Dynamic Relations between Order Imbalance, Volatility and Return of Top Losers

Yong-Chern Su*, HanChing Huang, Po-Hsin Kuo and Peiwen Chen

Abstract

Recently, many researches show that order imbalances have a significant relationship with stock returns, especially in speculative stocks. In this paper, we examine the relations between order imbalances, volatility and stock returns of top losers. Then, we develop a trading strategy based on the relations.

First, we apply GARCH (1,1) model with and without volatility to test whether it can fit our intraday data. We find that GARCH (1,1) explains the intraday price behaviors of speculative stocks in both cases. Then, we use multi-regression model to examine whether contemporaneous and lagged order imbalances have significant influences on stock returns. We find contemporaneous order imbalances have positive effect and lagged-one order imbalances have negative effect on stock returns. While controlling for contemporaneous order imbalances, only lagged-one order imbalances have a significantly negative effect on return. We also find that order imbalance and market capitalization have an insignificantly positive relation.

We develop a trading strategy based on the previous findings to make profit. We short sell when order imbalance is negative and buy back when order imbalance is positive. The empirical results show that if we don’t truncate the volume distribution, we can’t find abnormal returns. However, if we sift our data from trading volume, that is, above 99% volume, we document a significant profit based on our trading strategy.

In order to explain the significant profit based on dynamic relations between order imbalance and return of speculative top gainers, we examine the causality relationship between return and order imbalance. We find that order imbalance is a unidirectional indicator for predicting future returns. Especially, order imbalance is an extraordinary good indicator for price discovery in small firm size quartile.

Key words: order imbalance, information asymmetry, volatility, causality relationship

* Corresponding author, Email: ycsu@ccms.ntu.edu.tw, Address: 50 Lane 144 Sec. 4, Keelung Road, Taipei, Taiwan, Tel: 8862-33661089, Fax: 8862-23637714
1. Introduction

For many years, trading volume has provided the linkage between trading activity and returns. Investors are always seeking and finding reliable useful indicators to predict the movements of their holding stocks. Among many possible indicators, it has been showed that trading activity can be a great proxy to imply some private information. Just as researches of Lo and Wang (2000) and Karpoff (1987), they find the association of trading activities and stock returns. If no-informed investors can find a good indicator and use it to trade, they may earn abnormal profits even without private information. Therefore, the objective of this study is to shed further light on the dynamic relations between trading activities, volatility and stock returns.

There are many types of trading volume: trading number, trading shares and trading dollars. In Chordia, Roll and Subrahmanyam’s paper (2002), they find that trading volume have the most efficient effect on stock returns. They also use the indicator, order imbalance, to stand for the trading volume and they think that trading volume can reveal some private information behind the market markers and the big-deal traders. According to Lee and Ready (2000), order imbalance judges the direction of each trading volume. That’s an interesting thing that we can observe the inventory situation by the adjustment of bid-ask quotes of market makers. If market makers don’t have enough inventories on hand, they may adjust the quotes by increasing the ask price to reflect his inventory level. However, in the meanwhile, they also can do the same behavior by just cheating and misleading investors. Therefore, besides observing the order imbalances, investors also need to guess what the market makers’ situations are and how they think.

We examine whether trading volume have significant influence on stock returns and how long this effect lasts for. Based on Jiang Wang (2002), investors trade for two reasons: hedging for their stock holdings or speculating on their private information. They find that when investors trade stocks for hedging their positions with large trading volume, the patterns of stock returns will reverse in following periods but if the accompanying trading volume is small, it just refers to white noise. However, when investors trade stocks because of his private information, the patterns of stock returns will continue in following periods. That’s the so-called momentum effect and once it works, then we can buy the stocks and predict the movement of their stock prices. Therefore, we want to find the speculative stocks and observe the results of our research method. If it does really work,
then we may take this finding to develop a trading strategy and expect to earn profits.

In this study, we first apply the GARCH (1,1) model to test the intraday transaction data of NASDAQ from TAQ. We use top losers to examine the model and see whether the order imbalances can really have great influence on stock returns. Second, after examining the original model, we are interested in examining the relations between volatility and order imbalances. Since investors always care more about the returns but ignore the appending risks. Third, we want to know whether contemporaneous and lagged order imbalances do have significant influence on stock returns. According to Chordia and Subrahmanyam (2004), we ran multiple regressions to test whether this relationship exists. If we can prove the lagged order imbalances have explanatory power on current stock returns, then we can use prior order imbalances to predict the patterns of stock prices.

Fourth, we are curious about whether there exist some characteristics of our sample stocks. We investigate the small firm effect by testing whether the order imbalances of small firms have more influence on stock returns than large firms do. If the answer is yes, then investors shall choose small firms instead of large firms to make abnormal returns.

Fifth, we try to develop a trading strategy from the above findings. We have keen interest in observing the following actions: to short sell the stocks when the first seller-initiated order imbalance appeared and to buy back when the order imbalances become buyer-initiated. All our transactions ignore the transaction costs and taxes. Then we sieve out the volume with 50%, 90%, 99% respectively. After we trim off small volume and leave the large volume in our strategy, we want to know whether the trading strategy can beat the original individual stock returns.

Finally, we develop a story of dynamic lead-lag relationship to explain the abnormal return from our strategy. According to Chen and Wu (1999), we define five groups of dynamic relationship, including independency, the contemporaneous relationship, unidirectional relationship and feedback relationship. To determine a specific causal relationship, we use a systematic multiple hypotheses testing method. Unlike the traditional hypothesis testing, this testing method avoids the potential bias induced by restricting the causal relationship to a single alternative hypothesis.

2. Data

We collect our sample stocks data from the Center for Research in Security Prices (CRSP) and the NYSE Trades and Automated Quotations (TAQ) databases. We first
sought for the top losers of daily transactions within CRSP during December 2005, then match the corresponding intraday trading data on TAQ. We collect 61 stocks from the databases. Fifty stocks of our data are listed in NASDAQ market, while three stocks in AMEX market and eight stocks in NYSE market and all the stocks are able to be traded in the NASDAQ market.

Stocks are included and excluded depending on the following criteria:
1. The sample shall be included in both the Center for Research in Security Prices (CRSP) and the NYSE Trades and Automated Quotations (TAQ).
2. Among all the top losers, in order to avoid the manipulation, we exclude the data whose prices are below $2 and the daily trading number below 30 transactions.
3. For those stocks with negative bid-ask spread quotations, transaction prices, and the extreme quotes (e.g. the price is higher than 1,000) were excluded.
4. Since the stock trading characteristics differ from certificates, American Depository Receipts, shares of beneficial interest, companies incorporated outside the U.S, closed-end funds, preferred stocks and REITS. Therefore, we exclude these kinds of securities.
5. If there are stock splits, reverse splits, stock dividends, repurchase on the stocks during the sample period, we also eliminate them from our sample.

Our sample period are from December 1st, 2005 to December 31st, 2005. Moreover, our observation period each day is from 9:30 A.M. to 4:00 P.M. Since the trading behavior during market time and after market time is quite different, we use the regular hour trading data.

The average order imbalance volume of our data is -203,483 shares everyday; the average buyer-initiated order imbalance volume is 666,273 shares while average seller-initiated order imbalance volume is -869,756 shares; the average mean of order imbalance volume is -120 shares everyday; the average standard deviation of the order imbalance volume is really high, reaching for 2,124 shares. We find that the seller-initiated order imbalance is larger than the buyer-initiated order imbalance. It is self-explained that we choose our data by selecting top losers and negative order imbalances push stock prices move downward.

3. Methodology

We employ two different GARCH model to examine dynamic relations of return-order
imbalance and volatility-order imbalance.

3.1. Dynamic return-order imbalance GARCH(1,1) model:

\[ R_t = \alpha + \beta \cdot OI_t + \varepsilon_t, \]  \hspace{1cm} (1)

\[ \varepsilon_t|\Omega_{t-1} \sim N(0, h_t) \]

\[ h_t = \alpha_1 + B_1 h_{t-1} + C_1 \varepsilon_{t-1}^2 \]

Where \( R_t \) is the return in period \( t \), defined as \( \ln(P_t/P_{t-1}) \)

\( OI_t \) is the explanatory variable, “Order Imbalance” on stock returns

\( \beta \) is the coefficient of the impact of “Order Imbalance” on stock returns

\( \varepsilon_t \) means the residual of the stock return in period \( t \)

\( h_t \) is the conditional variance in period \( t \)

\( \Omega_{t-1} \) is the information set in period \( t-1 \)

\( \alpha, A_1, B_1, C_1 \) are coefficients

We use an order imbalance coefficient in our conditional mean equation. The \( \beta \) coefficient represents the impact of order imbalance on stock return.

3.2. Dynamic volatility-order imbalance GARCH(1,1) model:

Since investors always care more about the returns but ignore the appending risks, we have an interest to examine the dynamic relation between volatility and order imbalance.

\[ R_t = \alpha + \varepsilon_t, \]

\[ \varepsilon_t|\Omega_{t-1} \sim N(0, h_t) \]

\[ h_t = \alpha_1 + B_1 h_{t-1} + C_1 \varepsilon_{t-1}^2 + D_1 OI_t \]

Where \( R_t \) is the return in period \( t \), defined as \( \ln(P_t/P_{t-1}) \)

\( OI_t \) is the explanatory variable, “Order Imbalance” on stock returns

\( \varepsilon_t \) means the residual of the stock return in period \( t \)

\( h_t \) is the conditional variance in period \( t \)

\( \Omega_{t-1} \) is the information set in period \( t-1 \)

\( \alpha, A_1, B_1, C_1, D_1 \) are coefficients
We use an order imbalance coefficient in our conditional volatility equation. The coefficient represents the impact of order imbalance on stock volatility.

### 3.3 Dynamic causal relation between return and order imbalance

In order to clarify the dynamic lead-lag relationship, we employ a nested causality to explore the dynamic causal relation between return and order imbalance. According to Chen and Wu (1999), we define four relationship between two random variables, \( x_1 \) and \( x_2 \), in terms of constraints on the conditional variances of \( x_{1(T+1)} \) and \( x_{2(T+1)} \) based on various available information sets, where \( x_i = (x_{i1}, x_{i2}, ..., x_{iT}) \), \( i = 1, 2 \), are vectors of observations up to time period \( T \).

**Definition 1: Independency, \( x_1 \land x_2 \):**

\( x_1 \) and \( x_2 \) are independent if

\[
Var(x_{1(T+1)} \mid x_1) = Var(x_{1(T+1)} \mid x_1, x_2) = Var(x_{1(T+1)} \mid x_1, x_2, x_{2(T+1)})
\]  

(9)

and

\[
Var(x_{2(T+1)} \mid x_2) = Var(x_{2(T+1)} \mid x_1, x_2) = Var(x_{2(T+1)} \mid x_1, x_2, x_{1(T+1)})
\]  

(10)

**Definition 2: Contemporaneous relationship, \( x_1 \leftrightarrow x_2 \):**

\( x_1 \) and \( x_2 \) are contemporaneously related if

\[
Var(x_{1(T+1)} \mid x_1) = Var(x_{1(T+1)} \mid x_1, x_2)
\]  

(11)

\[
Var(x_{1(T+1)} \mid x_1, x_2) > Var(x_{1(T+1)} \mid x_1, x_2, x_{2(T+1)})
\]  

(12)

and

\[
Var(x_{2(T+1)} \mid x_2) = Var(x_{2(T+1)} \mid x_1, x_2)
\]  

(13)

\[
Var(x_{2(T+1)} \mid x_1, x_2) > Var(x_{2(T+1)} \mid x_1, x_2, x_{1(T+1)})
\]  

(14)

**Definition 3: Unidirectional relationship, \( x_1 \rightarrow x_2 \):**

There is a unidirectional relationship from \( x_1 \) to \( x_2 \) if

\[
Var(x_{1(T+1)} \mid x_1) = Var(x_{1(T+1)} \mid x_1, x_2)
\]  

(15)

and

\[
Var(x_{2(T+1)} \mid x_2) > Var(x_{2(T+1)} \mid x_1, x_2)
\]  

(16)
Definition 4: Feedback relationship, $x_1 \leq x_2$:

There is a feedback relationship between $x_1$ and $x_2$ if

$$\text{Var}(x_{1(T+1)} \mid x_1) > \text{Var}(x_{1(T+1)} \mid x_1, x_2)$$

and

$$\text{Var}(x_{2(T+1)} \mid x_2) > \text{Var}(x_{2(T+1)} \mid x_1, x_2)$$

To explore the dynamic relationship of a bi-variate system, we form the five statistical hypotheses in the Figure 7 where the necessary and sufficient conditions corresponding to each hypothesis are given in terms of constraints on the parameter values of the VAR model.

To determine a specific causal relationship, we use a systematic multiple hypotheses testing method. Unlike the traditional pair-wise hypothesis testing, this testing method avoids the potential bias induced by restricting the causal relationship to a single alternative hypothesis. To implement this method, we employ results of several pair-wise hypothesis tests. For instance, in order to conclude that $x_1 = x_2$, we need to establish that $x_1 \neq x_2$ and to reject that $x_1 \neq x_2$. To conclude that $x_1 < x_2$, we need to establish that $x_1 < x_2$ as well as $x_1 \neq x_2$ and also to reject $x_1 \wedge x_2$. In other words, it is necessary to examine all five hypotheses in a systematic way before we draw a conclusion of dynamic relationship. The following presents an inference procedure that starts from a pair of the most general alternative hypotheses.

Our inference procedure for exploring dynamic relationship is based on the principle that a hypothesis should not be rejected unless there is sufficient evidence against it. In the causality literature, most tests intend to discriminate between independency and an alternative hypothesis. The primary purpose of the literature cited above is to reject the independency hypothesis. On the contrary, we intend to identify the nature of the relationship between two financial series. The procedure consists of four testing sequences, which implement a total of six tests (denoted as (a) to (f)), where each test examines a pair of hypotheses.

The four testing sequences and six tests are summarized in a decision-tree flow chart in Figure 8. The inference procedure starts from executing tests (a) and (b), which result in one of the four possible outcomes, $E_1$, $E_2$, or $E_4$. The three outcomes, $E_1$, $E_2$, and $E_3$, that lead to the conclusions of $x_1 \leq x_2$, $x_1 = x_2$, and $x_1 \leq x_2$, respectively, will stop the procedure at the end of the first step. Nonetheless, when outcome $E_4$ is realized, tests (c)
and (d) will be implemented. There again one of the four possible outcomes, E₅, ..., or E₈, will be realized. The realization of outcomes E₅ and E₆, which respectively indicates $x₁ \leq x₂$, and $x₁ > x₂$, will stop the procedure at the end of Step 2. On the other hand, the realization of outcome E₇ would lead to test (e) in Step 3, which has the consequence of either outcome E₉ or outcome E₁₀. Outcome E₉ implies $x₁ \leq x₂$ and the procedure will stop. Either outcome E₈ from Step 2 or outcome E₁₀ from Step 3 will lead to test (f) in Step 4. This last step may generate two possible results, E₁₁ and E₁₂, which imply $x₁ < x₂$ and $x₁ \land x₂$, respectively.

4. Empirical Results

4.1 Dynamic relations between order imbalances and returns

We examine the dynamic relations between order imbalances and stock returns. Table 1 exhibits the empirical results. We find that mean of the coefficients of order imbalances is 7.39E-05, with a variance of the coefficients of order imbalances is 2.15E-07. From Panel A approximately 70% of the t values of order imbalances are significantly positive relation with the stock returns under 90%, 95% and 99% confidence level. This result is consistent with the daily results of previous studies. We document that in our intraday study with time varying model, order imbalance has a significantly positive relationship with stock return.

We also draw the distribution of the coefficients of the order imbalances, trying to find some characteristics of this proxy. In Panel B, we find an interesting phenomenon that about 95% coefficients of our 61 sample stocks are distributed in one tail. We attribute this situation to the property of our sample stocks. Since we select the daily top losers, the empirical results are self-explained.

4.2 Dynamic relations between order imbalances and volatility

We also have interested in examining the relation between volatility and order imbalances. We expect that high order imbalances are accompanied by large volatility.

The empirical results are exhibited in Table 2. It is intuitively that higher order imbalances cause higher volatility. Therefore, we expect to see a significantly positive relation between volatility and order imbalances. However, in Panel A of Table 2, we find about half of our samples have negative relation between volatility and order imbalances. We have two potential explanations:
1. The investors behavior

From the perspective of Kahneman and Tversky (1979), it explains how people make choices in situations where they have to decide between alternatives that involving risk. Given the same variation in value, there is a bigger impact of losses than that of gains. Therefore, investors tend to hold their stocks when stock price going up, but tend to overreact and sell them in panic. As a result, the negative relation between volatility and order imbalances of the half stocks may be attributed to the investors’ irrational behaviors.

2. The leverage effect

According to Christie (1982), the leverage effect refers to the well-established relationship between stock returns and both implied and realized volatility: volatility increases when the stock price falls. A standard explanation ties the phenomenon to the effect that a change in market valuation of a firm's equity has on the degree of leverage in its capital structure, with an increase in leverage producing an increase in stock volatility. That is, when stock price declines, then the market capitalization of a company drops off and the debt to equity ratio increases. Since the leverage gets highly geared, volatility has negative relationship with order imbalances.

4.3. Conditional contemporaneous return-order imbalance relation

From above, we know that current order imbalances do really have a significant impact on the stock returns. We want to know whether the previous order imbalances can also have influence on current stock returns, and if the answer is yes, how long do this previous order imbalances last for?

In Panel A of Table 3, we find that most of the contemporaneous order imbalances (above 80%) have positive influences on current stock returns and above 80% of lagged –one order imbalances have negative effect on current stock returns. The intercept, lagged two to four order imbalances don’t have such significant influences on current stock returns. This result is consistent with Chordia and Subrahmanyam (2004). We can’t add another additional information in our intraday study.

The contemporaneous relation between order imbalances and returns is consistent with both the inventory and asymmetry information stories of price formation. The negative coefficients of lagged-one imbalances are attributed to the reason that the most information on current stock returns is explained by contemporaneous order imbalance and auto-correlated lagged order imbalances cause reverse impacts on current stock
returns.

We compare the Panel B of Table 1 with Panel B of Table 3. The frequencies of the coefficients of contemporaneous order imbalances of these two models are similar. They both are concentrated in one tail. Again, the characteristics of our top loser stocks explain the empirical results.

4.4. Unconditional lagged return-order imbalance relation

In this section, we try to control the contemporaneous order imbalances and include the lagged-five order imbalances in this multiple regression model. We want to test whether lagged order imbalances have impacts on the current stock returns. If the relation between stock returns and lagged order imbalances is significant, then we can use this result to develop trading strategy.

In Panel A of Table 4, we find that only lagged-one order imbalances have significant effect on current stock returns and the relationship between these two variables is negative. This results is different from the findings of Chordia and Subrahmanyam (2004). They argue that about 77% of the coefficients on first lag of order imbalances are positive and significant. Their result indicates a predictive relation between lagged-one order imbalances and current stock returns. However, our result is quite different. It implies that controlling for the current period order imbalances, lagged-one order imbalances is negatively related to current returns and the price pressure reverse. We attribute this situation to three reasons as follows:

1. The data which Chordia employed in his empirical tests is daily data, while our data here is intraday data. It is intuitional that since the time lag interval in intraday data is much shorter than that in the daily data. Every period of intraday data is too short to reveal information timely.

2. Since we select data by choosing top losers in daily trades and top losers are speculative stocks. There is a momentum effect on the speculative stocks, based on Llorente, Michaely, Saar and Wang(2002). Speculative trades tend to continue themselves. Therefore, when order imbalances are positive and large, then the price will go up again. When price goes up, then the rate of return will decrease. As a result, the relation between order imbalances and stock returns is negative and significant.

3. There goes another possible explanation. Market makers have the responsibility to control and maintain the stability of stock markets. Therefore, when market makers
observed a huge positive order, they though this trade had private information and will keep the bid-ask quote down to prevent discretionary traders from manipulating the stock price. However, if bid-ask price is pressed lower, traders have a good opportunity to buy stocks with a lower price and the inventories of market makers will be reduced. It is quite dangerous that market makers have no inventories on hand because they have a good chance to be cornered. Once market makers were cornered, discretionary traders won and the stock price continued to go up. This is the unique characteristic of speculative stocks. Apparently, our sample stocks are top losers and they are speculative. However, the sample data which Chordia and Subrahmanyam employ include all stocks.

4.5. Small firm effect

It is plausible that inventory pressures from discretionary traders are different among big and small caps. We expected a significantly negative relationship between market capitalization and order imbalances, namely a small firm effect. Intuitively, small caps tend to be easier manipulated than big caps.

Panel A of Table 5 shows that the coefficient is slightly positive and not significant under 90%, 95% and 99% confidence level. We can not conclude that there exists a small firm effect between return and order imbalance. Instead of using original market capitalization, we substitute logged market cap for our independent variable. Since logged variables can give the regression more economic meanings. In Panel B of Table 5, we find that although the t value is larger. The relationship is still insignificant.

4.6 Trading Strategy based on return-order imbalance relation

Since there is evidence that contemporaneous order imbalances and current stock returns have positive relationship. A naive question has been raised whether we can develop a profitable trading strategy based on this result.

We first calculate the average return of our sample stocks. The average return of our 61 top losers is -18.76%. Then, we form two kinds of trading strategies, including a trading price basis and a bid-ask quote basis. In our trading strategy, we ignore the transaction costs and taxes. We short sell the stocks when the first corresponding negative order imbalances appeared and buy back the stocks when the first corresponding positive order imbalance show up. We trade this strategy based on four scenarios: no truncation, 50% truncation, 90% truncation and 99% truncation.

In Panel A of Table 6, we show the average return of each order imbalance truncation
trading strategy. The mean of no truncation, 50%, 90%, 99% truncations are -77.38%, -26.19%, -10.58% and 1.06%, respectively. Although the first three strategies have negative returns and even the first two are worse than the original average return, we observe the trend that when trimming the smaller order imbalances, the strategy yield a higher average return. When trimming off the 99% smaller order imbalances, the average return even becomes positive. This result is amazing and we document a successful trading strategy that turns daily top losers, with an average return of -18.76, to a positive return.

In Panel A of Table 6, we replace our trading strategy with a bid-ask quote basis. The Mean of no truncation, 50%, 90%, and 99% truncations are -167.85%, -35.26%, -12.39% and 2.59%, respectively. Again, we find that the first three scenarios are worse than the return of using transaction price. We attribute this result to the bid ask spread which market markers earned. Since investors buy stocks at ask price and sell stocks at bid price, bid-ask spread is market maker’s profit.

We use the paired comparisons test to see whether the order imbalance truncation trading strategy is better than no truncation trading strategy. We use no truncation and 99%-truncated strategy in our hypothesis testing:

\[
H_0: \mu_1 \geq \mu_2
\]

\[
H_1: \mu_1 < \mu_2
\]

Where \( \mu_1 \) is the mean of the no truncation rate of return

\( \mu_2 \) is the mean of the 99%-truncated rate of return

While the t-statistic with n-1 degree of freedom is computed as follows:

\[
t = \frac{\bar{d}}{S_{\bar{d}}}
\]

In the Panel B of Table 6 and Table 7, the t values of one-paired test are all significant. Therefore, we conclude that the 99%-truncated trading strategy earned a significant return than non-truncated one.

**4.7 Return-order imbalance causality relationship in explaining the successful trading strategy**

To explore the reason why a truncated order imbalance trading strategy earns a significant abnormal return, we employ a nested causality approach. In order to investigate a dynamic relationship between two variables, we impose the constraints in
the upper panel of Figure 1 on the VAR model. In Table 8, we present the empirical results of tests of hypotheses on the dynamic relationship in Figure 2. Panel A presents results for the entire sample. In the entire sample, we show that a unidirectional relationship from returns to order imbalances is 6.56% of the sample firms for the entire sample, while a unidirectional relationship from order imbalances to returns is 36.07%. The percentage of firms that fall into the independent category is 13.11%. Moreover, 31.15% of firms exhibit a contemporaneous relationship between returns and order imbalances. Finally, 13.11% of firms show a feedback relationship between returns and order imbalances. The percentage of firms carrying a unidirectional relationship from order imbalances to returns is almost six time than that from returns to order imbalances, suggesting that order imbalance is a good indicator for predicting future returns. It is consistent with many articles, which document that future daily returns could be predicted by daily order imbalances (Brown, Walsh, and Yuen (1997); Chordia and Subrahmanyam (2004)). In addition, the percentage of firms exhibiting a contemporaneous relationship is over twice than that reflecting a feedback relationship, indicating that the interaction between returns and order imbalances on the current period is larger than that over the whole period.

In order to provide the evidence showing the impact on the relation between returns and order imbalances, in Panels B, we divide firms into three groups according to the firm size. Then we test the multiple hypotheses of the relationship between returns and order imbalances. The results in Panel B indicate that the unidirectional relationship from order imbalances to returns is 40.00% in the small firm size quartile, while the corresponding number is 30.00% in the large firm size quartile during the entire sample period. The size-stratified results can be explained as follows. When the firm size is smaller, the percentage of firms exhibiting a unidirectional relationship from order imbalances to returns is larger, indicating that order imbalance is a better indicator for predicting returns in small firm size quartile.

5. Conclusions

With a view toward better understanding how stock prices move, many former researches have extensively explored the relation between trading activities and stock returns. We use order imbalances as a proxy for trading activities. This study undertakes an analysis of the relation between order imbalances and stock returns.
In this study, we first apply GARCH (1,1) model to examine how prices react to order imbalances. We find that contemporaneous order imbalances have positively significant influences on current stock returns no matter whether we add the volatility factor in the model or not.

Further, we extend our time interval to a longer horizon. We find contemporaneous order imbalances are strongly related to contemporaneous stock returns, but the relation between the lagged two to four order imbalances and current stock returns is not significant. After controlling for the contemporaneous order imbalances, the positive relationship even disappear and only lagged-one order imbalances have significantly negative impacts on current stock returns.

We test the relation between market capitalization and order imbalances to investigate whether there exists a small firm effect. The test help us to judge whether firms with some specific characteristics tend to have more significant order imbalances than other stocks without this characteristics. According to Jiang Wang (2002), small firm are easier to be influenced by informed traders. It is intuitional that the stock prices of firms with smaller capitalization are easier to be affected by adjusting the trading volume of those stocks, while the stock prices of firms with larger capitalization are more difficult to be manipulated. However, in our empirical test, we fail to find significant relationship between the market capitalization of our firms and their current order imbalance coefficients.

We develop a successful order imbalance truncated trading strategy based upon our empirical findings. Since our samples are daily top losers. Our strategy is to short sell when the first negative order imbalances appeared and buy back when the order imbalances become buyer-initiated. All actions ignore the transaction costs and taxes. No matter what kinds of scenarios we choose, by trading price or by bid-ask price, we all can attain abnormal returns by trimming 99% volumes. Our order imbalance truncated trading strategy yields a significant return.

In order to explore the reason why our order imbalance truncated trading strategy earns a significant return, we further investigate the dynamic causality relation between return and order imbalance. According to the nested causality empirical results, we find that order imbalance is a good indicator for predicting future returns. Moreover, order imbalance is a better indicator for price discovery in small firm size quartile.
References


## Figure 1. Hypotheses on the Dynamic Relationship of a Bivariate System

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>The VAR test</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₁ : ( x₁ \land x₂ )</td>
<td>( \varphi_{12} (L)= \varphi_{21} (L)=0 ), and ( \sigma_{12}=\sigma_{21}=0 )</td>
</tr>
<tr>
<td>H₂ : ( x₁ \leftarrow \rightarrow x₂ )</td>
<td>( \varphi_{12} (L)= \varphi_{21} (L)=0 )</td>
</tr>
<tr>
<td>H₃ : ( x₁ \nRightarrow x₂ )</td>
<td>( \varphi_{21} (L)=0 )</td>
</tr>
<tr>
<td>H₃* : ( x₂ \nRightarrow x₁ )</td>
<td>( \varphi_{12} (L)=0 )</td>
</tr>
<tr>
<td>H₄ : ( x₁ \lfrown\rhorightarrow x₂ )</td>
<td>( \varphi_{12} (L)\ast \varphi_{21} (L) \neq 0 )</td>
</tr>
<tr>
<td>H₅ : ( x₁ \nRightarrow \Rightarrow x₂ )</td>
<td>( \varphi_{21} (L)=0 ), and ( \sigma_{12}=\sigma_{21}=0 )</td>
</tr>
<tr>
<td>H₆ : ( x₁ \nRightarrow \Rightarrow x₁ )</td>
<td>( \varphi_{12} (L)=0 ), and ( \sigma_{12}=\sigma_{21}=0 )</td>
</tr>
<tr>
<td>H₇ : ( x₁ \leftarrow\leftarrow \rightarrow x₂ )</td>
<td>( \varphi_{12} (L)\ast \varphi_{21} (L) \neq 0 ), and ( \sigma_{12}=\sigma_{21}=0 )</td>
</tr>
</tbody>
</table>

The bivariate VAR model:

\[
\begin{bmatrix}
\phi_{11} (L) & \phi_{12} (L) \\
\phi_{21} (L) & \phi_{22} (L)
\end{bmatrix}
\begin{bmatrix}
x_{1t} \\
x_{2t}
\end{bmatrix}
= \begin{bmatrix}
\epsilon_{1t} \\
\epsilon_{2t}
\end{bmatrix}
\]

where \( x_{1t} \) and \( x_{2t} \) are mean adjusted variables. The first and second moments of the error structure, \( \sim \epsilon_{t} = (\sim \epsilon_{1t}, \sim \epsilon_{2t})' \), are that \( E(\sim \epsilon_{t}) = 0 \), and

\[
E(\sim \epsilon_{1t} \sim \epsilon_{1t+k}) = 0 \quad \text{for} \ k \neq 0 \quad \text{and} \quad E(\sim \epsilon_{1t} \sim \epsilon_{1t+k}) = \Sigma \quad \text{for} \ k = 0, \quad \text{where} \quad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}
\]

The causal relationship are defined as follows: \( \land \) is independency; \( \leftarrow \rightarrow \) is contemporaneous relationship; \( \nRightarrow \) is negation of a unidirectional relationship; \( \lfrown\rhd \) is feedback relationship; \( \nRightarrow \Rightarrow \) is negation of a strong unidirectional relationship where \( \sigma_{12}=\sigma_{21}=0 \); and \( \leftarrow\leftarrow\leftarrow \Rightarrow \) is a strong feedback relationship where \( \sigma_{12}=\sigma_{21}=0 \).
Figure 2. Test Flow Chart of a Multiple Hypothesis Testing Procedure

Test Sequence I
(a) H₃ vs. H₄
(b) H₃ * vs. H₄

→ E₁: (a) reject H₃, (b) reject H₃ *
    (x₁<=>x₂)

→ E₂: (a) reject H₃, (b) not reject H₃ *
    (x₁=>x₂)

→ E₃: (a) not reject H₃, (b) reject H₃ *
    (x₁<=x₂)

E₄: (a) not reject H₃ *
    (b) not reject H₃ *

Test Sequence II
(c) H₂ vs. H₃
(d) H₂ vs. H₃ *

→ E₅: (c) reject H₂, (b) not reject H₂
    (x₁<=x₂)

→ E₆: (c) not reject H₂, (b) reject H₂
    (x₁=>x₂)

E₇: (c) reject H₂
    (d) reject H₂

E₈: (c) not reject H₂, (b) not reject H₂

Test Sequence III
(e) H₂ vs. H₄

→ E₉: (e) reject H₂
    (x₁<=>x₂)

→ E₁₀: (e) not reject H₂

Test Sequence IV
(f) H₁ vs. H₂

→ E₁₁: (f) reject H₁

→ E₁₂: (f) not reject H₁

(x₁ <=> x₂)

Five groups of dynamic relationship are identified: independency (∧), the contemporaneous relationship (↔), unidirectional relationship (⇒ or ⇐) and feedback relationship (<=>). To determine a specific causal relationship, we use a systematic multiple hypotheses testing method. Unlike the traditional pairwise hypothesis testing, this testing method avoids the potential bias induced by restricting the causal relationship to a single alternative hypothesis. In implementing this method, we need to employ results of several pairwise hypothesis tests. For instance, in order to conclude that x₁<=>x₂, we need to establish that x₁<≠x₂ and to reject that x₁≠x₂. To conclude that x₁<=>x₂, we need to establish that x₁<≠x₂ as well as x₁≠x₂ and also to reject x₁∧x₂. In other words, it is necessary to examine all five hypotheses in a systematic way before a conclusion of dynamic relationship can be drawn.
Table 1. Results of dynamic return-order imbalance relation

\[ R_t = \alpha + \beta \cdot OI_t + \varepsilon_t, \quad \varepsilon_t \mid \Omega_{t-1} \sim N(0, h_t) \]

\[ h_t = A_t + B_t \cdot h_{t-1} + C_t \cdot \varepsilon_{t-1}^2 \]

Where \( R_t \) is the return in period \( t \), defined as \( \ln(P_t/P_{t-1}) \)

\( OI_t \) is the explanatory variable, “Order Imbalance” on stock returns

\( \beta \) is the coefficient of the impact of “Order Imbalance” on stock returns

\( \varepsilon_t \) means the residual of the stock return in period \( t \)

\( h_t \) is the conditional variance in period \( t \)

\( \Omega_{t-1} \) is the information set in period \( t-1 \)

\( \alpha, A_t, B_t, C_t \) are coefficients

Panel A. the significance test of the coefficients of conditional variance and order imbalances

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>Critical Value</th>
<th>B(1)</th>
<th>Spill</th>
<th>B(1)</th>
<th>Spill</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>&gt; 1.645</td>
<td>61</td>
<td>44</td>
<td>100%</td>
<td>72%</td>
</tr>
<tr>
<td>95%</td>
<td>&gt; 1.96</td>
<td>61</td>
<td>42</td>
<td>100%</td>
<td>69%</td>
</tr>
<tr>
<td>99%</td>
<td>&gt; 2.33</td>
<td>61</td>
<td>42</td>
<td>100%</td>
<td>69%</td>
</tr>
<tr>
<td>90%</td>
<td>&lt; -1.645</td>
<td>0</td>
<td>3</td>
<td>0%</td>
<td>5%</td>
</tr>
<tr>
<td>95%</td>
<td>&lt; -1.96</td>
<td>0</td>
<td>3</td>
<td>0%</td>
<td>5%</td>
</tr>
<tr>
<td>99%</td>
<td>&lt; -2.33</td>
<td>0</td>
<td>3</td>
<td>0%</td>
<td>5%</td>
</tr>
</tbody>
</table>

Panel B the Distribution of the coefficients of Contemporaneous order imbalances

![Frequency Distribution](image-url)
Table 2. Results of dynamic volatility-order imbalance relation

\[ R_t = \alpha + \varepsilon_t, \quad \varepsilon_t \mid \Omega_{t-1} \sim N(0, h_t) \]

\[ h_t = A_1 + B_1 h_{t-1} + C_1 \varepsilon_{t-1}^2 + D_1 OI_t \]

Where  
- \( R_t \) is the return in period t, defined as \( \ln(P_t/P_{t-1}) \)
- \( OI_t \) is the explanatory variable, “Order Imbalance” on stock returns
- \( \varepsilon_t \) means the residual of the stock return in period t
- \( h_t \) is the conditional variance in period t
- \( \Omega_{t-1} \) is the information set in period t-1
- \( \alpha, A_1, B_1, C_1, D_1 \) are coefficients

Panel A the significance test of the coefficients of conditional variance and order imbalances

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>Critical Value</th>
<th>B(1)</th>
<th>Spill</th>
<th>B(1)</th>
<th>Spill</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>&gt; 1.645</td>
<td>61</td>
<td>26</td>
<td>100%</td>
<td>43%</td>
</tr>
<tr>
<td>95%</td>
<td>&gt; 1.96</td>
<td>61</td>
<td>25</td>
<td>100%</td>
<td>41%</td>
</tr>
<tr>
<td>99%</td>
<td>&gt; 2.33</td>
<td>60</td>
<td>24</td>
<td>98%</td>
<td>39%</td>
</tr>
<tr>
<td>90%</td>
<td>&lt; -1.645</td>
<td>0</td>
<td>26</td>
<td>0%</td>
<td>43%</td>
</tr>
<tr>
<td>95%</td>
<td>&lt; -1.96</td>
<td>0</td>
<td>25</td>
<td>0%</td>
<td>41%</td>
</tr>
<tr>
<td>99%</td>
<td>&lt; -2.33</td>
<td>0</td>
<td>24</td>
<td>0%</td>
<td>39%</td>
</tr>
</tbody>
</table>
Table 3. Empirical results of conditional order imbalance regressions—lagged 0 through lagged 4

**Panel A Regression of Order Imbalance (OI) on return—lagged 0 through lagged 4**

\[ R_t = \alpha + \beta_1 OI_t + \beta_2 OI_{t-1} + \beta_3 OI_{t-2} + \beta_4 OI_{t-3} + \beta_5 OI_{t-4} + \epsilon, \]

Where \( R_t \) is the stock return in period \( t \), defined as \( \ln(P_t/P_{t-1}) \)

\( OI_t \) is the current order imbalance in period \( t \)

\( OI_{t-i}, i=1,2,3,4 \) are lagged order imbalance at time \( t-1, t-2, t-3, t-4 \) of the stock

\( \beta_i, i=1,2,3,4,5 \) are the coefficients of the impact of the current and lagged order imbalances

\( \epsilon_t \) is the residual of stock return in period \( t \)

1. **Significance test result of contemporaneous order imbalance (in numbers)**

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>Critical Value</th>
<th>intercept</th>
<th>( (OI)_t )</th>
<th>( (OI)_{t-1} )</th>
<th>( (OI)_{t-2} )</th>
<th>( (OI)_{t-3} )</th>
<th>( (OI)_{t-4} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>&gt;1.645</td>
<td>1</td>
<td>55</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>95%</td>
<td>&gt;1.96</td>
<td>1</td>
<td>50</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>99%</td>
<td>&gt;2.33</td>
<td>0</td>
<td>49</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>90%</td>
<td>&lt;-1.645</td>
<td>4</td>
<td>0</td>
<td>54</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>95%</td>
<td>&lt;-1.96</td>
<td>4</td>
<td>0</td>
<td>52</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>99%</td>
<td>&lt;-2.33</td>
<td>1</td>
<td>0</td>
<td>49</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

2. **Significance test of contemporaneous order imbalance (in percentage)**

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>Critical Value</th>
<th>intercept</th>
<th>( (OI)_t )</th>
<th>( (OI)_{t-1} )</th>
<th>( (OI)_{t-2} )</th>
<th>( (OI)_{t-3} )</th>
<th>( (OI)_{t-4} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>&gt;1.645</td>
<td>1.64%</td>
<td>90.16%</td>
<td>1.64%</td>
<td>3.28%</td>
<td>3.28%</td>
<td>8.20%</td>
</tr>
<tr>
<td>95%</td>
<td>&gt;1.96</td>
<td>1.64%</td>
<td>81.97%</td>
<td>1.64%</td>
<td>0.00%</td>
<td>3.28%</td>
<td>4.92%</td>
</tr>
<tr>
<td>99%</td>
<td>&gt;2.33</td>
<td>0.00%</td>
<td>80.33%</td>
<td>1.64%</td>
<td>0.00%</td>
<td>1.64%</td>
<td>1.64%</td>
</tr>
<tr>
<td>90%</td>
<td>&lt;-1.645</td>
<td>8.20%</td>
<td>0.00%</td>
<td>88.52%</td>
<td>8.20%</td>
<td>8.20%</td>
<td>8.20%</td>
</tr>
<tr>
<td>95%</td>
<td>&lt;-1.96</td>
<td>6.56%</td>
<td>0.00%</td>
<td>85.25%</td>
<td>4.92%</td>
<td>3.28%</td>
<td>4.92%</td>
</tr>
<tr>
<td>99%</td>
<td>&lt;-2.33</td>
<td>1.64%</td>
<td>0.00%</td>
<td>80.33%</td>
<td>1.64%</td>
<td>3.28%</td>
<td>1.64%</td>
</tr>
</tbody>
</table>
Table 4. Empirical results of unconditional order imbalance regressions—lagged 1 through lagged 5

Panel A Regression of Order Imbalance (OI) on return—lagged 1 through lagged 5

\[ R_t = \alpha + \beta_1 OI_{t-1} + \beta_2 OI_{t-2} + \beta_3 OI_{t-3} + \beta_4 OI_{t-4} + \beta_5 OI_{t-5} + \epsilon_t \]

Where \( R_t \) is the stock return in period t, defined as \( \ln(P_t/P_{t-1}) \)

\( OI_t \) is the current order imbalance in period t

\( OI_{t-i}, i=1,2,3,4,5 \) are lagged order imbalance at time t-1, t-2, t-3, t-4, t-5 of the stock

\( \beta_i, i=1,2,3,4,5 \) are the coefficients of the impact of the current and lagged order imbalances

\( \epsilon_t \) is the residual of stock return in period t

1. Significance test result of lagged order imbalance

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>Critical Value</th>
<th>intercept</th>
<th>(OI)t-1</th>
<th>(OI)t-2</th>
<th>(OI)t-3</th>
<th>(OI)t-4</th>
<th>(OI)t-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>&gt; 1.645</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>95%</td>
<td>&gt; 1.96</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>99%</td>
<td>&gt; 2.33</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>90%</td>
<td>&lt; -1.645</td>
<td>17</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>95%</td>
<td>&lt; -1.96</td>
<td>11</td>
<td>50</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>99%</td>
<td>&lt; -2.33</td>
<td>4</td>
<td>47</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

2. Significance test result of lagged order imbalance

<table>
<thead>
<tr>
<th>Confidence Level</th>
<th>Critical Value</th>
<th>intercept</th>
<th>(OI)t-1</th>
<th>(OI)t-2</th>
<th>(OI)t-3</th>
<th>(OI)t-4</th>
<th>(OI)t-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>90%</td>
<td>&gt; 1.645</td>
<td>1.64%</td>
<td>1.59%</td>
<td>3.17%</td>
<td>4.76%</td>
<td>7.94%</td>
<td>3.17%</td>
</tr>
<tr>
<td>95%</td>
<td>&gt; 1.96</td>
<td>1.64%</td>
<td>1.59%</td>
<td>0.00%</td>
<td>1.59%</td>
<td>7.94%</td>
<td>3.17%</td>
</tr>
<tr>
<td>99%</td>
<td>&gt; 2.33</td>
<td>0.00%</td>
<td>1.59%</td>
<td>0.00%</td>
<td>1.59%</td>
<td>4.76%</td>
<td>3.17%</td>
</tr>
<tr>
<td>90%</td>
<td>&lt; -1.645</td>
<td>27.87%</td>
<td>82.54%</td>
<td>1.59%</td>
<td>3.17%</td>
<td>4.76%</td>
<td>1.59%</td>
</tr>
<tr>
<td>95%</td>
<td>&lt; -1.96</td>
<td>18.03%</td>
<td>79.37%</td>
<td>1.59%</td>
<td>3.17%</td>
<td>1.59%</td>
<td>0.00%</td>
</tr>
<tr>
<td>99%</td>
<td>&lt; -2.33</td>
<td>6.56%</td>
<td>77.05%</td>
<td>1.59%</td>
<td>0.00%</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>
Table 5. Small Firm Effect Testing result

Panel A. Relation between Market Capitalization and Order Imbalance Coefficient

\[ \alpha_i = \theta_0 + \theta_1 (MarketCap_i) + \epsilon_i \]

Where \( \alpha_i \) is the coefficients of the order imbalances of each stock, \( \theta_0 \) is the intercept.

\( \theta_1 \) is the coefficient of the market cap of each stock, \( \epsilon_i \) is the residual of the stock.

<table>
<thead>
<tr>
<th>Regression Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t value</th>
<th>P value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.60033E-05</td>
<td>6.67E-05</td>
<td>0.839387</td>
<td>-7.8E-05</td>
<td>5.60033E-05</td>
</tr>
<tr>
<td>X1</td>
<td>2.74523E-08</td>
<td>4.58E-08</td>
<td>0.599096</td>
<td>-6.4E-08</td>
<td>2.74523E-08</td>
</tr>
</tbody>
</table>

Panel B. Relation between Log Market Capitalization and Order Imbalance Coefficient

\[ \alpha_i = \theta_0 + \theta_1 \ln(MarketCap_i) + \epsilon_i \]

Where \( \alpha_i \) is the coefficients of the order imbalances of each stock, \( \theta_0 \) is the intercept.

\( \theta_1 \) is the coefficient of the market cap of each stock, \( \epsilon_i \) is the residual of the stock.

<table>
<thead>
<tr>
<th>Regression Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t value</th>
<th>P value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.81766E-04</td>
<td>0.000226</td>
<td>0.42363</td>
<td>-1.81766E-04</td>
<td>0.000226</td>
</tr>
<tr>
<td>X1</td>
<td>4.71684E-05</td>
<td>4.02E-05</td>
<td>1.174519</td>
<td>4.71684E-05</td>
<td>4.02E-05</td>
</tr>
</tbody>
</table>
Table 6. Return from Speculative Trading Strategy under Truncated Order Imbalance Distribution (trading price basis)

Panel A. Average Return of All Sample Stocks under 0%, 50%, 90% and 99% order imbalance truncations

<table>
<thead>
<tr>
<th>Average Daily Return</th>
<th>Not truncated</th>
<th>50% truncated</th>
<th>90% truncated</th>
<th>99% truncated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Return</td>
<td>-18.76%</td>
<td>-77.38%</td>
<td>-26.19%</td>
<td>-10.58%</td>
</tr>
</tbody>
</table>

Panel B. Significant test on with and without speculative trading strategies

<table>
<thead>
<tr>
<th></th>
<th>X1</th>
<th>X2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.7738</td>
<td>0.0106</td>
</tr>
<tr>
<td>Variance</td>
<td>1.1452</td>
<td>0.0065</td>
</tr>
<tr>
<td>Numbers</td>
<td>61</td>
<td>61</td>
</tr>
<tr>
<td>Pearson Correlation Coefficient</td>
<td>0.1986</td>
<td></td>
</tr>
<tr>
<td>Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Degree of Freedom</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>t value</td>
<td>-5.7951</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tailed</td>
<td>1.35E-07</td>
<td></td>
</tr>
<tr>
<td>Critical Value: one-tailed</td>
<td>1.6706</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tailed</td>
<td>2.69E-07</td>
<td></td>
</tr>
<tr>
<td>Critical Value: two-tailed</td>
<td>2.0003</td>
<td></td>
</tr>
</tbody>
</table>
Table 7. Return from Speculative Trading Strategy under Truncated Order Imbalance Distribution (bid-ask basis)

Panel A Average Return of All Sample Stocks under 0%, 50%, 90% and 99% order imbalance truncations

<table>
<thead>
<tr>
<th>Average Daily Return</th>
<th>Not truncated</th>
<th>50% truncated</th>
<th>90% truncated</th>
<th>99% truncated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Return</td>
<td>-18.76%</td>
<td>-167.85%</td>
<td>-35.26%</td>
<td>-12.39%</td>
</tr>
</tbody>
</table>

Panel B Significant test on with and without speculative trading strategies

<table>
<thead>
<tr>
<th></th>
<th>X1</th>
<th>X2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-1.6785</td>
<td>0.0259</td>
</tr>
<tr>
<td>Variance</td>
<td>2.4583</td>
<td>0.0066</td>
</tr>
<tr>
<td>Numbers</td>
<td>61</td>
<td>61</td>
</tr>
<tr>
<td>Pearson Correlation Coefficient</td>
<td>-0.0569</td>
<td></td>
</tr>
<tr>
<td>Mean Difference</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Degree of Freedom</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>t value</td>
<td><strong>-8.4541</strong></td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) one-tailed</td>
<td>4.1327E-12</td>
<td></td>
</tr>
<tr>
<td>Critical Value: one-tailed</td>
<td>1.6706</td>
<td></td>
</tr>
<tr>
<td>P(T&lt;=t) two-tailed</td>
<td>8.2653E-12</td>
<td></td>
</tr>
<tr>
<td>Critical Value: two-tailed</td>
<td>2.0003</td>
<td></td>
</tr>
</tbody>
</table>
Table 8 Dynamic Nested Causality Relationship between Returns and Order Imbalances (in percentage)

<table>
<thead>
<tr>
<th></th>
<th>x₁ ∧ x₂</th>
<th>x₁ ←− x₂</th>
<th>x₁ ⇒ x₂</th>
<th>x₁ ← x₂</th>
<th>x₁ &lt;=&gt; x₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: All Trade Size</td>
<td>13.11</td>
<td>31.15</td>
<td>6.56</td>
<td>36.07</td>
<td>13.11</td>
</tr>
<tr>
<td>Panel B: Firm Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small Firm Size</td>
<td>15.00</td>
<td>30.00</td>
<td>0.00</td>
<td>40.00</td>
<td>15.00</td>
</tr>
<tr>
<td>Medium Firm Size</td>
<td>9.52</td>
<td>33.33</td>
<td>4.76</td>
<td>38.10</td>
<td>14.29</td>
</tr>
<tr>
<td>Large Firm Size</td>
<td>15.00</td>
<td>30.00</td>
<td>15.00</td>
<td>30.00</td>
<td>10.00</td>
</tr>
</tbody>
</table>

The causal relationships are defined as follows: ∧ is independency; ←− is contemporaneous relationship; ≠ is negation of a unidirectional relationship; <=> is feedback relationship; ≠> is negation of a strong unidirectional relationship where σ₁₂ = σ₂₁ = 0; and ◄<> is a strong feedback relationship where σ₁₂ = σ₂₁ = 0