]	[0.33]	[-3.46]	[-2.91]
OI_{t-4}	-9.257***	-31.807	-0.066***	-0.032***	-0.052***	0.003	-1.494***	-0.068
	[-2.84]	[-1.28]	[-4.20]	[-2.81]	[-5.41]	[0.43]	[-5.37]	[-0.34]
OI_{t-5}	-9.779**	-9.537*	-0.084***	-0.054***	-0.063***	-0.001	-1.885***	-0.241
	[-2.51]	[-1.67]	[-5.69]	[-4.56]	[-6.41]	[-0.10]	[-6.77]	[-1.14]
$BASPRD_{t-1}$	2.492***	0.188	2.578***	0.209*	2.219***	0.053	2.261***	0.072
	[16.50]	[1.63]	[16.80]	[1.79]	[14.80]	[0.46]	[15.06]	[0.63]
$TURN_{t-1}$	0.240***	0.099***	0.218***	0.088***	0.186***	0.098***	0.147***	0.104***
	[15.70]	[9.70]	[14.84]	[8.88]	[14.50]	[10.17]	[11.23]	[10.37]
RET_{t-1}	-19.405***	-2.434*	-19.374***	-2.288*	-23.242***	-4.799***	-25.518***	-5.580***
	[-48.58]	[-1.91]	[-48.62]	[-1.80]	[-57.26]	[-3.74]	[-60.71]	[-4.32]
RET_{t-2}	-5.421***	-1.834***	-5.441***	-1.831***	-6.551***	-2.083***	-7.118***	-1.858***
	[-32.75]	[-16.76]	[-33.28]	[-16.54]	[-35.22]	[-17.58]	[-34.77]	[-14.71]
RET_{t-3}	-2.029***	-1.615***	-2.067***	-1.652***	-2.170***	-1.655***	-2.178***	-1.477***
	[-18.87]	[-16.35]	[-19.02]	[-16.65]	[-18.53]	[-15.63]	[-17.68]	[-13.43]
RET_{t-4}	-0.700***	-1.304***	-0.720***	-1.318***	-0.565***	-1.378***	-0.479***	-1.329***
	[-6.59]	[-13.30]	[-6.68]	[-13.26]	[-4.70]	[-13.40]	[-3.65]	[-12.39]
RET_{t-5}	-0.200*	-0.886***	-0.189*	-0.863***	0.006	-0.924***	0.145	-0.882***
	[-1.95]	[-9.47]	[-1.80]	[-9.08]	[0.05]	[-9.52]	[1.21]	[-8.76]
RET^2_{t-1}	51.231***	44.350***	51.377***	44.407***	53.539***	45.172***	46.931***	37.627***
	[29.65]	[22.46]	[29.57]	[22.53]	[30.60]	[22.33]	[28.07]	[19.80]
RET^2_{t-2}	8.657***	3.202**	8.978***	3.610**	9.460***	4.159***	10.341***	6.535***
	[7.22]	[2.09]	[7.46]	[2.34]	[7.99]	[2.71]	[8.76]	[4.19]
RET^2_{t-3}	3.826***	1.749	3.754***	2.084	4.122***	2.005	5.186***	3.944***
	[3.43]	[1.23]	[3.32]	[1.46]	[3.77]	[1.43]	[4.70]	[2.75]
RET^2_{t-4}	4.806***	5.111***	4.991***	4.879***	5.066***	5.834***	6.337***	7.291***
	[4.26]	[3.57]	[4.36]	[3.42]	[4.53]	[4.07]	[5.61]	[4.96]
RET^2_{t-5}	2.710***	3.568***	3.077***	3.638***	2.273**	3.577***	3.688***	5.120***
	[2.67]	[2.66]	[3.01]	[2.69]	[2.24]	[2.67]	[3.64]	[3.78]
ALPHA	-0.202***	-0.082***	-0.201***	-0.081***	-0.188***	-0.072***	-0.185***	-0.072***
	[-18.09]	[-3.44]	[-18.02]	[-3.40]	[-17.17]	[-3.01]	[-17.00]	[-3.00]
ADJRSQ	0.083***	0.047***	0.083***	0.046***	0.090***	0.050***	0.095***	0.054***
	[61.05]	[73.91]	[59.87]	[69.90]	[65.44]	[72.64]	[68.44]	[71.11]
OBS	4,020,525	3,874,686	3,998,364	3,852,701	4,145,379	3,997,558	4,145,379	3,997,558

Table X
Return predictability in different exchange markets

This table shows the results of Fama-MacBeth regression to measure return predictability of HFOIs and LFOIs within each exchange markets. We separate whole sample dataset into two subsamples based on exchange market. Panel A is for NYSE and AMEX exchange markets. Panel B is for NASDAQ exchange market. The sample period is from 01 January 1993 to 31 December 2013. The first row represents which kind of lagged OI uses in a regression model. The dependent variable is either raw returns or mid-quote return at time t. The second row represents which type of return is the dependent variable of a regression model. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2011). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. TURN is daily turnover rate for each stock, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted bid-ask spread, the difference of bid and ask prices divided by the average of bid and ask prices. We only use bid-ask spread within the interval between 0 and 0.5. Also bid and ask prices are close price at each day. RET is daily stock returns. RET^2 is daily squared stock returns. ALPHA is the intercept of regression models. ADJRSQ is adjusted R-square. OBS is the number of observation in each regression model. Each subscript of variables indicates its time lag. In brackets, we report t-statistics of the average coefficient over sample period based on New-West (1987) standard errors. ***, **, and * indicate statistical significance at the 1, 5, and 10 per cent level, respectively.

Table. X (Continued)

Panel A. NYSE and AMEX

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
OI_{t-1}	2.059**	4.867***	0.056***	0.059***	0.074***	0.106***	2.295***	3.366***
	[2.42]	[4.52]	[6.45]	[6.50]	[18.20]	[18.80]	[13.99]	[16.67]
OI_{t-2}	-4.943***	-2.747	-0.043***	-0.017	-0.019***	-0.005	-1.032***	-0.567
	[-6.76]	[-0.84]	[-5.73]	[-0.43]	[-4.58]	[-0.27]	[-5.62]	[-0.71]
OI_{t-3}	-3.001***	-4.790**	-0.025***	-0.035***	-0.009***	-0.009**	-0.398***	-0.409
	[-4.62]	[-2.14]	[-3.60]	[-2.61]	[-3.20]	[-2.37]	[-3.49]	[-1.51]
OI_{t-4}	-3.384***	-5.454***	-0.022***	-0.030***	-0.004	-0.006*	0.005	-0.567
	[-4.62]	[-3.39]	[-3.22]	[-3.07]	[-1.31]	[-1.88]	[0.04]	[-1.01]
OI_{t-5}	-1.864***	-1.532**	-0.018**	-0.016**	0.001	0.002	0.152	0.217*
	[-2.64]	[-2.04]	[-2.38]	[-2.05]	[0.37]	[0.71]	[1.35]	[1.72]
$BASPRD_{t-1}$	2.793***	2.172***	2.829***	2.183***	2.717***	2.103***	2.814***	2.179***
	[11.58]	[10.04]	[11.66]	[10.06]	[11.24]	[9.72]	[11.70]	[10.17]
$TURN_{t-1}$	0.081***	0.093***	0.080***	0.109***	0.085***	0.094***	0.089***	0.099***
	[14.79]	[5.22]	[14.33]	[3.30]	[16.95]	[5.32]	[17.15]	[4.50]
RET_{t-1}	-0.686***	0.134	-0.776***	0.106	-1.623***	-1.029***	-1.925***	-1.351***
	[-5.35]	[0.96]	[-6.05]	[0.80]	[-11.98]	[-5.73]	[-12.32]	[-6.86]
RET_{t-2}	-0.887***	-0.885***	-0.881***	-0.869***	-0.688***	-0.386	-0.442***	-0.105
	[-8.57]	[-8.11]	[-8.47]	[-7.96]	[-6.45]	[-1.54]	[-3.82]	[-0.41]
RET_{t-3}	-0.645***	-0.710***	-0.640***	-0.722***	-0.549***	-0.642***	-0.486***	-0.610***
	[-5.98]	[-6.49]	[-5.94]	[-6.40]	[-5.00]	[-5.29]	[-4.14]	[-4.60]
RET_{t-4}	-0.562***	-0.582***	-0.551***	-0.571***	-0.558***	-0.763***	-0.612***	-0.835***
	[-5.79]	[-6.00]	[-5.67]	[-5.84]	[-5.72]	[-3.67]	[-5.87]	[-3.94]
RET_{t-5}	-0.307***	-0.336***	-0.282***	-0.319***	-0.375***	-0.700**	-0.396***	-0.774***
	[-3.38]	[-3.68]	[-3.08]	[-3.43]	[-3.98]	[-2.39]	[-3.86]	[-2.63]
RET^2_{t-1}	16.440***	21.475***	16.381***	21.070***	16.582***	20.215***	14.743***	15.605***
	[7.98]	[8.19]	[7.90]	[8.03]	[8.48]	[7.84]	[7.51]	[6.39]
RET^2_{t-2}	3.074***	2.377	2.846**	2.423	2.024*	-7.821	3.863***	-3.760
	[2.75]	[1.47]	[2.55]	[1.46]	[1.80]	[-0.92]	[3.31]	[-0.44]
RET^2_{t-3}	1.885	4.327**	1.805	4.422**	0.878	2.763	2.362**	4.718***
	[1.59]	[2.16]	[1.54]	[2.39]	[0.75]	[1.48]	[1.98]	[2.58]
RET^2_{t-4}	2.772***	2.934**	2.648**	3.015**	2.448**	1.961	2.988***	3.996***
	[2.58]	[2.10]	[2.51]	[2.24]	[2.29]	[1.31]	[2.81]	[2.75]
RET^2_{t-5}	3.489***	3.794***	3.572***	3.348**	3.122***	2.970**	3.960***	3.963***
	[3.53]	[2.64]	[3.59]	[2.29]	[3.20]	[1.99]	[4.05]	[2.67]
ALPHA	-0.060***	-0.051***	-0.059***	-0.049***	-0.053***	-0.041***	-0.056***	-0.044***
	[-12.97]	[-10.80]	[-12.69]	[-10.55]	[-11.24]	[-7.65]	[-12.10]	[-8.25]
ADJRSQ	0.043***	0.041***	0.043***	0.041***	0.045***	0.043***	0.048***	0.047***
	[39.01]	[35.10]	[38.72]	[35.10]	[39.85]	[36.47]	[40.61]	[36.44]
OBS	8,372,963	8,037,733	8,371,187	8,035,968	8,398,435	8,061,443	8,398,435	8,061,443

Table. X (Continued)

Panel B. NASDAQ

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
OI_{t-1}	-64.685	82.134	0.200***	0.067***	0.410***	0.138***	14.341***	3.551***
	[-0.14]	[0.77]	[19.65]	[11.02]	[26.38]	[32.53]	[27.78]	[29.23]
OI_{t-2}	-321.76	-42.512	-0.026***	-0.034***	0.009**	-0.030***	0.428***	-1.312***
	[-0.85]	[-0.41]	[-3.41]	[-6.70]	[2.10]	[-9.86]	[3.12]	[-12.97]
OI_{t-3}	1138.20*	396.561	-0.039***	-0.026***	-0.028***	-0.011***	-0.940***	-0.516***
	[1.71]	[1.14]	[-5.83]	[-5.23]	[-8.30]	[-4.20]	[-8.64]	[-6.16]
OI_{t-4}	-47.372	-173.37	-0.040***	-0.018***	-0.030***	-0.004	-0.835***	-0.039
	[-0.42]	[-1.58]	[-6.21]	[-3.54]	[-8.88]	[-1.57]	[-8.03]	[-0.52]
OI_{t-5}	40.426	2.344	-0.041***	-0.023***	-0.022***	-0.001	-0.685***	0.021
	[0.69]	[0.29]	[-6.57]	[-4.61]	[-7.07]	[-0.30]	[-6.95]	[0.29]
$BASPRD_{t-1}$	2.846***	0.810***	2.948***	0.839***	2.438***	0.606***	2.401***	0.638***
	[18.57]	[6.27]	[18.52]	[6.37]	[16.41]	[4.79]	[16.20]	[5.04]
$TURN_{t-1}$	0.125***	0.061***	0.111***	0.052***	0.099***	0.070***	0.069***	0.080***
	[14.69]	[10.71]	[13.39]	[9.19]	[15.28]	[12.89]	[12.24]	[14.42]
RET_{t-1}	-9.293***	0.091	-9.286***	0.152	-12.872***	-1.704***	-15.421***	-2.257***
	[-32.18]	[0.42]	[-31.75]	[0.71]	[-37.47]	[-7.36]	[-40.55]	[-8.31]
RET_{t-2}	-2.430***	-1.734***	-2.450***	-1.731***	-2.874***	-1.559***	-3.209***	-1.219***
	[-23.30]	[-20.69]	[-23.29]	[-20.40]	[-23.40]	[-18.53]	[-22.95]	[-13.19]
RET_{t-3}	-1.115***	-1.264***	-1.104***	-1.243***	-0.956***	-1.167***	-0.861***	-1.037***
	[-15.25]	[-17.13]	[-14.87]	[-16.73]	[-12.27]	[-15.37]	[-9.89]	[-12.80]
RET_{t-4}	-0.586***	-0.997***	-0.565***	-0.983***	-0.370***	-0.997***	-0.269***	-1.025***
	[-8.00]	[-13.46]	[-7.66]	[-13.07]	[-4.54]	[-12.74]	[-2.96]	[-12.16]
RET_{t-5}	-0.382***	-0.754***	-0.364***	-0.731***	-0.230***	-0.791***	-0.135	-0.811***
	[-5.36]	[-10.38]	[-4.99]	[-9.91]	[-2.96]	[-10.41]	[-1.58]	[-10.15]
RET^2_{t-1}	24.096***	28.557***	24.530***	28.740***	26.611***	29.211***	23.766***	25.456***
	[22.85]	[23.41]	[23.27]	[23.62]	[24.48]	[23.15]	[24.09]	[20.76]
RET^2_{t-2}	4.588***	4.641***	4.740***	4.837***	4.244***	3.465***	5.215***	6.136***
	[7.14]	[4.88]	[7.41]	[5.04]	[6.75]	[3.63]	[8.24]	[6.45]
RET^2_{t-3}	4.132***	6.980***	4.232***	7.051***	3.887***	6.316***	4.864***	8.135***
	[6.58]	[7.46]	[6.78]	[7.50]	[6.36]	[6.83]	[8.04]	[8.68]
RET^2_{t-4}	4.823***	8.487***	4.840***	8.010***	4.482***	7.883***	5.287***	9.535***
	[8.16]	[9.53]	[8.14]	[9.01]	[7.80]	[8.96]	[9.17]	[10.74]
RET^2_{t-5}	4.082***	6.899***	4.110***	6.941***	3.992***	6.699***	4.706***	7.700***
	[6.53]	[7.37]	[6.49]	[7.33]	[6.49]	[7.17]	[7.66]	[8.26]
ALPHA	-0.091***	-0.030***	-0.089***	-0.028***	-0.073***	-0.018***	-0.070***	-0.021***
	[-16.17]	[-6.98]	[-15.83]	[-6.55]	[-13.28]	[-4.15]	[-12.93]	[-4.83]
ADJRSQ	0.041***	0.029***	0.040***	0.028***	0.044***	0.030***	0.047***	0.031***
	[47.79]	[56.77]	[46.28]	[56.71]	[49.54]	[57.14]	[51.15]	[56.68]
OBS	12,171,944	11,837,540	12,150,758	11,816,527	12,338,297	12,002,632	12,338,997	12,002,632

Table XI
Return predictability in period subsamples

This table shows the results of Fama-MacBeth regression to measure return predictability of HFOIs and LFOIs within each size subsamples. We separate whole sample dataset into three subsamples based on subperiod. Panel A is for early subperiod from 1993 to 2000. Panel B is for midst subperiod from 2001 to 2006. Panel C is for late subperiod from 2007 to 2013. The whole sample period is from 01 January 1993 to 31 December 2013. The first row represents which kind of lagged OI uses in a regression model. The dependent variable is either raw returns or mid-quote return at time t. The second row represents which type of return is the dependent variable of a regression model. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2011). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. TURN is daily turnover rate for each stock, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted bid-ask spread, the difference of bid and ask prices divided by the average of bid and ask prices. We only use bid-ask spread within the interval between 0 and 0.5. Also bid and ask prices are close price at each day. RET is daily stock returns. RET^2 is daily squared stock returns. ALPHA is the intercept of regression models. ADJRSQ is adjusted R-square. OBS is the number of observation in each regression model. Each subscript of variables indicates its time lag. In brackets, we report t-statistics of the average coefficient over sample period based on New-West (1987) standard errors. ***, **, and * indicate statistical significance at the 1, 5, and 10 per cent level, respectively.

Table. XI (Continued)

Panel A. Early period (1993-2000)

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
OI_{t-1}	36.487***	9.994***	0.355***	0.123***	0.689***	0.203***	22.833***	3.839***
	[20.53]	[12.34]	[33.68]	[20.89]	[45.98]	[36.81]	[33.01]	[21.18]
OI_{t-2}	-1.935***	-3.947***	-0.012*	-0.033***	0.005	-0.058***	-0.251	-2.824***
	[-2.74]	[-6.66]	[-1.88]	[-6.79]	[0.79]	[-13.12]	[-1.23]	[-22.77]
OI_{t-3}	-4.543***	-1.428***	-0.055***	-0.024***	-0.053***	-0.018***	-2.051***	-0.958***
	[-8.15]	[-3.16]	[-9.59]	[-5.45]	[-9.63]	[-4.26]	[-11.68]	[-7.37]
OI_{t-4}	-4.437***	-1.523***	-0.045***	-0.014***	-0.051***	-0.006	-1.668***	-0.139
	[-7.15]	[-2.99]	[-7.64]	[-3.09]	[-10.10]	[-1.57]	[-10.44]	[-1.20]
OI_{t-5}	-3.982***	-1.200**	-0.047***	-0.019***	-0.042***	-0.002	-1.351***	0.074
	[-7.25]	[-2.48]	[-9.12]	[-4.51]	[-8.43]	[-0.60]	[-9.14]	[0.68]
$BASPRD_{t-1}$	2.811***	0.858***	2.786***	0.844***	2.466***	0.698***	2.386***	0.697***
	[20.26]	[7.82]	[19.95]	[7.67]	[18.09]	[6.38]	[17.68]	[6.38]
$TURN_{t-1}$	0.253***	0.150***	0.235***	0.140***	0.187***	0.144***	0.133***	0.153***
	[24.45]	[19.26]	[22.55]	[17.84]	[21.74]	[19.29]	[15.79]	[20.36]
RET_{t-1}	-12.130***	3.606***	-12.257***	3.564***	-16.995***	1.841***	-19.712***	2.106***
	[-28.34]	[15.95]	[-28.71]	[15.67]	[-36.13]	[7.67]	[-37.84]	[8.43]
RET_{t-2}	-3.224***	-2.341***	-3.268***	-2.351***	-3.770***	-1.809***	-3.841***	-0.994***
	[-23.83]	[-17.31]	[-24.33]	[-17.22]	[-25.54]	[-15.07]	[-22.56]	[-7.99]
RET_{t-3}	-1.146***	-1.560***	-1.124***	-1.539***	-0.821***	-1.308***	-0.501***	-1.043***
	[-6.02]	[-13.62]	[-5.89]	[-13.37]	[-4.72]	[-9.16]	[-2.86]	[-7.27]
RET_{t-4}	-0.912***	-1.140***	-0.890***	-1.148***	-0.224	-1.165***	0.039	-1.149***
	[-5.88]	[-9.33]	[-5.74]	[-9.34]	[-1.34]	[-11.43]	[0.23]	[-11.02]
RET_{t-5}	-0.679***	-0.921***	-0.663***	-0.912***	-0.006	-0.957***	0.188	-1.008***
	[-2.82]	[-6.61]	[-2.74]	[-6.53]	[-0.04]	[-8.39]	[1.21]	[-8.74]
RET^2_{t-1}	34.142***	27.191***	34.569***	27.533***	26.337***	28.177***	20.662***	23.724***
	[7.80]	[18.88]	[7.91]	[19.18]	[3.58]	[24.38]	[2.82]	[20.07]
RET^2_{t-2}	0.544	1.605	0.923	1.772*	1.720	0.655	2.427	4.198***
	[0.20]	[1.60]	[0.34]	[1.76]	[0.75]	[0.69]	[1.06]	[4.37]
RET^2_{t-3}	1.854**	1.506	1.945**	1.476	-3.825	-0.870	-2.633	0.817
	[2.08]	[1.44]	[2.17]	[1.41]	[-0.75]	[-0.42]	[-0.51]	[0.39]
RET^2_{t-4}	1.229**	2.730***	1.241**	2.708***	12.212	6.873*	13.119	7.873*
	[2.22]	[2.84]	[2.25]	[2.82]	[1.12]	[1.67]	[1.20]	[1.91]
RET^2_{t-5}	0.850	1.427	0.798	1.449	0.935*	1.725*	1.906***	2.433***
	[1.53]	[1.50]	[1.43]	[1.52]	[1.69]	[1.82]	[3.48]	[2.60]
ALPHA	-0.149***	-0.065***	-0.146***	-0.062***	-0.125***	-0.054***	-0.115***	-0.056***
	[-21.11]	[-9.80]	[-20.56]	[-9.36]	[-17.57]	[-8.03]	[-15.96]	[-8.40]
ADJRSQ	0.033***	0.021***	0.032***	0.021***	0.038***	0.022***	0.041***	0.023***
	[38.69]	[33.98]	[38.01]	[33.82]	[39.54]	[34.66]	[40.36]	[34.13]
OBS	9,870,129	9,643,287	98,699,865	9,643,029	9,983,065	9,754,770	9,983,065	9,754,770

Table. XI (Continued)

Panel B. Midst period (2001-2006)

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
OI_{t-1}	0.403	0.413	0.058***	0.013	0.143***	0.085***	5.885***	3.004***
	[0.45]	[0.56]	[6.27]	[1.49]	[17.64]	[16.65]	[21.86]	[15.58]
OI_{t-2}	-6.740***	-5.981***	-0.074***	-0.064***	-0.013***	-0.019***	-0.620***	-0.894***
	[-8.79]	[-9.17]	[-7.70]	[-8.38]	[-3.19]	[-4.19]	[-4.47]	[-6.34]
OI_{t-3}	-2.960***	-2.537***	-0.041***	-0.038***	-0.006	-0.002	-0.262*	-0.200
	[-3.52]	[-3.73]	[-4.55]	[-4.93]	[-1.62]	[-0.51]	[-1.91]	[-1.52]
OI_{t-4}	-1.051	-1.331**	-0.026***	-0.026***	-0.007*	-0.002	-0.056	0.055
	[-1.33]	[-1.99]	[-2.94]	[-3.54]	[-1.68]	[-0.44]	[-0.42]	[0.46]
OI_{t-5}	-1.737**	-1.104	-0.024***	-0.016**	0.002	0.004	0.041	0.144
	[-2.19]	[-1.57]	[-2.81]	[-2.13]	[0.49]	[1.06]	[0.31]	[1.13]
$BASPRD_{t-1}$	2.815***	2.121***	2.817***	2.099***	2.701***	2.009***	2.682***	2.045***
	[10.70]	[9.28]	[10.76]	[9.20]	[10.09]	[8.82]	[10.08]	[9.00]
$TURN_{t-1}$	0.027***	0.022***	0.015*	0.014*	0.023***	0.021***	0.027***	0.033***
	[3.26]	[2.98]	[1.70]	[1.73]	[3.10]	[2.99]	[4.01]	[4.72]
RET_{t-1}	-4.244***	-1.169***	-4.325***	-1.178***	-6.282***	-2.571***	-7.815***	-3.186***
	[-22.63]	[-6.51]	[-23.15]	[-6.56]	[-29.63]	[-13.93]	[-32.89]	[-15.04]
RET_{t-2}	-1.373***	-1.337***	-1.377***	-1.339***	-1.423***	-1.238***	-1.312***	-0.932***
	[-9.93]	[-8.99]	[-9.94]	[-8.96]	[-10.44]	[-8.94]	[-8.90]	[-6.16]
RET_{t-3}	-0.796***	-0.975***	-0.773***	-0.943***	-0.771***	-1.009***	-0.695***	-0.930***
	[-6.86]	[-7.75]	[-6.65]	[-7.44]	[-6.53]	[-7.91]	[-5.46]	[-7.11]
RET_{t-4}	-0.575***	-0.852***	-0.537***	-0.810***	-0.525***	-0.900***	-0.563***	-0.950***
	[-4.78]	[-6.52]	[-4.44]	[-6.15]	[-4.07]	[-6.86]	[-3.97]	[-6.88]
RET_{t-5}	-0.364***	-0.507***	-0.355***	-0.486***	-0.440***	-0.628***	-0.426***	-0.639***
	[-3.13]	[-4.04]	[-3.01]	[-3.82]	[-3.69]	[-5.16]	[-3.29]	[-4.97]
RET^2_{t-1}	16.295***	31.734***	16.279***	31.559***	18.161***	32.624***	17.347***	29.081***
	[11.47]	[14.94]	[11.59]	[15.14]	[12.76]	[14.95]	[12.47]	[14.18]
RET^2_{t-2}	4.584***	5.829***	4.292***	5.321***	3.950***	4.545***	4.605***	6.470***
	[4.94]	[3.98]	[4.67]	[3.65]	[4.31]	[3.12]	[5.00]	[4.38]
RET^2_{t-3}	5.174***	9.328***	5.276***	9.084***	4.778***	8.251***	5.420***	10.307***
	[5.62]	[6.18]	[5.83]	[6.02]	[5.28]	[5.39]	[6.00]	[6.58]
RET^2_{t-4}	6.398***	9.040***	6.250***	8.558***	6.384***	8.610***	6.764***	9.794***
	[6.68]	[5.79]	[6.48]	[5.50]	[6.44]	[5.45]	[6.83]	[6.32]
RET^2_{t-5}	5.782***	9.485***	5.719***	9.241***	5.409***	8.993***	5.805***	9.649***
	[5.18]	[5.61]	[5.12]	[5.49]	[4.92]	[5.33]	[5.35]	[5.71]
ALPHA	-0.034***	-0.020***	-0.031***	-0.016***	-0.027***	-0.012**	-0.027***	-0.014***
	[-6.15]	[-3.90]	[-5.45]	[-3.11]	[-4.76]	[-2.24]	[-4.79]	[-2.65]
ADJRSQ	0.029***	0.026***	0.028***	0.025***	0.029***	0.026***	0.032***	0.028***
	[26.96]	[33.15]	[26.66]	[33.37]	[28.73]	[32.92]	[29.91]	[32.67]
OBS	5,360,555	4,979,642	5,360,249	4,979,341	5,388,132	5,006,497	5,388,132	5,006,497

Table. XI (Continued)

Panel C. Late period (2007-2013)

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
OI_{t-1}	-3.873***	0.396	0.060***	0.029***	0.086***	0.060***	4.656***	3.200***
	[-4.00]	[0.52]	[5.15]	[3.32]	[21.63]	[18.05]	[26.34]	[22.71]
OI_{t-2}	-2.066**	-2.501***	-0.007	-0.015*	-0.002	-0.003	0.192	0.015
	[-2.34]	[-3.62]	[-0.57]	[-1.82]	[-0.67]	[-1.30]	[1.38]	[0.14]
OI_{t-3}	-0.887	-0.797	-0.018*	-0.012	-0.002	-0.001	-0.077	-0.043
	[-1.01]	[-1.20]	[-1.83]	[-1.62]	[-0.97]	[-0.41]	[-0.64]	[-0.44]
OI_{t-4}	-2.596***	-1.793***	-0.013	-0.006	-0.004*	-0.000	-0.137	0.036
	[-3.43]	[-2.91]	[-1.37]	[-0.79]	[-1.66]	[-0.02]	[-1.24]	[0.38]
OI_{t-5}	-1.481*	-1.710***	-0.018*	-0.017**	-0.001	0.001	-0.095	-0.005
	[-1.79]	[-2.70]	[-1.68]	[-2.07]	[-0.24]	[0.62]	[-0.85]	[-0.06]
$BASPRD_{t-1}$	2.278***	0.051	2.655***	0.230	1.805***	-0.168	1.848***	-0.088
	[7.02]	[0.19]	[7.63]	[0.83]	[5.88]	[-0.65]	[5.98]	[-0.34]
$TURN_{t-1}$	0.009*	-0.004	0.010*	-0.003	0.038***	0.021***	0.035***	0.022***
	[1.72]	[-0.84]	[1.87]	[-0.68]	[6.70]	[4.05]	[6.81]	[4.59]
RET_{t-1}	-4.194***	-2.436***	-4.043***	-2.308***	-6.387***	-4.079***	-8.394***	-5.580***
	[-18.55]	[-11.72]	[-17.50]	[-11.07]	[-25.31]	[-18.75]	[-28.89]	[-21.37]
RET_{t-2}	-0.746***	-0.662***	-0.756***	-0.648***	-0.893***	-0.667***	-1.237***	-0.850***
	[-4.77]	[-4.47]	[-4.82]	[-4.33]	[-5.28]	[-4.35]	[-6.57]	[-4.92]
RET_{t-3}	-0.649***	-0.712***	-0.645***	-0.716***	-0.576***	-0.681***	-0.613***	-0.702***
	[-4.40]	[-4.87]	[-4.29]	[-4.86]	[-3.79]	[-4.66]	[-3.66]	[-4.50]
RET_{t-4}	-0.293**	-0.542***	-0.307**	-0.536***	-0.229	-0.543***	-0.229	-0.596***
	[-1.99]	[-3.54]	[-2.07]	[-3.45]	[-1.51]	[-3.60]	[-1.46]	[-3.77]
RET_{t-5}	-0.240*	-0.347**	-0.227	-0.328**	-0.229	-0.417***	-0.161	-0.391**
	[-1.65]	[-2.38]	[-1.53]	[-2.23]	[-1.50]	[-2.75]	[-1.01]	[-2.48]
RET^2_{t-1}	13.039***	20.800***	13.269***	20.517***	13.905***	20.596***	14.165***	17.779***
	[7.91]	[9.08]	[8.08]	[8.94]	[8.45]	[8.54]	[8.69]	[7.65]
RET^2_{t-2}	3.534***	6.475***	3.670***	6.900***	2.068*	4.546**	3.066***	6.714***
	[3.37]	[3.54]	[3.50]	[3.78]	[1.96]	[2.48]	[2.88]	[3.65]
RET^2_{t-3}	4.100***	9.492***	4.100***	9.865***	3.285***	8.010***	4.003***	9.926***
	[3.67]	[5.78]	[3.70]	[5.98]	[3.04]	[4.97]	[3.75]	[6.10]
RET^2_{t-4}	4.894***	10.294***	5.007***	9.626***	4.023***	9.009***	4.656***	11.428***
	[5.19]	[6.58]	[5.40]	[6.19]	[4.37]	[5.87]	[5.06]	[7.32]
RET^2_{t-5}	4.800***	10.655***	4.979***	10.620***	4.177***	9.842***	4.645***	11.509***
	[4.77]	[6.58]	[4.93]	[6.56]	[4.26]	[6.14]	[4.72]	[7.30]
ALPHA	-0.025***	-0.012*	-0.026***	-0.012**	-0.014**	-0.000	-0.019***	-0.006
	[-3.96]	[-1.95]	[-4.13]	[-2.06]	[-2.15]	[-0.07]	[-2.91]	[-1.00]
ADJRSQ	0.039***	0.032***	0.040***	0.032***	0.041***	0.034***	0.044***	0.036***
	[27.14]	[28.83]	[26.71]	[29.40]	[28.42]	[29.24]	[29.70]	[29.85]
OBS	5,314,223	5,252,344	5,291,831	5,230,125	5,365,535	5,302,808	5,365,535	5,302,808

Table XII
Return predictability of Asymmetric Order Imbalance

This table shows the results of Fama-MacBeth regression to measure return predictability of asymmetric order imbalances. The sample period is from 01 January 1993 to 31 December 2013. The first row represents which kind of lagged OI uses in a regression model. The dependent variable is either raw returns or mid-quote return at time t. The second row represents which type of return is the dependent variable of a regression model. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2013). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. OI[+] (OI[-]) is equivalant to order imbalance if order imbalance is positive (negative), and otherwise OI[+] (OI[-]) is zero. TURN is daily turnover rate for each stock, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted bid-ask spread, the difference of bid and ask prices divided by the average of bid and ask prices. We only use bid-ask spread within the interval between 0 and 0.5. Also bid and ask prices are close price at each day. RET is daily stock returns. RET^2 is daily squared stock returns. ALPHA is the intercept of regression models. ADJRSQ is adjusted R-square. OBS is the number of observation in each regression model. Panel A reports daily-based Fama-MacBeth regression for return predictability. Panel B reports weekly-based Fama-MacBeth regression. Each subscript of variables indicates its time lag. In brackets, we report t-statistics of the average coefficient over sample period based on New-West (1987) standard errors. ***, **, and * indicate statistical significance at the 1, 5, and 10 per cent level, respectively.

Table. XII (Continued)

	HFOI1		HFOI2		LFOI1		LFOI2	
	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q	RAW	MID-Q
$OI[+]_{t-1}$	23.709***	14.672***	0.323***	0.189***	0.438***	0.234***	11.872***	4.992***
	[16.70]	[16.07]	[27.37]	[19.81]	[29.08]	[26.63]	[27.92]	[26.45]
$OI[-]_{t-1}$	-2.120*	-2.814***	0.031**	-0.040***	0.100***	-0.001	8.757***	2.112***
	[-1.82]	[-3.52]	[2.43]	[-4.48]	[7.53]	[-0.12]	[18.18]	[9.87]
$OI[+]_{t-2}$	-3.359***	-4.558***	0.000	-0.016**	-0.022***	-0.043***	0.285*	-0.999***
	[-4.18]	[-6.81]	[0.02]	[-2.41]	[-4.90]	[-11.38]	[1.94]	[-7.31]
$OI[-]_{t-2}$	-4.357***	-4.725***	-0.063***	-0.065***	0.009*	-0.011***	-1.622***	-2.238***
	[-6.35]	[-8.29]	[-7.14]	[-9.76]	[1.81]	[-2.77]	[-6.63]	[-12.06]
$OI[+]_{t-3}$	-5.063***	-2.481***	-0.032***	-0.008	-0.031***	-0.014***	-0.379***	-0.216**
	[-6.75]	[-4.27]	[-4.29]	[-1.22]	[-7.11]	[-4.06]	[-2.90]	[-1.97]
$OI[-]_{t-3}$	-1.896***	-2.197***	-0.046***	-0.042***	-0.003	0.002	-1.355***	-0.792***
	[-2.89]	[-3.98]	[-5.58]	[-6.32]	[-0.76]	[0.40]	[-7.36]	[-4.89]
$OI[+]_{t-4}$	-4.005***	-1.630***	-0.027***	-0.009	-0.024***	-0.008**	-0.174	0.227**
	[-5.43]	[-2.70]	[-3.32]	[-1.33]	[-6.00]	[-2.27]	[-1.36]	[1.99]
$OI[-]_{t-4}$	-2.174***	-2.114***	-0.036***	-0.026***	-0.004	0.007*	-0.962***	-0.236
	[-3.19]	[-3.87]	[-4.60]	[-4.30]	[-1.03]	[1.73]	[-5.87]	[-1.62]
$OI[+]_{t-5}$	-4.646***	-2.316***	-0.041***	-0.016***	-0.031***	-0.014***	-0.172	0.196*
	[-6.03]	[-3.67]	[-5.43]	[-2.61]	[-7.29]	[-3.79]	[-1.41]	[1.71]
$OI[-]_{t-5}$	-1.041	-1.213**	-0.025***	-0.024***	0.015***	0.021***	-0.439***	0.167
	[-1.53]	[-2.21]	[-3.05]	[-3.67]	[3.53]	[5.06]	[-2.62]	[1.13]
$BASPRD_{t-1}$	2.493***	0.955***	2.508***	0.941***	2.249***	0.764***	2.342***	0.894***
	[17.87]	[7.73]	[17.95]	[7.62]	[15.80]	[6.02]	[16.71]	[7.16]
$TURN_{t-1}$	0.081***	0.047***	0.066***	0.032***	-0.014*	-0.002	0.036***	0.049***
	[12.20]	[9.18]	[9.52]	[5.96]	[-1.66]	[-0.21]	[7.24]	[11.19]
RET_{t-1}	-6.528***	-0.117	-6.608***	-0.128	-9.127***	-1.620***	-11.203***	-2.148***
	[-28.46]	[-0.61]	[-29.39]	[-0.66]	[-33.71]	[-8.40]	[-36.09]	[-9.52]
RET_{t-2}	-1.819***	-1.242***	-1.846***	-1.262***	-2.003***	-1.085***	-2.030***	-0.713***
	[-19.51]	[-5.13]	[-19.94]	[-5.21]	[-19.99]	[-4.48]	[-17.94]	[-2.90]
RET_{t-3}	-0.887***	-1.053***	-0.884***	-1.049***	-0.796***	-1.040***	-0.656***	-0.911***
	[-10.75]	[-10.84]	[-10.69]	[-10.78]	[-9.38]	[-10.68]	[-6.98]	[-8.88]
RET_{t-4}	-0.569***	-1.047***	-0.565***	-1.039***	-0.475***	-1.112***	-0.416***	-1.170***
	[-8.17]	[-5.30]	[-8.10]	[-5.25]	[-6.34]	[-5.62]	[-4.96]	[-5.81]
RET_{t-5}	-0.358***	-0.911***	-0.354***	-0.902***	-0.328***	-1.014***	-0.270***	-1.058***
	[-5.36]	[-3.18]	[-5.23]	[-3.15]	[-4.52]	[-3.53]	[-3.35]	[-3.68]
RET^2_{t-1}	20.137***	24.910***	20.751***	25.676***	21.843***	23.301***	18.997***	18.472***
	[13.83]	[16.78]	[14.24]	[17.34]	[14.75]	[15.47]	[12.96]	[11.67]
RET^2_{t-2}	2.936***	-4.067	2.796***	-4.281	2.766***	-4.556	2.428***	-3.450
	[5.29]	[-0.49]	[5.09]	[-0.51]	[4.98]	[-0.54]	[4.05]	[-0.41]
RET^2_{t-3}	3.270***	6.318***	2.972***	5.729***	3.328***	6.872***	2.879***	7.422***
	[5.80]	[7.58]	[5.30]	[6.87]	[5.94]	[8.01]	[4.97]	[7.95]
RET^2_{t-4}	3.925***	6.903***	3.523***	6.401***	4.164***	8.380***	3.381***	7.934***
	[7.49]	[8.43]	[6.81]	[7.78]	[7.91]	[10.08]	[6.05]	[8.69]
RET^2_{t-5}	4.054***	7.062***	3.724***	6.547***	4.751***	9.788***	4.247***	8.938***
	[7.48]	[8.36]	[6.86]	[7.70]	[8.70]	[11.66]	[7.65]	[9.81]
ALPHA	-0.068***	-0.032***	-0.071***	-0.033***	-0.051***	-0.018***	-0.054***	-0.024***
	[-16.15]	[-7.14]	[-16.57]	[-7.45]	[-11.74]	[-3.97]	[-13.25]	[-5.42]
ADJRSQ	0.035***	0.028***	0.034***	0.027***	0.037***	0.029***	0.041***	0.032***
	[48.62]	[50.96]	[47.92]	[50.06]	[51.97]	[50.49]	[54.51]	[50.67]
OBS	20,544,907	19,875,273	20,521,945	19,852,495	20,736,732	20,064,705	20,736,732	20,064,705

Table XIII
Order Imbalance Porftolio Investment Strategy

This table shows the performance, daily Sharpe ratio, and Fama-French three factor alpha of investment strategy based on order imbalance. The sample period is from 01 January 1993 to 31 December 2013. The first row represents which kind of OI is used in the investment simulation. The rows from Highest OI to Lowest OI are the equal-weighted average return of decile portfolios contructed by the amount of order imbalance in the previous day. Highest OI is top decile portfolio and Lowest OI is bottom decile portfolio. In the row, High OI-Low OI, we report the performance of investment strategy which long Highest OI portfolio and short Lowest OI portfolio. FF3 Alpha row documents Fama-French three factor alphas. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 aredefined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2011). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. In this table, we employ OI residual to contruct decile portfolio, by regressing OI on contemporaneous return, squared return, turnover, and bid-ask spread. FF3 Alpha is the intercept in the regression of portfolio permance on Fama-French three factors including excess market return, Small-Minus-Big, and High-Minus-Low. In brackets, this talbe report t-statistics of the average coefficient over sample period based on New-West (1987) standard errors.

	HFOI1 Residual	HFOI2 Residual	LFOI1 Residual	LFOI2 Residual
Highest OI	0.311 %	0.320 %	0.499 %	0.613 %
9	0.251	0.282	0.317	0.200
8	0.157	0.195	0.179	0.082
7	0.092	0.120	0.101	0.058
6	0.061	0.067	0.064	0.065
5	0.049	0.033	0.059	0.077
4	0.033	0.003	0.038	0.083
3	0.006	-0.032	-0.015	0.070
2	-0.009	-0.026	-0.085	-0.008
Lowest OI	0.132	0.120	-0.079	-0.162
High OI - Low OI	0.178 %	0.200 %	0.579 %	0.776 %
	[23.79]	[33.30]	[70.66]	[80.38]
Daily Sharpe Ratio	30.84 %	43.47 %	95.75 %	109.44 %
FF3 Alpha (x1000)	1.790	1.990	5.790	7.760
	[24.53]	[33.50]	[70.98]	[80.42]

Table XIV
Return predictability around earnings announcements

This table shows the results of Fama-MacBeth regression to measure return predictability of HFOIs and LFOIs around earnings announcement. The sample period is from 01 January 1993 to 31 December 2013. The first row represents which kind of lagged OI uses in a regression model. The dependent variable is raw cumulative abnormal returns at time t. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2013). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. TURN is daily turnover rate for each stock, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted bid-ask spread, the difference of bid and ask prices divided by the average of bid and ask prices. We only use bid-ask spread within the interval between 0 and 0.5. Also bid and ask prices are close price at each day. RET is daily stock returns. RET^2 is daily squared stock returns. ALPHA is the intercept of regression models. ADJRSQ is adjusted R-square. OBS is the number of observation in each regression model. Panel A reports test results of regressions of $CAR_{t,t}$, the cumulative abnormal return at earnings announcement date. Panel B reports test results of regressions of $CAR_{t-1,t+1}$, the cumulative abnormal return from a day before to a day after earnings announcement date. Panel C reports test results of regressions of $CAR_{t-2,t+2}$, the cumulative abnormal return from two days before to two days after earnings announcement date. Abnormal returns are residual terms in the regression of raw returns on Fama-French three factors. Each subscript of variables indicates its time lag. In brackets, we report t-statistics of the average coefficient over sample period based on New-West (1987) standard errors. ***, **, and * indicate statistical significance at the 1, 5, and 10 per cent level, respectively.

Table. XIV (Continued)

Panel A. $CAR_{t,t}$

	HFOI1	HFOI2	LFOI1	LFOI2
OI_{t-1}	0.013	0.001***	0.001***	0.026***
	[0.38]	[2.77]	[5.79]	[7.24]
OI_{t-2}	-0.097***	-0.001**	-0.000***	-0.019***
	[-2.65]	[-2.17]	[-3.56]	[-5.05]
OI_{t-3}	-0.165***	-0.001***	-0.000***	-0.017***
	[-4.51]	[-2.77]	[-2.72]	[-4.41]
OI_{t-4}	-0.008	0.000	-0.000	-0.002
	[-0.25]	[0.01]	[-0.39]	[-0.48]
OI_{t-5}	-0.181***	-0.002***	0.000	-0.003
	[-5.49]	[-4.02]	[0.17]	[-0.88]
$BASPRD_{t-1}$	0.072***	0.073***	0.072***	0.072***
	[14.00]	[14.08]	[14.23]	[14.17]
$TURN_{t-1}$	0.000**	0.000*	0.000	0.000
	[2.38]	[1.82]	[1.46]	[1.33]
RET_{t-1}	-0.115***	-0.116***	-0.124***	-0.131***
	[-35.55]	[-35.57]	[-35.28]	[-33.88]
RET_{t-2}	-0.065***	-0.065***	-0.060***	-0.054***
	[-18.52]	[-18.42]	[-15.68]	[-13.02]
RET_{t-3}	-0.043***	-0.044***	-0.040***	-0.034***
	[-12.55]	[-12.56]	[-10.65]	[-8.26]
RET_{t-4}	-0.036***	-0.036***	-0.036***	-0.035***
	[-10.46]	[-10.49]	[-9.61]	[-8.52]
RET_{t-5}	-0.027***	-0.027***	-0.030***	-0.027***
	[-8.03]	[-7.83]	[-7.93]	[-6.73]
RET^2_{t-1}	0.039***	0.040***	0.042***	0.044***
	[11.51]	[11.54]	[12.24]	[12.59]
RET^2_{t-2}	0.091***	0.090***	0.082***	0.082***
	[5.14]	[5.08]	[4.63]	[4.63]
RET^2_{t-3}	-0.021**	-0.022**	-0.021**	-0.024**
	[-2.13]	[-2.16]	[-2.13]	[-2.36]
RET^2_{t-4}	0.058***	0.057***	0.058***	0.057***
	[4.64]	[4.52]	[4.59]	[4.57]
RET^2_{t-5}	0.023***	0.015*	0.018**	0.017**
	[2.94]	[1.86]	[2.33]	[2.16]
ALPHA	0.001***	0.001***	0.001***	0.001***
	[6.90]	[7.14]	[7.19]	[7.39]
ADJRSQ	0.008***	0.008***	0.008***	0.009***
	[0.01]	[0.01]	[0.01]	[0.01]
OBS	251,640	251,376	253,434	253,434

Table. XIV (Continued)

Panel B. $CAR_{t-1,t+1}$

	HFOI1	HFOI2	LFOI1	LFOI2
OI_{t-2}	0.021	0.002***	0.001***	0.038***
	[0.36]	[3.07]	[5.77]	[6.49]
OI_{t-3}	-0.174***	-0.000	-0.000**	-0.020***
	[-3.06]	[-0.27]	[-2.13]	[-3.28]
OI_{t-4}	0.004	0.000	0.000	0.003
	[0.07]	[0.74]	[0.15]	[0.56]
OI_{t-5}	-0.348***	-0.002***	-0.000	-0.007
	[-6.76]	[-3.65]	[-0.16]	[-1.24]
OI_{t-6}	-0.043	0.000	-0.001***	-0.023***
	[-0.85]	[0.67]	[-3.21]	[-3.83]
$BASPRD_{t-2}$	0.072***	0.073***	0.073***	0.073***
	[8.87]	[9.06]	[9.26]	[9.26]
$TURN_{t-2}$	-0.001***	-0.001***	-0.001***	-0.001***
	[-3.68]	[-4.25]	[-4.31]	[-4.35]
RET_{t-2}	-0.191***	-0.194***	-0.206***	-0.214***
	[-35.22]	[-35.42]	[-34.77]	[-33.25]
RET_{t-3}	-0.120***	-0.123***	-0.119***	-0.113***
	[-22.28]	[-22.56]	[-20.10]	[-17.59]
RET_{t-4}	-0.095***	-0.097***	-0.097***	-0.098***
	[-17.83]	[-17.97]	[-16.55]	[-15.37]
RET_{t-5}	-0.081***	-0.082***	-0.083***	-0.079***
	[-15.28]	[-15.34]	[-14.30]	[-12.47]
RET_{t-6}	-0.063***	-0.065***	-0.057***	-0.051***
	[-11.81]	[-12.15]	[-9.77]	[-8.00]
RET^2_{t-2}	0.167***	0.169***	0.182***	0.179***
	[6.06]	[6.13]	[6.58]	[6.48]
RET^2_{t-3}	0.011	0.014	0.012	0.011
	[0.74]	[0.93]	[0.76]	[0.70]
RET^2_{t-4}	0.036*	0.037*	0.035*	0.035*
	[1.75]	[1.80]	[1.74]	[1.73]
RET^2_{t-5}	0.054***	0.041***	0.041***	0.039***
	[4.43]	[3.33]	[3.38]	[3.17]
RET^2_{t-6}	0.052***	0.053***	0.050***	0.049***
	[4.88]	[4.92]	[4.66]	[4.51]
ALPHA	0.002***	0.002***	0.002***	0.002***
	[8.81]	[8.79]	[8.87]	[8.98]
ADJRSQ	0.010**	0.010***	0.010**	0.010**
	[0.01]	[0.01]	[0.01]	[0.01]
OBS	246,890	246,617	248,656	248,656

Table. XIV (Continued)

Panel C. $CAR_{t-2,t+2}$

	HFOI1	HFOI2	LFOI1	LFOI2
OI_{t-3}	-0.013	0.002***	0.001***	0.031***
	[-0.21]	[2.73]	[3.86]	[4.70]
OI_{t-4}	-0.003	0.001	0.000	0.006
	[-0.05]	[1.03]	[1.33]	[0.87]
OI_{t-5}	-0.298***	-0.002***	-0.000	-0.013**
	[-5.28]	[-3.27]	[-0.37]	[-2.02]
OI_{t-6}	-0.009	-0.000	-0.000*	-0.020***
	[-0.16]	[-0.01]	[-1.68]	[-2.95]
OI_{t-7}	-0.052	-0.000	-0.001***	-0.027***
	[-0.85]	[-0.12]	[-3.57]	[-4.06]
$BASPRD_{t-3}$	0.091***	0.092***	0.091***	0.090***
	[10.34]	[10.49]	[10.55]	[10.50]
$TURN_{t-3}$	-0.001***	-0.001***	-0.001***	-0.001***
	[-4.45]	[-4.83]	[-4.87]	[-4.86]
RET_{t-3}	-0.210***	-0.213***	-0.223***	-0.230***
	[-35.81]	[-35.99]	[-34.64]	[-32.94]
RET_{t-4}	-0.161***	-0.164***	-0.166***	-0.166***
	[-27.61]	[-27.77]	[-25.93]	[-23.82]
RET_{t-5}	-0.123***	-0.124***	-0.125***	-0.117***
	[-21.37]	[-21.25]	[-19.61]	[-16.92]
RET_{t-6}	-0.106***	-0.107***	-0.102***	-0.094***
	[-18.25]	[-18.22]	[-15.96]	[-13.43]
RET_{t-7}	-0.087***	-0.088***	-0.076***	-0.070***
	[-14.91]	[-14.90]	[-12.02]	[-10.13]
RET^2_{t-3}	0.122***	0.126***	0.129***	0.131***
	[7.32]	[7.52]	[7.76]	[7.84]
RET^2_{t-4}	0.114***	0.116***	0.114***	0.112***
	[5.29]	[5.39]	[5.31]	[5.21]
RET^2_{t-5}	0.096***	0.084***	0.085***	0.082***
	[7.14]	[6.30]	[6.35]	[6.08]
RET^2_{t-6}	0.076***	0.076***	0.075***	0.073***
	[6.43]	[6.41]	[6.34]	[6.13]
RET^2_{t-7}	-0.037*	-0.036	-0.041*	-0.039*
	[-1.70]	[-1.64]	[-1.87]	[-1.79]
ALPHA	0.002***	0.002***	0.002***	0.002***
	[6.53]	[6.48]	[6.61]	[6.76]
ADJRSQ	0.012**	0.012**	0.012**	0.012**
	[0.01]	[0.01]	[0.01]	[0.01]
OBS	251,583	251,295	253,355	253,355

Table XV
Return predictability at market level

This table shows the results of Fama-MacBeth regression to measure return predictability of HFOIs and LFOIs at market level. The sample period is from 01 January 1993 to 31 December 2013. The first row represents which kind of lagged OI uses in a regression model. High frequency order imbalances (hereafter, HFOIs) are obtained from Trades and Quotes (TAQ) dataset. Low frequency stock characteristics come from the Center for Research in Security Prices (CRSP) at daily level, including stock returns, the number of shares outstanding, stock prices, bid and ask close price, and trading volume. HFOI1 and HFOI2 are defined as order imbalances (hereafter, OIs) divided by the number of shares outstanding. We estimate the OIs in HFOI1 by using Lee-Ready (1991) algorithm, while we estimate OIs in HFOI2 following Easley, Lopez de Prado, and O'Hara (2011). LFOI stands for Low Frequency Order Imbalance. LFOI1 is the sign of daily returns multiplied by daily turnover rates for each stock. LFOI2 is daily returns multiplied by daily turnover rate for each stock. TURN is daily averaged turnover rate over NYSE, AMEX, and NASDAQ, and turnover rate is defined as trading volume over the number of shares outstanding. BASPRD is the adjusted and averaged bid-ask spread over NYSE, AMEX, and NASDAQ. VWRET is daily value-weighted market returns. VWRET^2 is daily squared valued-weighted market returns. SPRET is daily S&P returns. SPRET^2 is daily squared S&P returns. ALPHA is the intercept of regression models. ADJRSQ is adjusted R-square. OBS is the number of observation in each regression model. Panel A reports test results of regression of value-weighted market returns. Panel B reports test results of regression of S&P returns. Each subscript of variables indicates its time lag. In brackets, we report t-statistics of the average coefficient over sample period based on New-West (1987) standard errors. ***, **, and * indicate statistical significance at the 1, 5, and 10 per cent level, respectively.

Table. XV (Continued)

Panel A. Value-weighted market return				
	HFOI1	HFOI2	LFOI1	LFOI4
OI_{t-1}	1.295	0.031**	-0.001	-0.082**
	[1.34]	[2.55]	[-0.66]	[-2.01]
OI_{t-2}	-0.078	-0.001	0.001	0.024
	[-0.08]	[-0.07]	[0.92]	[0.59]
OI_{t-3}	-1.550	-0.022*	-0.001	-0.025
	[-1.54]	[-1.75]	[-1.28]	[-0.62]
OI_{t-4}	-0.069	-0.011	0.001	0.051
	[-0.07]	[-0.89]	[0.71]	[1.26]
OI_{t-5}	-0.593	-0.002	-0.001	-0.073*
	[-0.62]	[-0.17]	[-0.88]	[-1.80]
$BASPRD_{t-1}$	-0.034	-0.056	-0.037	-0.032
	[-0.66]	[-1.05]	[-0.73]	[-0.62]
$TURN_{t-1}$	-0.002**	-0.002*	-0.001	-0.001
	[-2.13]	[-1.93]	[-1.64]	[-1.58]
$VWRET_{t-1}$	-0.048***	-0.061***	-0.016	0.038
	[-2.76]	[-3.49]	[-0.57]	[1.01]
$VWRET_{t-2}$	-0.046***	-0.047***	-0.066**	-0.063*
	[-2.60]	[-2.68]	[-2.34]	[-1.70]
$VWRET_{t-3}$	0.021	0.024	0.038	0.029
	[1.18]	[1.34]	[1.33]	[0.78]
$VWRET_{t-4}$	-0.006	0.005	-0.027	-0.052
	[-0.33]	[0.28]	[-0.94]	[-1.41]
$VWRET_{t-5}$	-0.047***	-0.049***	-0.031	0.009
	[-2.77]	[-2.88]	[-1.10]	[0.25]
$VWRET^2_{t-1}$	1.295***	1.304***	1.229***	1.230***
	[3.04]	[3.05]	[2.89]	[2.89]
$VWRET^2_{t-2}$	-0.380	-0.368	-0.447	-0.435
	[-0.90]	[-0.87]	[-1.07]	[-1.04]
$VWRET^2_{t-3}$	0.939**	0.946**	0.863**	0.873**
	[2.20]	[2.22]	[2.02]	[2.05]
$VWRET^2_{t-4}$	-0.733*	-0.770*	-0.783*	-0.734*
	[-1.79]	[-1.88]	[-1.91]	[-1.79]
$VWRET^2_{t-5}$	0.277	0.249	0.268	0.198
	[0.68]	[0.61]	[0.65]	[0.48]
MON	0.002*	0.001	0.001	0.001
.	[1.90]	[1.44]	[1.36]	[1.31]
TUE	0.002**	0.002**	0.001**	0.001**
.	[2.57]	[2.16]	[2.01]	[1.97]
WED	0.002**	0.002**	0.001*	0.001*
.	[2.37]	[2.06]	[1.83]	[1.82]
THR	0.002**	0.002*	0.001*	0.001*
.	[2.20]	[1.88]	[1.78]	[1.78]
FRI	0.002**	0.002*	0.001*	0.001*
.	[2.32]	[1.92]	[1.83]	[1.81]
ADJRSQ	0.010	0.010	0.009	0.010
OBS	5,284	5,284	5,284	5,284

Table. XV (Continued)

Panel B. S&P return

	HFOI1	HFOI2	LFOI1	LFOI4
OI_{t-1}	1.060	0.030**	0.000	-0.047
	[1.10]	[2.45]	[0.09]	[-1.33]
OI_{t-2}	0.141	0.002	0.001	0.013
	[0.14]	[0.15]	[0.87]	[0.37]
OI_{t-3}	-1.629	-0.023*	-0.001	-0.011
	[-1.62]	[-1.82]	[-0.81]	[-0.32]
OI_{t-4}	0.130	-0.010	0.001	0.048
	[0.13]	[-0.76]	[1.07]	[1.35]
OI_{t-5}	-0.850	-0.009	-0.001	-0.085**
	[-0.88]	[-0.74]	[-1.50]	[-2.41]
$BASPRD_{t-1}$	-0.022	-0.053	-0.027	-0.024
	[-0.42]	[-1.00]	[-0.52]	[-0.48]
$TURN_{t-1}$	-0.002**	-0.002**	-0.001*	-0.001*
	[-2.22]	[-2.23]	[-1.69]	[-1.66]
$SPRET_{t-1}$	-0.081***	-0.097***	-0.071***	-0.029
	[-4.71]	[-5.50]	[-2.62]	[-0.89]
$SPRET_{t-2}$	-0.057***	-0.060***	-0.072***	-0.061*
	[-3.28]	[-3.39]	[-2.68]	[-1.89]
$SPRET_{t-3}$	0.015	0.017	0.018	0.010
	[0.84]	[0.98]	[0.66]	[0.31]
$SPRET_{t-4}$	-0.011	0.001	-0.037	-0.051
	[-0.62]	[0.03]	[-1.35]	[-1.58]
$SPRET_{t-5}$	-0.042**	-0.041**	-0.017	0.018
	[-2.53]	[-2.39]	[-0.64]	[0.56]
$SPRET^2_{t-1}$	1.307***	1.328***	1.250***	1.235***
	[3.21]	[3.24]	[3.08]	[3.04]
$SPRET^2_{t-2}$	-0.315	-0.285	-0.373	-0.381
	[-0.78]	[-0.71]	[-0.93]	[-0.95]
$SPRET^2_{t-3}$	0.915**	0.935**	0.858**	0.887**
	[2.24]	[2.29]	[2.11]	[2.18]
$SPRET^2_{t-4}$	-0.821**	-0.840**	-0.863**	-0.864**
	[-2.09]	[-2.14]	[-2.19]	[-2.19]
$SPRET^2_{t-5}$	0.483	0.457	0.454	0.395
	[1.23]	[1.16]	[1.15]	[1.00]
MON	0.002**	0.002*	0.001	0.001
.	[2.18]	[1.91]	[1.59]	[1.60]
TUE	0.002***	0.002**	0.001**	0.001**
.	[2.61]	[2.43]	[1.97]	[1.97]
WED	0.002**	0.002**	0.001	0.001
.	[2.17]	[2.16]	[1.56]	[1.59]
THR	0.002**	0.002*	0.001	0.001
.	[1.99]	[1.94]	[1.47]	[1.53]
FRI	0.002*	0.002*	0.001	0.001
.	[1.92]	[1.86]	[1.34]	[1.41]
ADJRSQ	0.014	0.015	0.014	0.015
OBS	5,284	5,284	5,284	5,284