# An Investigation of Short Sellers' Behavior around Short Interest Announcement Dates

Wei Wei

# An Investigation of Short Sellers' Behavior around Short Interest Announcement Dates


by

Wei Wei


A thesis submitted to the faculty of The University of Mississippi in partial fulfillment of the requirements of the Sally McDonnell Barksdale Honors College.

Oxford

April 2008

Approved by

Advisor: Dr. Bonnie Van Ness

Reader: Dr. John Conlon

Reader: Dr. Ken Cyree

# ABSTRACT

In this thesis, we examine the traders shorting behavior before and after the short interest settlement day. We investigate whether short sellers influence short interest by excessively opening new positions before short interest data collection. We find no evidence suggesting the existence of short sellers' short interest influencing activities. However, in the course of examining the differences among daily short levels, we demonstrate the danger and consequences of ignoring the panel structure of the data. Plus, we find results that are inconsistent with our understanding of the short interest collection process. Thus, we also give detailed discussions on panel data and on the short interest collection process.

ACKNOWLEDGEMENT

Many thanks to Dr. Bonnie Van Ness, Dr. Robert Van Ness, Dr. Andriy Shkilko, Dr. Xin Dang, and others who gave support during my preparation for this thesis. I especially want to thank Dr. John Conlon for teaching me all the essences in good academic research and good writing. I also thank Mr. Ben Blau for his comments and advices on my thesis and for providing me with the data for control variables used in this thesis.

# TABLE OF CONTENTS

# I. INTRODUCTION

Under the traditional definition of investing, an investor buys an asset, holds it while it appreciates in value, then sells the asset to make a profit. Short selling is the exact opposite. Short-sellers sell an asset, which they do not own, at the current price, and then buy it back at a future price to cover their short positions. Thus, short-sellers profit only when the price of the asset goes down. While short sales are allowed for most stocks traded in U.S. markets, the proceeds of short sales are generally held as collateral. Also, if the price rises rather than falls, short-sellers need to put down additional assets as collateral to maintain their short positions; otherwise an early recall may be demanded by the lender. In addition, during the time period of this study, NASDAQ Short Sale Regulations prohibit NASD members from short selling a NASDAQ-listed stock at or below the inside best bid when that bid is lower than the previous inside best bid[1] (NYSE follows a similar requirement but uses the previous trade price instead of bid price). Therefore, short selling is more costly and constrained than buying stocks (Zhou, 2007).

Many previous studies provide evidence that short-sellers are successful in identifying the securities that will underperform the market in the near future (see Asquith and Meulbroek (1995), Asquith et al. (2005), and Desai et al. (2002),). Others such as Senchack and Starks (1993) look at the effect of short interest announcements on

[1] As on July 5, 2007, SEC cancelled the up-tick rule on short sales. However, the up-tick rule was still in effect during our sample period from 2005 to 2006.

short-term stock returns. However, due to the lack of data on daily short sale transactions, previous studies only use the short interests of stocks published in the short interest announcement.

The short interest of a stock is the total number of shares that are shorted and still outstanding up to the day when such information is collected. By outstanding, we mean that the stocks are not yet repurchased by the short-sellers and returned to the lenders. The short interest data is collected on a monthly basis. In each month, the market makers are required to report the number of shares shorted on all recorded current short positions to the exchange on the "settlement day", which is usually on the 15th of each month or on the preceding trading day if the 15th is not a trading day. The exchange then calculates the short interest of each stock and sends the information to its wire service on the "dissemination day", which is usually the seventh trading day after the "settlement day". Finally, the short interest information is printed in newspapers and released to the public on the "publication day", which is the next trading day after the "dissemination day". According to NASDAQ, the monthly short interest information will be available on its Trader Web site for subscribed users after 4 p.m. on dissemination day. Table I-1 gives the detailed short interest announcement schedule for each month in 2005 and 2006. On the schedule, there is also a trade date which is always three days before the settlement date. Although NASDAQ does not start the importance of this date, we will show later that this date may actually mark the end of a short interest collection cycle[2]. Figure I-1 illustrates the timeline of how short interest data is recorded and published.

[2] We will discuss the ambiguity in the actual significance of the trade date and in the informational content of the short interest data in later sections.

In the beginning of 2005, after the implementation of SEC Regulation SHO, daily short sale transaction data became available to academic researchers. With this new source of information, researchers are now able to examine the level of short selling on a daily basis. Previously, Diether et al. (2007) and Boehmer et al. (2007) both use daily short transaction data in their research. However they all examine the relation between daily short selling levels and stock returns. In our research, we take a look at short-sellers' shorting behavior on a daily basis. We test whether short-sellers intentionally influence the short interest measure by excessively opening new short positions before the settlement day in order to drive the price down. More specifically, we want to see whether higher levels of short selling occur on trading days before the settlement day and lower levels of short selling occur on trading days after the settlement day. We examine whether high short level in one day is followed by high short level on other days. Finally, we investigate the relations between stocks' short interest and each day's short selling levels during the month.

We find no evidence indicating the existence of short interest influencing activity. In fact, we find that the average daily short selling level does not vary significantly from day to day across all months. Therefore, on any particular day, the position of that day relative to the settlement day in the month does not affect short-sellers' shorting behavior in general. Nevertheless, in the process of examining the differences among daily short levels, we demonstrate the danger and consequences of ignoring the panel structure of the data. Also, we obtain results that are inconsistent with our understanding of the short interest collection process.

The paper proceeds in eight sections. In the next section, we discuss the relevant literature on both the behavior of short-sellers and the predictive power of short interest. We then develop our hypotheses in Section III. We present our data and descriptive statistics in Section IV. We discuss our methods for testing the hypotheses and show the empirical results from our tests in Section V. We give our conclusions in Section VI. Then, in Section VII, we discuss our mistake from ignoring the panel structure in our data. Finally, in Section VIII, we review the findings that contradict our understanding of the short interest collection process and describe unsolved or not yet understood parts of our testing methods and empirical results.

## II. RELATED LITERATURE

There are several perspectives on the relation between short interest and future stock returns. According to Diamond and Verrecchia (1987), short selling is costly and short-sellers oftentimes face high risk. Thus, for a rational trader to be willing to take on such a risky position, he must believe that he is better informed about the potential downward movements in future security prices. Hence, a high short interest or short interest ratio[3] may reveal adverse information, implying a negative relation between short interest and stock returns. In practice, many investors consider high short interest a signal for negative future returns on the stock. Thus, the level of short interest affects investors' selling decisions.

However, there are other alternative views on the relation between short interest and stock returns. One is that a high level of short interest is considered a bullish signal because it represents latent demand, which will eventually transform into purchases to cover the short positions (Desai at el (2002)). Another view is that short selling may be unrelated to stock returns because some short selling is motivated by hedging strategies, arbitrage transactions, and tax-related reasons (Brent, Morse, and Stice (1990)) rather than by an anticipated drop in price. For example, traders may short sell securities in

[3] The short interest ratio is short interest divide by the total shares outstanding.

which they have a long position in order to pay less taxes with a short term gain. Further, a derivatives trader might short the underlying stock to replicate the payoff from an option. Finally, a market-maker might routinely short a stock as part of his inventory management (Purnanandam and Seyhun (2007)). The short positions described above do not reveal negative information regarding the stock.

Nevertheless, most academic studies find that short interest is negatively correlated with future stock returns. While some studies (e.g., Figlewski (1981), Brent, et al. (1990), Figlewski and Webb (1993), and Boehme et al. (2006)) find little relation between short interest and subsequent stock returns, several recent studies document that a higher short interest ratio predicts lower stock returns. For example, Asquith and Meulbroek (1995) find evidence that short-sellers, as a group, are able to identify securities that subsequently underperform the market. Boehmer et al. (2007) find that heavily shorted NYSE stocks underperform lightly shorted stocks by 1.16% on a risk-adjusted basis over the following 20 trading days on a value-weighted basis from 2000 to 2004. Diether et al. (2007) find that an increase in short activity by 10% of share volume is associated with a future decline in returns by 0.94 (0.72) percent per month on the NYSE (NASDAQ). Desai et al. (2002) find that heavily shorted firms experience negative abnormal returns ranging from -0.76% to -1.13% after controlling for market, size, book-to-market (B/M) and momentum factors. They also find that heavily shorted firms are more likely to be liquidated or delisted in the 36 month after being heavily shorted than their size-, book-to-market-, and momentum-matched control firms. Finally, Asquith et al. (2005) examine stocks with short interest ratios greater than 2.5%, 5%, and 10%, and stocks in the 95% and 99% percentiles based on short interest ratios and find that the abnormal monthly returns range from -0.28% to -1.25% on an equally weighted basis. They also find an increasing negative abnormal return among the highly shorted stocks when institutional holdings decrease. The monthly abnormal return for the portfolio with the highest demand and the least supply (99% percentile and lowest institutional ownership) is -2.15%.

## III. HYPOTHESES

The finance literature suggests that short-sellers as a group are more informed than the average trader and that high short interest corresponds with negative returns in the near future. In practice, traders tend to be influenced by the negative signals from short interest announcement. If short-sellers excessively open new short positions on some stocks and keep these positions open before the short interest settlement day, the total short volume of those short positions is included in the stocks' short interests for that month. Consequently, the short interest of those stocks will be higher than usual. When the short interest announcement comes out, other traders see those abnormally high short interests and start to short or sell the same stocks. As a result, the prices of those stocks will fall and the short-sellers can realize their gain by covering those short positions. Therefore, we believe that short-sellers have an incentive to influence stocks' short interests.

In order to influence the level of short interest of a stock, the short-sellers have to short before the short interest settlement day, otherwise the transaction will not be accounted for in the short interest of the current month. On the other hand, since longer holding periods impose more risk on the holders of short positions, short-sellers who short for influencing purposes would not want to short too early, either. Hence, we suspect that influencing short selling transactions would generally occur within a couple of days before the short interest settlement day in each month. For the convenience in our later discussion, we call these days the "influencing period". Consequently, there should be an abnormally high level of daily short selling in the influencing period.

However, an abnormally high daily level of shorting right before the settlement day alone is insufficient to prove that short-sellers intentionally influence short interests of some stocks. If short-sellers are truly trying to influence the short interests of some stocks, and expect the price on those stocks to fall after the announcement, they should also try to cash in on the gains by shorting more shares right before the short interest announcement day, the publication day. Thus, we expect to see stocks that have high daily short levels during influencing period to also have high daily short levels on the dissemination day or on the day before the dissemination day.

Finally, if our theory about short-sellers influencing activity is correct, then those stocks with relatively high daily short levels during the influencing period should also have relatively high short interests in that month for the following reasons. First, if we assume that on average the durations of short positions opened on each day are approximately the same, then naturally a bigger portion of the short positions opened on days closer to settlement day would remain open on the settlement day. Second, since we conjecture that the short-sellers excessively open new short positions and keep these positions open through settlement day, the short interest should pick up more of the positions opened closer to settlement day. Therefore, overall the short interests in a month should be more positively correlated with the daily short levels during the influencing period than with those of any other days (among all the days before settlement day).

In summary, we develop the following hypotheses to detect the existence of the short interest influencing activities on NASDAQ:

1. Daily short levels on average are significantly higher on days right before the short interest settlement day.

2. The daily short levels on days right before settlement day are most positively correlated with those on days right before publication day (among all the days between settlement day and publication day).

3. The short interests of a month are most positively correlated with the daily short levels on days right before settlement day.

We test these hypotheses in the rest of the paper.

# IV. DATA AND SAMPLE

To test our hypothesis, we need daily short transaction data and monthly short interest data. NASDAQ-listed stocks are traded on several different stock venues including the NASDAQ Stock Exchange, the American Stock Exchange (AMEX), the Chicago Stock Exchange (CHX), the National Stock Exchange (NSX), the Archipelago Electronic Stock Exchange (ARCAEX), and the Alternative Display Facility (ADF).[4] To obtain complete daily short volume, we use intraday short sale transaction data for all six market centers from the Trades and Quotes (TAQ) database.[5] We obtain daily returns, trade volume, high, low, and close prices, as well as market capitalization and shares outstanding from the Center of Research Securities Prices (CRSP) database. We also obtain monthly short interests on all stocks from NASDAQ.

Our sample time period is from January of 2005 to December of 2006 (503 trading days). We start with all NASDAQ-listed Class A common stocks with share codes 10 and 11 in CRSP. We assemble the intraday short transaction data into daily aggregates and merge the short sale data with the CRSP data. We exclude daily observations when the daily short volume exceeds the daily trade volume. We also omit observations with daily trade volume less than 100 shares. The filtered sample consists

[4] The Alternative Display Facility is run by NASD as a temporary alternative to the SuperMontage. The ADF reports its transactions separately from NASDAQ.

[5] The short sale transaction data are made available by each market in compliance with SEC Regulation SHO.

of 1,366,105 observations for 3,384 stocks, which includes many infrequently traded stocks on the NASDAQ. Less liquid stocks are subject to higher shorting constraints. In fact, many infrequently traded stocks cannot be shorted at all. Thus, we only use the filtered sample to obtain descriptive statistics for all NASDAQ stocks in general. For our empirical tests, we further trimmed our sample down to 742,428 observations for 1,476 stocks that are traded every day. For convenience in later discuss, we refer to those 1,476 stocks as active stocks hereafter.

Also included in the sample are 100 stocks that experienced temporary ticker changes during our sample period. Pursuant to Section 19(b) of the Securities Exchange Act of 1934, a stock will be traded under a temporary ticker when it fails to file its quarterly report, annual report, or other required documents. Also, a stock which goes through a reverse split may sometimes be traded under a temporary ticker for a short period, usually no more than 20 trading days. These stocks are generally excluded from most academic studies on NASDAQ stocks. However, we think that stocks with temporary ticker changes may be good targets for short-sellers. In fact, we find that many of these stocks have above average Short-to-Trade Ratios during our sample period and especially around the time of the ticker change. Hence, we manually match the temporary tickers with their regular tickers and keep those observations in our sample.

The descriptive statistics for all stocks with daily volume of 100 shares or more and for the active stocks are presented in Panels A and B in Table IV-1. The descriptive statistics for the short interest ratio are given in Table IV-2. We find that the average daily Short-to-Trade Ratio across all active stocks is about 0.30528 over the two years. Diether et al. (2007) also find in their sample that on average short volume is about 31.3

percent of NASDAQ trade volume in 2005. Further, Zhou (2007) calculates the monthly average short interest ratio in his data sample from 1995 to 2000. The minimum value of the average monthly short interest ratio in a year is about 1.02% and the maximum value of the average monthly short interest ratio is approximately 1.69%. Zhou shows that the average short interest ratio has been increasing steadily over the years. In our sample from 2005 and 2006, average monthly short interest ratio is about 5%. This confirms the notion that short interest ratio is in general increasing over time from.

# V. METHODS AND RESULTS:

## V.1 Dates and Measures:

In order to test our hypotheses, we first need to represent each trading day in terms of its relative position to the settlement day of the month. Thus, we label each day in each month relative to the short interest settlement date. For example, the settlement date is day S, the first day before the settlement date is day S-1, the first day after the settlement date is day S+1, and so on. We choose to keep an event window of 19 days from day S-9 to day S+9 so as not to double count days. Under this labeling method, the trade day will be day S-3, the dissemination day will be day S+7, and the publication day will be day S+8.

Second, we define our measures for daily short level. The easiest and most intuitive measure is the Short-to-Trade Ratio calculated as $\frac{ShortVolume}{TradeVolume}$. However, such a measure is affected by the differences in average daily short levels among stocks. Thus, we use two standardized measures in addition to Short-to-Trade Ratio.

The first measure, $DSTTR_y$ (daily short-to-trade ratio standardized by the current year), is standardized using the average Short-to-Trade Ratio for the current year; and is computed as:

$$DSTTR_{y\ i,t} = \frac{\left(\frac{ShortVolume}{TradeVolume}\right)_{i,t} - mean\left(\frac{ShortVolume}{TradeVolume}\right)_{\bullet\bullet}}{Std.dev.\left(\frac{ShortVolume}{TradeVolume}\right)_{\bullet\bullet}}$$

Thus, for each stock, we take its Short-to-Trade Ratio on day *i* in month *t* and subtract its average Short-to-Trade Ratio for the current year. Then we divide the difference by its own standard deviation of the Short-to-Trade Ratios for the year in which the observation occures. The standardized measure, $DSTTR_y$ should be similarly distributed with a zero mean and a unit variance (see Lakonishok and Vermaelen (1986), Koski and Scruggs (1998), and Sias (2004) for use of the similar standardized measures). Thus, t-statistics will test whether $DSTTR_y$ is significantly different from zero, which is the mean.

In summary, from day S-9 to day S+9, we now have 19 event days in our data. On each event day (from day S-9 to day S+9), we have 1,476 stocks; and, for each stock, we have 24 $DSTTR_y$' s (one from each month). So totally, we have 35,424 observations of $DSTTR_y$' s on event each.

However, in order for us to consider the yearly mean the normal short level, we have to assume that short-sellers are actually forward looking. Since the previous month is the most immediate reference for the short-sellers, it may be that to them, the short levels in the previous month are the normal short levels. Therefore, we create a second measure, $DSTTR_{pm}$ (daily short-to-trade ratio standardized by the previous month), using the average Short-to-Trade Ratio and the standard deviation for the previous month. This is calculated as:

$$DSTTR_{pm\ i,t} = \frac{\left(\frac{ShortVolume}{TradeVolume}\right)_{i,t} - mean\left(\frac{ShortVolume}{TradeVolume}\right)_{\bullet,t-1}}{Std.dev.\left(\frac{ShortVolume}{TradeVolume}\right)_{\bullet,t-1}}$$

For each stock, we take its Short-to-Trade Ratio on day *i* in month *t* and subtract its average Short-to-Trade Ratio in month *t-1*, then divide the difference by its standard deviation of the Short-to-Trade Ratios in month *t-1*. Since we do not have data from Dec, 2004, we cannot calculate $DSTTR_{pm}$*'s* in Jan 2005. Thus, we only have 23 $DSTTR_{pm}$*'s* for each stock on each day. Also, when using this measure, the $DSTTR_{pm}$*'s* will be slightly auto-correlated with each other.

In summary, from day S-9 to day S+9, we now have 19 days in our data. On each event day, we have 1,476 stocks; and, for each stock, we have 23 $DSTTR_{pm}$' s (one from each month excluding Jan, 2005). So in total, we have 33948 observations of $DSTTR_{pm}$' s on each event day.

## V.2 Results for average daily short levels

To test our first hypothesis, we first perform t-tests on the three daily short level measures. For daily Short-to-Trade Ratio, we test the differences between the average daily Short-to-Trade Ratio on day *i* and mean of daily Short-to-Trade Ratios on all other days. More specifically, each time, we set all Short-to-Trade Ratios on event day S-*i* as the treatment group and the Short-to-Trade Ratios on rest of the days as the control group. Then we perform a t-test with the null hypothesis that the mean of treatment group equals the mean of the control group. We perform the same procedures for all 19 event days. If our hypothesis is true, we expect to see a most positive and significant difference for days in the influencing period.

Since $DSTTR_y$ is a measure standardized by the yearly mean, if there is no special trading pattern from day to day, the average $DSTTR_y$'s on all labeled days should equal zero, which is the population mean over the year. Therefore, for each labeled day, we perform a t-test on $DSTTR_y$ across months and stocks with the null hypothesis that $\mu = 0$. We then perform t-tests in the same matter on $DSTTR_{pm}$. However, since the population mean for $DSTTR_{pm}$ is unknown, we decide to use the null hypothesis that $\mu$ equals the mean of all $DSTTR_{pm}$ in our sample. We acknowledge that we are testing $DSTTR_{pm}$ with a random variable rather than the true population mean. Thus, there is a risk of obtaining incorrect results due to sampling errors. If our hypothesis is true, we expect to see the most positive and significant mean on days in the influencing period for both $DSTTR_y$ and $DSTTR_{pm}$.

The results for the two-group t-tests on the Short-to-Trade Ratio are shown in Table V-1. The results for the daily means and t-statistics of $DSTTR_y$ and $DSTTR_{pm}$ are shown in Table V-2. To help visualize the patterns in these daily short level measures over time, we plot the daily means of the Short-to-Trade Ratio in Figure V-1 and the daily means of $DSTTR_y$ and $DSTTR_{pm}$ in Figure V-2. We see that day S-2 has the largest average level in all three measures. However, day S-5, day S+2, and day S+4 also have significantly high average levels in all three measures[6].

[6] As we will show later in our paper, those significant results are obtained from ignoring the panel structure of the data.

Next, we attempt to test whether day S-2 has the highest short level after controlling for other potential factors that may influence the daily level of short selling. First we run a multiple linear regression for the Short-to-Trade Ratio using the following model:

$$\begin{aligned}\text{Short-to-Trade} = {} & \beta_0 + \beta_1 Return_{i,t} + \beta_2 ln(Market\ Cap) + \beta_3 PriceVolatility_{i,t} + \beta_4 Pre9 \\ & + \beta_5\, Pre8 + \beta_6\, Pre7 + \beta_7\, Pre6 + \beta_6\, Pre5 + \beta_9\, Pre4 + \beta_{10}\, Pre3 \\ & + \beta_{11}\, Pre1 + \beta_{12}\, Pre0 + \beta_{13}\, Post1 + \beta_{14}\, Post2 + \beta_{15}\, Post3 + \beta_{16}\, Post4 \\ & + \beta_{17} Post5 + \beta_{18}\, Post6 + \beta_{19}\, Post7 + \beta_{20}\, Post8 + \beta_{21}\, Post9 + \varepsilon_{i,t}\end{aligned}$$

(1)

We regress the Short-to-Trade Ratio against the daily return, the natural log of the daily market Capitalization (in thousands), and the price volatility for stock *i* on day *t*, which is calculated as the difference between the daily high price and the daily low price divided by the daily high price. In order to compare the Short-to-Trade Ratio on day S-2 with the Short-to-Trade Ratio on other days, we create 18 dummy variables for the days from day S-9 to day S+9, for example, pre9 represents day S-9, pre0 represents day S, and post9 represents day S+9. The each dummy is equal to one for its corresponding day and zero for all other days. Since we want to see how the Short-to-Trade Ratios on other days compare with the Short-to-Trade Ratio on day S-2, we leave the dummy variable for day S-2, pre2, out of model.

Similarly, we regress our $DSTTR_y$ and $DSTTR_{pm}$ measures against the same independent variables as we do in model (1).

Since we choose day S-2 as a reference level in each regression, if our first hypothesis is true, then we expect to find negative coefficients with significant t-statistics for all 18 dummy variables. The results for the dummy variables of all three regressions are shown in Table V-3.

We see that while all the coefficients are negative, the t-statistics on day S-5 and day S+2 are not very significant. Therefore, we conclude that day S-2 has the highest daily short level, but its level is statistically no different from that of day S-5 and day S+2[7].

## V.3 Results for correlations among daily short levels

Since the short interest announcement is posted on the NASDAQ web site after trading hours on day S+7. If short-sellers want to further short on the stocks that they previously influenced on day S-2, they are most likely to do so within day S+6 and day S+7. Therefore, to test our second hypothesis, we first calculate the simple correlations of the daily short level measures between any two days in our event window. Since $DSTTR_y$ is a standardized measure, we simply calculate the correlations between all $DSTTR_y$*'s* on day *i* and all $DSTTR_y$*'s* on day *j*, where *i* and *j* are integers from -9 to 9. We also use the same method to calculate correlations for $DSTTR_{pm}$. However, the Short-to-Trade Ratio is not a standardized measure. Thus, to avoid picking up the correlations of average Short-to-Trade Ratios among stocks, we first calculate the correlation of the Short-to-Trade Ratios between two days on by each stock. Then we take the average of the correlations in all stocks. If our hypothesis is true, then we expect to see high correlations in daily short level between day S-2 and day S+6 or day S+7. We also

[7] Again, we will later show that those results are obtained from using incorrect method for regression.

expect to see that the correlations between those days are higher than the correlations between any other two days with the same time distance. Table V- 4 presents the correlations between daily short level on day S-2 and daily short level on each of the days in our event window. Since day S+6 is 8 days after day S-2 and day S+7 is 9 days after day S-2. We also show the correlations between daily short levels on days that are 8 days apart in Table V- 5 and the correlations between daily short levels on days that are 9 days apart in Table V- 6.

From Table V- 4, we see that for all three measures, the correlation between day S-2 and day S+6 are relatively small compared to those between day S-2 and each of the earlier days. This is also true for correlations between day S-2 and day S+7. In addition, from Table V- 5, we see that the correlation between day S-2 and day S+6 is not the highest among all the correlations between days that are 8 days apart. In fact, the correlation between day S-2 and day S+6 is among the lowest of all the correlations for $DSTTR_{pm}$. Similarly in Table V-6, the correlation between day S-2 and day S+7 is not the highest among all correlations between days with the same time distance. This suggests that the stocks heavily shorted on day S-2 are not heavily shorted on day S+6 or on day S+7. Therefore, our second hypothesis is not supported by the data.

## V.4 Results for correlation between daily short level and short interest

In order to test our third hypothesis, we need to define our measures for estimating the level of short interest. The most commonly used one is the short interest ratio defined as SI / SO, where SI is the short interest in shares and SO is the number of shares outstanding.

To remove the effect of the difference in the average short interest ratio among stocks, we use a measure that is similar to the standardized short interest measure that Purnanandam and Seyhun (2007) implement in their research. Our standardized measure is calculated as:

$$Standardized\ SI_t = \frac{\left(\frac{SI}{SO}\right)_t - mean\left(\frac{SI}{SO}\right)}{std.dev.\left(\frac{SI}{SO}\right)}$$

On a stock by stock basis, we subtract the average short interest ratio for the current year from the short interest ratio in month t. Then we divide the difference by the standard deviation of the short interest ratio for the current year. Similar to $DSTTR_y$, *Standardized SI* should be similarly distributed with a zero mean and a unit variance.

For each stock on each day, we match the 24 Short-to-Trade Ratios and $DSTTR_y$'s with their 24 short interests and *Standardized SI*'s by month. We also repeat the same matching procedure for $DSTTR_{pm}$. However, since we do not have $DSTTR_{pm}$'s in Jan, 2005. We delete the short interests and *Standardized SI*'s in Jan 2005 before we match them with their $DSTTR_{pm}$'s.

We calculate the correlations between the short interest and the daily short level on each of the 19 days. More specifically, for each event day, we calculate the correlation between daily Short-to-Trade Ratio and monthly short interest ratio. We also calculate the correlation between Short-to-Trade Ratio and *Standardized SI*. Then we repeat the same procedures for $DSTTR_y$ and $DSTTR_{pm}$. The results of these correlations are shown in Table IV-7. Since we find that the high daily short levels occur on day S-2, we expect the short selling activities for short interest influencing purposes to occur most likely on day S-2. Thus, we expect to see the largest correlation between short interest and the daily short level on day S-2, if our influence theory is correct.

Surprisingly, we see that no matter which set of correlations we choose to look at, the correlation for day S-2 is obvious not the highest among all the correlations. In fact, the correlation for day S-2 is lower than all correlations for days before day S-2 and only slightly higher than some of the correlations for days after day S-2. Our finding suggests that the high daily short level on day S-2 does not have a positive influence on short interest. Therefore, our third hypothesis does not hold.

## V.5 Revisiting results for average daily short level

So far, our results show that both hypothesis two and hypothesis three are false. Thus, the high daily short level on day S-2 do not seem to be a result of short-sellers' short interest influencing practices. Now, we start to wonder why day S-2 would be any different from other event days. When we take a look at Table V.1 again, we see that all the differences in the Short-to-Trade Ratio are small. Even for day S-2, the difference is only about 0.003, which is less than 1% of the average short-t-trade ratio on day S-2. We ask why the t-statistics is significant for such a small value of the difference.

We start to suspect that we may get significant t-statistics both from the t-tests and from the regressions because we ignore the panel structure[8] of our data. We believe that our data is subject to a time series random effect. Thus, to obtain the correct coefficients and t-statistics for our model (1), we re-test the model with a Time Series Cross Section Regression using a two-way random effects specification. The two-way random effects method controls for both cross-sectional and time-series random effects. We also re-test the model with the same regression but only controlling the cross-sectional effects. We repeat the same procedures for $DSTTR_y$ and $DSTTR_{pm}$. Table V-8 shows the regression results for the Short-to-Trade Ratio. Table V-9 shows the regression results for $DSTTR_y$. Table V-10 shows the regression results for $DSTTR_{pm}$.

We see that for both the Short-to-Trade Ratio and $DSTTR_y$, the coefficients of all the dummy variables do not change much as we move from the ordinary least squares method, to the one-way random effects method, and to the two-way random effects method. However, the t-statistics from both the ordinary least squares method and the one-way random effects method are about ten times larger than those from the two-way random effects method. Hence, we see that the significant t-statistics in Table V-1, Table V-2, and Table V-3 are indeed incorrect results obtained from improper use of ordinary least squares on panel data.

Thus, as shown in Table V-8 and Table V-9, day S-2 is not significantly different from any of the other days in the month. Therefore, our first hypothesis is actually not

---

[8] We will discuss panel data issues and problems with ignoring the panel structure in more detail in section VII.

supported. In fact, in unreported estimates, we find that none of the event days is significantly different from the others when using the two-way random effects method.

## VI. CONCLUSION

Based on our results, all three hypotheses turn out to be false. Thus, we conclude that short-sellers do not seem to influence short interest by excessively opening short positions right before the settlement day. In fact, there is no significant difference in daily short selling levels on any of the days in a month. Nevertheless, we have demonstrated the consequences of ignoring panel structure in cross-sectional time-series data. As we will show in the next section, we also learn to understand the rationale behind our mistake, which results in obtaining incorrect significant t-statistics.

# VII. PANEL DATA AND TWO-WAY EFFECTS

Panel data, also called cross-sectional time series data, are data where information about multiple cases is collected at different time periods. Thus, the information contained in panel data goes in two-dimensions. In one dimension, the cross-sectional information tells us the differences between all cases at a fix point of time. In the other dimension, the time-series information tells us how each case itself changes over time. As we will see in later discussion, this two-dimensional structure complicates the way we deal with panel data and make it very easy to obtain incorrect results.

Under an ordinary linear model, the independent variable $Y_i$ is expresses as:

$$Y_i = \alpha + \sum_{j=1}^{K} \beta_j X_{ji} + \varepsilon_i ,$$

where $\alpha$ is the intercept, each $X_j$ is one independent variable, and $\varepsilon_i$ is the error term. Our goal is to estimate the coefficients, $\beta_j$ 's.

In a panel data, the error term becomes $u_{it} = \zeta_i + \omega_t + \varepsilon_{it}$. Thus, we have:

$$Y_{it} = \alpha + \sum_{j=1}^{K} \beta_j X_{jit} + \zeta_i + \omega_t + \varepsilon_{it}$$

$\zeta_i$ is an error term that is different for each case in the cross-section but constant through all time period for one specific case. $\omega_t$ is an error term that is the same for all cases at a specific point of time but varies from one time period to another. Finally, $\varepsilon_{it}$ is the error term that is different for each case at any given time. Thus, $\zeta_i$ constitutes the purely cross-sectional effect and $\omega_t$ constitutes the purely time-series effect.

Let us recall that we have 24 months of data. Each month has the same 19 event days (day S-9 to day S-9). Plus, on each event day, we have 1,476 stocks. So, in our case, $\zeta_i$ accounts for the effect of any news that affects only one individual stock and affects that stock the same way throughout all 24 months. $\omega_t$ accounts for the effect of any news on the macro-level that affects all 1,476 stocks the same way in one day; but on some day all stocks are affect in one way while on other days all stocks are affect in other ways. For example, sometimes traders are all optimistic about the stock market thus short sells less; other times traders are all pessimistic about the stock market thus short sells more. As we can see, for each event day in each month, we have one $\omega_t$. Thus, overall for each event day, we have 24 $\omega_t$ 's.

It may happen that, for one event day, more than half of the 24 $\omega_t$ 's are positive. For example, maybe in 17 out of the 24 months, short traders are all pessimistic about the market on day S-2. Thus, in those 17 months, they short a lot on day S-2. Then those usually high short levels on all stocks on day S-2 should be explained by the positive

$\omega_t$ 's on day S-2. Now if we ignore the panel structure and use the ordinary least squares method, we are assuming that $\varepsilon_{it}$ is the only error term present in our data when, in reality, $\zeta_i$ and $\omega_t$ are also present. Since we omit $\omega_t$ , the time-series effects previously explained by $\omega_t$ , now have to be explained either by $\varepsilon_{it}$ 's or by the coefficient, $\beta_j$. Thus, we are forcing OLS to think that either those short levels are usually high because all the $\varepsilon_{it}$ 's coincidentally move together on day S-2's in all 17 months or something is special about the coefficient of $x_{jit}$ . Since on each day S-2, we have 1,476 observations for $\varepsilon_{it}$ , it is extremely unlikely that on all those 17 days the $\varepsilon_{it}$ 's for all stocks move together. Therefore, OLS naturally chooses to believe that there is something special about $\beta_j$, the coefficient of $x_{jit}$ .

In essence, since $\omega_t$ is the same for all stocks on a particular day. It is repeated 1476 times on each day. So, we are forcing OLS to think that the variance from 24 observations is the variance from 35,424 observations. Thus, we trick OLS to believe that we have more observations than we really have. Consequently, the t-statistics we obtained from OLS regression is magnified. Therefore, we get all these significant results when they should not be significant at all.

Further, since $\zeta_i$ represents the effect on each individual stock on all days, it is the same from day to day for a specific stock. Thus, it should not make any difference when we are comparing daily short level from day to day. The goal of our regression is to check on the differences between the average daily short level on day S-2 and that on each of the other days; and the dummy variables are daily dummies. Therefore, the omission of $\zeta_i$ should not have much effects on our results.

Indeed, as we can see in Table V-8 and Table V-9, the results from the ordinary least squares method and from the one-way random effects method (controlling for cross-sectional effects) are similar. This means that the cross-sectional random effects in our data do not have much effect on our estimations of coefficients for our dummy variables. Thus, the time-series random effect is the main reason for the breakdown of the ordinary least squares method.

# VIII. REMAINING PUZZLES

## VIII.1 Ambiguity about the Trade Date

In a recent paper, Zhou (2007) suggests that the short interest reported on the settlement day is based on the short positions that are open on and before the trade day. Thus, he uses the trade date as the basic event day of his study. Henry (2006) also suggests the same idea about trade date in his research.

However, NASDAQ claims that the short interest data is collected on the settlement day and that the dealers report all the short positions as of the settlement day. Also, in the reporting instructions posted for its member firms, FINRA[9] states: "Trade Date is provided for reference purposes only. Positions are to be reported as of settlement date (Filing Due Dates, 2008)." Plus, Purnanandam and Siphon (2007) state that the NYSE also collects data on short interest from all member firms as of the 15th day of every month or the next business day if it is a holiday. Hence, we believe that the outstanding short positions opened after trade day are also included in the short interest calculation as long as those short transactions are recorded on the books before the dealers report to NASDAQ. Therefore, we still use the settlement day as the basic event day in our study.

---

[9] FINRA is the trademark of Financial Industry Regulatory Authority, Inc.

As we see from Table V-7, the correlations on day S-1 and day S-2 seem to be too small for days right before the collection of short interest data. So, either almost of the short positions opened on day S-2 are closed before the settlement day or maybe the outstanding short positions opened on day S-2 are not counted in the short interest measure. We also see that the correlations on day S-3 and on days before S-3 are relatively bigger than those on the other days. Thus, our finding seems to support Zhou and Henry's view that the trade date is the real mark off point for the short interest reporting period. Therefore, further investigation is required to clarify the exact function of the trade date and exact period of coverage of the short interest announcement.

## VIII. 2 Problems with $DSTTR_{pm}$

If we take a look at the last column in Table V-7, we notice that, for most of the days, the correlation between monthly *Standardized SI* and $DSTTR_{pm}$ (daily short-to-trade ratio standardized by the previous month) is negative. Also, in the second last column in Table V-7, we see that, for all the days, the correlation between monthly *Short Interest Ratio* and $DSTTR_{pm}$ is negative. Such results are very counterintuitive and fairly different from the results we obtained when using Short-to-Trade Ratio and $DSTTR_y$ (daily short-to-trade ratio standardized by the current year) as the measure for daily short level. Since $DSTTR_{pm}$ is standardized with the mean and standard deviation of the previous month, we do expect the results for $DSTTR_{pm}$ to be somewhat different from the results for the Short-to-Trade Ratio and $DSTTR_y$. However, at this point, we do not know exactly why we would obtain such odd results for $DSTTR_{pm}$ in this particular way, namely they are almost all negative.

In Table V-10, we also notice that for $DSTTR_{pm}$, all three regression methods produce similarly significant results. However, we expect the t-statistics from the two-way random effects method to be insignificant. As discussed in the previous paragraph, since $DSTTR_{pm}$ is standardized with the mean and standard deviation of previous month, it may behave differently from the Short-to-Trade Ratio and $DSTTR_{y}$. However, in this particular case, we do not know why such a method of standardization would result in significant t-statistics even under the two-way random effects method.

# References


Arnold, T., A.W. Butler, T.F. Crack, and Y. Zhang, 2005, The Information Content of Short Interest: A Natural Experiment, Journal of Business 78, 1307-1335.

Asquith P., Pathak P. A., and J. R. Ritter, 2005, Short interest, institutional ownership, and stock returns, Journal of Financial Economics, 78, 243-276.

Asquith, Paul, and Lisa Meulbroek, 1995. An empirical investigation of short interest, unpublished M.I.T. working paper.

Boehme, Rodney D., Bartley R. Danielsen, and Sorin M. Sorescu, 2006. Short sale constraints, differences of opinion, and overvaluation, Journal of Financial and Quantitative Analysis, Jun2006, Vol. 41 Issue 2, p455-487

Boehmer, E., C. Jones, and X. Zhang, 2007. Which Shorts are Informed, forthcoming Journal of Finance.

Brent, A., D. Morse, and E.K. Stice, 1990, Short Interest: Explanations and Tests. Journal of Financial and Quantitative Analysis 25, 273-289.

Chen, Honghui, and Vijay Singal, 2003, Role of speculative short sales in price formation: the case of weekend effect, *Journal of Finance* 58, 685-705.

Covrig, Vicentiu and Ng, Lilian, 2004, Volume autocorrelation, information, and investor trading, *Journal of Banking & Finance*, 2004, vol. 28, issue 9, pages 2155-2174.

Desai, Hemang, K. Ramesh, S.R. Thiagarajan, and B. V. Balachandran, 2002. An investigation of the informational role of short interest in the Nasdaq market, *Journal of* Finance 57, 2263-2287.

Diamond, D.W., Verrecchia, R.E., 1987. Constraints on short-selling and asset price adjustment to private information. *Journal of Financial Economics* 18, 277–311.

Diether, K.B., Lee, Kuan-Hui and Werner, Ingrid M., 2007, Can Short-Sellers Predict Returns? Daily Evidence. Forthcoming, *Review of Financial Studies*.

Fama, Eugene F. James and MacBeth, D., 1973, Risk, Return, and Equilibrium: Empirical Tests. *The Journal of Political Economy*, Vol. 81, No. 3 (May - Jun., 1973), pp. 607-636.

Figlewski, Stephen, 1981, The informational effects of restrictions on short sales; Some empirical evidence, *Journal of Financial and Quantitative Analysis* 16, 463-476.

Figlewski, S. and G.P. Webb (1993): Options, Short Sales, and Market Completeness. *Journal of Finance* 48, 761-777.

*Filing Due Dates* 2008, Regulatory Systems Financial Industry Regulatory Authority, Inc., <http://www.finra.org/RegulatorySystems/RegulationFilingApplications/p005712>.

Koski, J.L., and J.T. Scruggs, 1998, Who Trades Around the Ex-Dividend Day? Evidence from NYSE Audit File Data. *Financial Management* 27, 3, 58-72.

Lakonishok, J., and T. Vermaelen, 1986, Tax-Induced Trading Around Ex-dividend Days. *Journal of Financial Economics* 3, 287-319.

Senchack, A.J., Jr., and Laura T. Starks, 1993, Short-sale restrictions and market reaction to short-interest announcements. *Journal of Financial and Quantitative Analysis* 28, 177-194.

Purnanandam, Amiyatosh and Seyhun, Nejat, 2007, Shorts and Insiders. working paper.

Sias, R.W., 2004, Institutional Herding. *Review of Financial Studies* 17, 165-206.

Zhou Xin, 2007, Understanding the information content of short interest. Job Market Paper.

## Table I-1: Timeline for short data collection for NASDAQ in 2005, 2006

The "trade date" is for reference only. The "settlement date" is the date when the market makers report the number of shares they sold short and are still outstanding. The "dissemination date" is the date information is sent to wire services. The "publication date" is the date when short positions are printed in newspapers.

| | Trade Date | Settlement Date | Dissemination Date | Publication Date |
|---|---|---|---|---|
| Panel A. Schedules for 2005 | | | | |
| January | 01/11/05 | 01/14/05 | 01/26/05 | 01/27/05 |
| February | 02/10/05 | 02/15/05 | 02/25/05 | 02/28/05 |
| March | 03/10/05 | 03/15/05 | 03/24/05 | 03/28/05 |
| April | 04/12/05 | 04/15/05 | 04/26/05 | 04/27/05 |
| May | 05/10/05 | 05/13/05 | 05/24/05 | 05/25/05 |
| June | 06/10/05 | 06/15/05 | 06/24/05 | 06/27/05 |
| July | 07/12/05 | 07/15/05 | 07/26/05 | 07/27/05 |
| August | 08/10/05 | 08/15/05 | 08/24/05 | 08/25/05 |
| September | 09/12/05 | 09/15/05 | 09/26/05 | 09/27/05 |
| October | 10/11/05 | 10/14/05 | 10/25/05 | 10/26/05 |
| November | 11/09/05 | 11/15/05 | 11/25/05 | 11/28/05 |
| December | 12/12/05 | 12/15/05 | 12/27/05 | 12/28/05 |
| Panel B. Schedules for 2006 | | | | |
| January | 01/10/06 | 01/13/06 | 01/25/06 | 01/26/06 |
| February | 02/10/06 | 02/15/06 | 02/27/06 | 02/28/06 |
| March | 03/10/06 | 03/15/06 | 03/24/06 | 03/27/06 |
| April | 04/10/06 | 04/13/06 | 04/25/06 | 04/26/06 |
| May | 05/10/06 | 05/15/06 | 05/24/06 | 05/25/06 |
| June | 06/12/06 | 06/15/06 | 06/26/06 | 06/27/06 |
| July | 07/11/06 | 07/14/06 | 07/25/06 | 07/26/06 |
| August | 08/10/06 | 08/15/06 | 08/24/06 | 08/25/06 |
| September | 09/12/06 | 09/15/06 | 09/26/06 | 09/27/06 |
| October | 10/10/06 | 10/13/06 | 10/24/06 | 10/25/06 |
| November | 11/09/06 | 11/15/06 | 11/27/06 | 11/28/06 |
| December | 12/12/06 | 12/15/06 | 12/27/06 | 12/28/06 |

## Table IV-1: Descriptive Statistics

The table report descriptive statistics for all NASDAQ-listed common stocks. Price is the daily closing price. Price volatility is calculated by taking the difference between the daily high price and the daily low price and divides the difference by the daily high price. Short to trade, short for Short to Trade Ratio, is calculated as daily short volume over daily trade volume.

| | Return | Short Volume | Trade Volume | Market Cap (in thousands) | Shares Outstanding | Price Volatility | number of trade | short to trade |
|---|---|---|---|---|---|---|---|---|
| Panel A. All NASDAQ Stocks (excluding trades with volume less than 100) | | | | | | | | |
| Mean | 0.000321 | 176,899 | 555,212 | 1,007,803 | 49936.91 | 0.03781 | 1,408.87 | 0.24690 |
| St. Dev | 0.00305 | 923,496 | 2,789,091 | 6,756,541 | 275585.18 | 0.01969 | 4,714.37 | 0.12136 |
| Min | -0.07165 | 0 | 335.88 | 2,266.23 | 735.000 | 0.000659 | 2.71 | 0 |
| Max | 0.04720 | 21,565,569 | 68,153,221 | 271,599,349 | 10420817 | 0.17820 | 106,817 | 0.52854 |
| N | 3384 | 3384 | 3384 | 3384 | 3384 | 3384 | 3384 | 3384 |
| Panel B. Active NASDAQ Stocks | | | | | | | | |
| Mean | 0.000446 | 329,163 | 982,279 | 1,813,887 | 83852.15 | 0.03836 | 2,497.22 | 0.30528 |
| St. Dev | 0.00117 | 1,364,351 | 4,111,917 | 10,052,339 | 410128.39 | 0.01417 | 6,770.01 | 0.09770 |
| Min | -0.005822 | 583.94 | 7,431.71 | 10,595.96 | 1711.00 | 0.012218 | 45.25 | 0.01385 |
| Max | 0.005799 | 21,565,569 | 68,153,221 | 271,599,349 | 10420817 | 0.10053 | 106,817 | 0.48095 |
| N | 1476 | 1476 | 1476 | 1476 | 1476 | 1476 | 1476 | 1476 |

## Table IV-2: Descriptive Statistics for Short Interest Ratio

The following statistics are for the active stocks, stocks that are traded every day. We have 35424 observations (24 month with 1476 stocks in each month). In Panel A, we give the basic statistics for the Short Interest Ratio. In Panel B, we show the quantiles for the Short Interest Ratio.

| Panel A. Basic Measures of the Short Interest Ratio | | | |
|---|---|---|---|
| Mean | 0.053435 | Std Deviation | 0.056232 |
| Median | 0.036825 | Variance | 0.003162 |
| Mode | 0 | Range | 0.582462 |

| Panel B. Short Interest Ratio Quantiles | |
|---|---|
| Quantile | Estimate |
| 100% Max | 0.582462377 |
| 99% | 0.277716429 |
| 95% | 0.162003648 |
| 90% | 0.120353439 |
| 75% Q3 | 0.069812235 |
| 50% Median | 0.036825144 |
| 25% Q1 | 0.017438838 |
| 10% | 0.005203846 |
| 5% | 0.00198401 |
| 1% | 0.000269419 |
| 0% Min | 0 |

Table V-1: Average Daily Short-to-Trade Ratio and T-Statistics

The results presented in this table come from a t-test between the treatment group and control group. The treatment group contains all Short-to-Trade Ratios on day S-*i*. The control group contains all the Short-to-Trade Ratios on the remaining days. The first two columns show the mean and standard deviation of the daily Short-to-Trade Ratios on each day. The last two columns show the difference between the means of two groups and its t-statistics for each day.

| Event | Short-to-Trade Ratio | | | |
|---|---|---|---|---|
| Date | Mean | Std.Dv | Difference | t-statistics |
| S-9 | 0.3058 | 0.1649 | 0.00046 | (0.52) |
| S-8 | 0.3036 | 0.1639 | -0.00190** | (-2.12) |
| S-7 | 0.3039 | 0.1642 | -0.00160* | (-1.79) |
| S-6 | 0.3060 | 0.1651 | 0.00069 | (0.76) |
| S-5 | 0.3080 | 0.1664 | 0.00300*** | (3.10) |
| S-4 | 0.3041 | 0.1664 | -0.00140 | (-1.54) |
| S-3 | 0.3066 | 0.1664 | 0.00100 | (1.40) |
| S-2 | 0.3086 | 0.0661 | 0.00340*** | (3.82) |
| S-1 | 0.3067 | 0.1647 | 0.00100 | (1.60) |
| S | 0.3066 | 0.1655 | 0.00100 | (1.41) |
| S+1 | 0.3067 | 0.1658 | 0.00100 | (1.51) |
| S+2 | 0.3083 | 0.1662 | 0.00310*** | (3.43) |
| S+3 | 0.3059 | 0.1647 | 0.00055 | (0.61) |
| S+4 | 0.3069 | 0.1656 | 0.00200* | (1.83) |
| S+5 | 0.3055 | 0.1660 | 0.00015 | (0.17) |
| S+6 | 0.3008 | 0.1670 | -0.00480*** | (-5.31) |
| S+7 | 0.3022 | 0.1668 | -0.00330*** | (-3.64) |
| S+8 | 0.3017 | 0.1655 | -0.00390*** | (-4.30) |
| S+9 | 0.3041 | 0.1652 | -0.00130 | (-1.43) |

## Table V-2: Average Daily *DSTTRy* and *DSTTRpm* with T-Statistics

The daily mean and t-statistics for *DSTTRy* and *DSTTRpm* are all presented in the columns of this table. The daily mean and t-statistics for *DSTTRy* and *DSTTRpm* come from simple t-tests. For t-test on *DSTTRy*, we use H0: $\mu = 0$. However, since *DSTTRpm* has a mean of 0.04858, we use H0: $\mu = 0.04858$ instead of 0 when conducting the t-test on *DSTTRpm*.

| Event | *DSTTRy* | | *DSTTRpm* | |
|---|---|---|---|---|
| Date | Mean | T- Statistics | Mean | T- Statistics |
| S-9 | 0.0027 | (0.5053) | 0.0542 | (0.7686) |
| S-8 | -0.0132*** | (-2.5038) | 0.0267*** | (-2.5777) |
| S-7 | -0.0104** | (-1.9698) | 0.0402 | (-0.4813) |
| S-6 | 0.0061 | (1.1489) | 0.0451 | (-0.4822) |
| S-5 | 0.0220*** | (4.1329) | 0.0721*** | (3.0386) |
| S-4 | -0.0081* | (-1.5150) | 0.0374 | (-1.1848) |
| S-3 | 0.0079* | (1.4836) | 0.0551 | (0.8246) |
| S-2 | 0.0246*** | (4.6866) | 0.0913** | (2.1999) |
| S-1 | 0.0119** | (2.2473) | 0.0633* | (1.5940) |
| S | 0.0102** | (1.9210) | 0.0518 | (0.3957) |
| S+1 | 0.0059 | (1.1009) | 0.0535 | (0.5383) |
| S+2 | 0.0222*** | (4.1999) | 0.0673*** | (2.5374) |
| S+3 | 0.0071* | (1.3418) | 0.0624** | (1.7473) |
| S+4 | 0.0157*** | (2.9850) | 0.0612* | (1.5621) |
| S+5 | 0.0012 | (0.2342) | 0.0523 | (0.4291) |
| S+6 | -0.0366*** | (-6.8584) | -0.0035*** | (-6.5569) |
| S+7 | -0.0265*** | (-4.9084) | 0.0101*** | (-4.4899) |
| S+8 | -0.0283*** | (-5.3109) | 0.0284 | (-1.2562) |
| S+9 | -0.0036 | (-0.6875) | 0.0542 | (0.6550) |

## Table V-3: Regression Results for Each Individual Event Date

The first two columns show the coefficients and t-statistics for dummy variables in model (1) with Short-to-Trade Ratio as our dependent variable. The second two columns show the coefficients and t-statistics for dummy variables in model (1) with *DSTTRy* as our dependent variable. The last two columns show the coefficients and t-statistics for dummy variables in model (1) with *DSTTRpm* as our dependent variable.

| Event | *Short-to-Trade Ratio* | | *DSTTRy* | | *DSTTRpm* | |
|---|---|---|---|---|---|---|
| Dates | coefficients | T-statistics | coefficients | T-statistics | coefficients | T-statistics |
| S-9 | -0.0048*** | (-4.2258) | -0.0353*** | (-4.7254) | -0.0578*** | (-3.9548) |
| S-8 | -0.0059*** | (-5.1798) | -0.0432*** | (-5.7789) | -0.0746*** | (-5.1088) |
| S-7 | -0.0060*** | (-5.3379) | -0.0439*** | (-5.8729) | -0.0655*** | (-4.4876) |
| S-6 | -0.0039*** | (-3.4459) | -0.0242*** | (-3.2393) | -0.0548*** | (-3.7508) |
| S-5 | -0.0019* | (-1.6855) | -0.0078 | (-1.0472) | -0.0282* | (-1.9332) |
| S-4 | -0.0057*** | (-5.0059) | -0.0377*** | (-5.0431) | -0.0612*** | (-4.1936) |
| S-3 | -0.0029** | (-2.5450) | -0.0236*** | (-3.1531) | -0.0479*** | (-3.2837) |
| S-2 | 0.0000 | | 0.0000 | | 0.0000 | |
| S-1 | -0.0028** | (-2.4594) | -0.0195*** | (-2.6136) | -0.0392*** | (-2.6828) |
| S | -0.0038*** | (-3.3848) | -0.0281*** | (-3.7612) | -0.0583*** | (-3.9892) |
| S+1 | -0.0034*** | (-2.9761) | -0.0274*** | (-3.6704) | -0.0493*** | (-3.3761) |
| S+2 | -0.0012 | (-1.0491) | -0.0092 | (-1.2262) | -0.0357** | (-2.4462) |
| S+3 | -0.0045*** | (-3.9907) | -0.0319*** | (-4.2710) | -0.0511*** | (-3.4956) |
| S+4 | -0.0023** | (-1.9894) | -0.0128* | (-1.7083) | -0.0367** | (-2.5131) |
| S+5 | -0.0040*** | (-3.5182) | -0.0267*** | (-3.5663) | -0.0468*** | (-3.2067) |
| S+6 | -0.0101*** | (-8.9512) | -0.0736*** | (-9.8432) | -0.1134*** | (-7.7616) |
| S+7 | -0.0082*** | (-7.2031) | -0.0609*** | (-8.1429) | -0.0939*** | (-6.4324) |
| S+8 | -0.0080*** | (-7.0530) | -0.0578*** | (-7.7302) | -0.0710*** | (-4.8630) |
| S+9 | -0.0057*** | (-5.0695) | -0.0372*** | (-4.9799) | -0.0510*** | (-3.4938) |

## Table V- 4: Correlations between day -2 and each of the other days

This table shows the correlation between daily short level on day -2 and daily short level on each of the days in our event window. The first column presents the correlation results when daily short level is measure by Short-to-Trade Ratio. The second column presents the correlation results when daily short level is measure by *DSTTRy*. The third column presents the correlation results when daily short level is measure by *DSTTRpm*.

| | Short-to-Trade<br>S-2 | *DSTTRy*<br>S-2 | *DSTTRpm*<br>S-2 |
|---|---|---|---|
| Correlation with s-9 | 0.2218 | 0.1685 | 0.0683 |
| Correlation with s-8 | 0.2338 | 0.1805 | 0.0825 |
| Correlation with s-7 | 0.2560 | 0.2071 | 0.0662 |
| Correlation with s-6 | 0.2737 | 0.2268 | 0.1063 |
| Correlation with s-5 | 0.2957 | 0.2464 | 0.1201 |
| Correlation with s-4 | 0.3303 | 0.2874 | 0.1245 |
| Correlation with s-3 | 0.4051 | 0.3680 | 0.1668 |
| Correlation with s-2 | 1.0000 | 1.0000 | 1.0000 |
| Correlation with s-1 | 0.4036 | 0.3716 | 0.5152 |
| Correlation with s | 0.3414 | 0.3012 | 0.4789 |
| Correlation with s+1 | 0.2848 | 0.2447 | 0.6289 |
| Correlation with s+2 | 0.2805 | 0.2332 | 0.1168 |
| Correlation with s+3 | 0.2606 | 0.2119 | 0.1096 |
| Correlation with s+4 | 0.2401 | 0.1920 | 0.4353 |
| Correlation with s+5 | 0.2216 | 0.1723 | 0.2677 |
| Correlation with s+6 | 0.2039 | 0.1514 | 0.1027 |
| Correlation with s+7 | 0.1975 | 0.1460 | 0.1535 |
| Correlation with s+8 | 0.1953 | 0.1418 | 0.8529 |
| Correlation with s+9 | 0.1883 | 0.1361 | 0.4039 |

## Table V- 5: Correlations between days that are 8 days apart

This table shows the correlation between daily short levels on days that are 8 day apart. The first column presents the correlation results when daily short level is measure by Short-to-Trade Ratio. The second column presents the correlation results when daily short level is measure by *DSTTRy*. The third column presents the correlation results when daily short level is measure by *DSTTRpm*.

| | Short-to-Trade | *DSTTRy* | *DSTTRpm* |
|---|---|---|---|
| Corr (dayS-9, dayS-1) | 0.1890 | 0.1395 | 0.1281 |
| Corr (dayS-8, day S) | 0.1973 | 0.1453 | 0.1440 |
| Corr (dayS-7, dayS+1) | 0.2092 | 0.1639 | 0.0695 |
| Corr (dayS-6, dayS+2) | 0.2331 | 0.1831 | 0.2468 |
| Corr (dayS-5, dayS+3) | 0.2187 | 0.1680 | 0.2203 |
| Corr (dayS-4, dayS+4) | 0.2155 | 0.1695 | 0.2191 |
| Corr (dayS-3, dayS+5) | 0.2169 | 0.1711 | 0.2200 |
| Corr (dayS-2, dayS+6) | 0.2039 | 0.1514 | 0.1027 |
| Corr (dayS-1, dayS+7) | 0.2027 | 0.1558 | 0.2101 |
| Corr (day S, dayS+8) | 0.2086 | 0.1574 | 0.4441 |
| Corr (dayS+1, dayS+9) | 0.2044 | 0.1601 | 0.3766 |

## Table V- 6: Correlations between days that are 9 days apart

This table shows the correlation between daily short levels on days that are 9 day apart. The first column presents the correlation results when daily short level is measure by Short-to-Trade Ratio. The second column presents the correlation results when daily short level is measure by *DSTTRy*. The third column presents the correlation results when daily short level is measure by *DSTTRpm*.

| | Short-to-Trade | *DSTTRy* | *DSTTRpm* |
|---|---|---|---|
| Corr (dayS-9, day S) | 0.2002 | 0.1458 | 0.1620 |
| Corr (dayS-8, dayS+1) | 0.1920 | 0.1463 | 0.1310 |
| Corr (dayS-7, dayS+2) | 0.2163 | 0.1669 | 0.0881 |
| Corr (dayS-6, dayS+3) | 0.2130 | 0.1671 | 0.2268 |
| Corr (dayS-5, dayS+4) | 0.1992 | 0.1456 | 0.2039 |
| Corr (dayS-4, dayS+5) | 0.2116 | 0.1634 | 0.1693 |
| Corr (dayS-3, dayS+6) | 0.1945 | 0.1460 | 0.2249 |
| Corr (dayS-2, dayS+7) | 0.1975 | 0.1460 | 0.1535 |
| Corr (dayS-1, dayS+8) | 0.1908 | 0.1429 | 0.4382 |
| Corr (day S, day+S9) | 0.2113 | 0.1615 | 0.3368 |

## Table V- 7: Correlations between daily short level and short interest

The first two columns show the correlation between Short-to-Trade Ratio and Short Interest Ratio and the correlation between Short-to-Trade Ratio and *Standardized SI* for each day. The second two columns show the correlation between *DSTTRy* and Short Interest Ratio and the correlation between *DSTTRy* and *Standardized SI* for each day. The last two columns show the correlation between *DSTTRpm* and Short Interest Ratio and the correlation between *DSTTRpm* and *Standardized SI* for each day.

| Event | Short-to-Trade | | *DSTTRy* | | *DSTTRpm* | |
|---|---|---|---|---|---|---|
| Date | SI/SO | StSi/SO | SI/SO | StSi/SO | SI/SO | StSi/SO |
| S-9 | 0.2377 | 0.0922 | 0.0334 | 0.1162 | -0.0106 | -0.0006 |
| S-8 | 0.2417 | 0.0881 | 0.0365 | 0.1149 | -0.0090 | 0.0000 |
| S-7 | 0.2408 | 0.0971 | 0.0341 | 0.1246 | -0.0074 | 0.0012 |
| S-6 | 0.2327 | 0.0980 | 0.0268 | 0.1274 | -0.0167 | 0.0167 |
| S-5 | 0.2229 | 0.0929 | 0.0168 | 0.1200 | -0.0268 | 0.0065 |
| S-4 | 0.2291 | 0.0950 | 0.0232 | 0.1245 | -0.0181 | 0.0068 |
| S-3 | 0.2286 | 0.0869 | 0.0244 | 0.1121 | -0.0186 | -0.0026 |
| S-2 | 0.2214 | 0.0582 | 0.0174 | 0.0753 | -0.0143 | -0.0198 |
| S-1 | 0.2280 | 0.0580 | 0.0182 | 0.0706 | -0.0260 | -0.0349 |
| S | 0.2100 | 0.0308 | 0.0013 | 0.0407 | -0.0399 | -0.0573 |
| S+1 | 0.2281 | 0.0392 | 0.0230 | 0.0510 | -0.0200 | -0.0476 |
| S+2 | 0.2190 | 0.0333 | 0.0135 | 0.0433 | -0.0277 | -0.0635 |
| S+3 | 0.2252 | 0.0450 | 0.0135 | 0.0544 | -0.0337 | -0.0554 |
| S+4 | 0.2209 | 0.0315 | 0.0102 | 0.0385 | -0.0302 | -0.0596 |
| S+5 | 0.2136 | 0.0340 | 0.0034 | 0.0404 | -0.0348 | -0.0566 |
| S+6 | 0.2197 | 0.0373 | 0.0107 | 0.0456 | -0.0301 | -0.0528 |
| S+7 | 0.2110 | 0.0328 | 0.0018 | 0.0365 | -0.0370 | -0.0531 |
| S+8 | 0.2194 | 0.0254 | 0.0087 | 0.0332 | -0.0197 | -0.0363 |
| S+9 | 0.2242 | 0.0264 | 0.0140 | 0.0326 | -0.0292 | -0.0597 |

SI/SO stands for Short Interest Ratio, StSi/SO stands for Standardized Short Interest Ratio, *Standardized SI*.

Table V- 8: Regression Results with Short-to-Trade Ratio

All results presented in this table are from the regression with Short-to-Trade Ratio as our independent variable. The first two columns show the coefficients and t-statistics of each dummy variable from ordinary least square method. The second two columns show the coefficients and t-statistics of each dummy variable from one-way random effects method (controlling only for cross-sectional random effects). The third two columns show the coefficients and t-statistics of each dummy variable from two-way random effects method (controlling for both cross-sectional and time-series random effects).

| Event | OLS Method | | 1-Way Random | | 2-Way Random | |
|---|---|---|---|---|---|---|
| Dates | coefficients | t-statistics | coefficients | t-statistics | coefficients | t-statistics |
| S-9 | -0.0048*** | (-4.2258) | -0.0046*** | (-4.6891) | -0.0049 | (-0.4380) |
| S-8 | -0.0059*** | (-5.1798) | -0.0059*** | (-5.9763) | -0.0060 | (-0.5367) |
| S-7 | -0.0060*** | (-5.3379) | -0.0060*** | (-6.0811) | -0.0062 | (-0.5521) |
| S-6 | -0.0039*** | (-3.4459) | -0.0034*** | (-3.4714) | -0.0037 | (-0.3280) |
| S-5 | -0.0019* | (-1.6855) | -0.0014 | (-1.3750) | -0.0016 | (-0.1437) |
| S-4 | -0.0057*** | (-5.0059) | -0.0053*** | (-5.3095) | -0.0055 | (-0.4880) |
| S-3 | -0.0029** | (-2.5450) | -0.0030*** | (-2.9916) | -0.0030 | (-0.2722) |
| S-2 | 0.0000 | | | | | |
| S-1 | -0.0028** | (-2.4594) | -0.0026*** | (-2.6586) | -0.0028 | (-0.2473) |
| S | -0.0038*** | (-3.3848) | -0.0036*** | (-3.6799) | -0.0039 | (-0.3478) |
| S+1 | -0.0034*** | (-2.9761) | -0.0030*** | (-3.0109) | -0.0032 | (-0.2876) |
| S+2 | -0.0012 | (-1.0491) | -0.0011 | (-1.1293) | -0.0012 | (-0.1102) |
| S+3 | -0.0045*** | (-3.9907) | -0.0045*** | (-4.5201) | -0.0047 | (-0.4199) |
| S+4 | -0.0023** | (-1.9894) | -0.0022** | (-2.1888) | -0.0022 | (-0.2010) |
| S+5 | -0.0040*** | (-3.5182) | -0.0034*** | (-3.4523) | -0.0036 | (-0.3239) |
| S+6 | -0.0101*** | (-8.9512) | -0.0092*** | (-9.2963) | -0.0096 | (-0.8628) |
| S+7 | -0.0082*** | (-7.2031) | -0.0075*** | (-7.6042) | -0.0079 | (-0.7021) |
| S+8 | -0.0080*** | (-7.0530) | -0.0075*** | (-7.5603) | -0.0077 | (-0.6886) |
| S+9 | -0.0057*** | (-5.0695) | -0.0055*** | (-5.6007) | -0.0057 | (-0.5124) |

Table V- 9: Regression Results with *DSTTRy*

All results presented in this table are from the regression with *DSTTRy* as our independent variable. The first two columns show the coefficients and t-statistics of each dummy variable from ordinary least square method. The second two columns show the coefficients and t-statistics of each dummy variable from one-way random effects method (controlling only for cross-sectional random effects). The third two columns show the coefficients and t-statistics of each dummy variable from two-way random effects method (controlling for both cross-sectional and time-series random effects).

| Event | OLS Method | | 1-Way Random | | 2-Way Random | |
|---|---|---|---|---|---|---|
| Dates | coefficients | T-statistics | coefficients | T-statistics | coefficients | T-statistics |
| S-9 | -0.0353*** | (-4.7254) | -0.0353*** | (-4.7262) | -0.0367 | (-0.4539) |
| S-8 | -0.0432*** | (-5.7789) | -0.0432*** | (-5.7817) | -0.0435 | (-0.5382) |
| S-7 | -0.0439*** | (-5.8729) | -0.0439*** | (-5.8746) | -0.0447 | (-0.5526) |
| S-6 | -0.0242*** | (-3.2393) | -0.0242*** | (-3.2362) | -0.0253 | (-0.3130) |
| S-5 | -0.0078 | (-1.0472) | -0.0078 | (-1.0429) | -0.0090 | (-0.1115) |
| S-4 | -0.0377*** | (-5.0431) | -0.0377*** | (-5.0400) | -0.0387 | (-0.4786) |
| S-3 | -0.0236*** | (-3.1531) | -0.0236*** | (-3.1548) | -0.0241 | (-0.2977) |
| S-2 | 0.0000 | | | | | |
| S-1 | -0.0195*** | (-2.6136) | -0.0195*** | (-2.6115) | -0.0205 | (-0.2530) |
| S | -0.0281*** | (-3.7612) | -0.0281*** | (-3.7590) | -0.0297 | (-0.3679) |
| S+1 | -0.0274*** | (-3.6704) | -0.0274*** | (-3.6665) | -0.0288 | (-0.3566) |
| S+2 | -0.0092 | (-1.2262) | -0.0092 | (-1.2255) | -0.0099 | (-0.1227) |
| S+3 | -0.0319*** | (-4.2710) | -0.0319*** | (-4.2708) | -0.0334 | (-0.4130) |
| S+4 | -0.0128* | (-1.7083) | -0.0128* | (-1.7076) | -0.0133 | (-0.1640) |
| S+5 | -0.0267*** | (-3.5663) | -0.0266*** | (-3.5601) | -0.0278 | (-0.3439) |
| S+6 | -0.0736*** | (-9.8432) | -0.0735*** | (-9.8334) | -0.0762 | (-0.9420) |
| S+7 | -0.0609*** | (-8.1429) | -0.0608*** | (-8.1368) | -0.0627 | (-0.7755) |
| S+8 | -0.0578*** | (-7.7302) | -0.0577*** | (-7.7254) | -0.0590 | (-0.7292) |
| S+9 | -0.0372*** | (-4.9799) | -0.0372*** | (-4.9781) | -0.0384 | (-0.4746) |

## Table V- 10: Regression Results with *DSTTRpm*

All results presented in this table are from the regression with *DSTTRpm* as our independent variable. The first two columns show the coefficients and t-statistics of each dummy variable from ordinary least square method. The second two columns show the coefficients and t-statistics of each dummy variable from one-way random effects method (controlling only for cross-sectional random effects). The third two columns show the coefficients and t-statistics of each dummy variable from two-way random effects method (controlling for both cross-sectional and time-series random effects).

| Event | OLS Method | | 1-Way Random | | 2-Way Random | |
|---|---|---|---|---|---|---|
| Dates | coefficients | T-statistics | coefficients | T-statistics | coefficients | T-statistics |
| S-9 | -0.0578*** | (-3.9548) | -0.0578*** | (-3.9703) | -0.0580*** | (-3.9863) |
| S-8 | -0.0746*** | (-5.1088) | -0.0746*** | (-5.1245) | -0.0748*** | (-5.1415) |
| S-7 | -0.0655*** | (-4.4876) | -0.0656*** | (-4.5010) | -0.0657*** | (-4.5169) |
| S-6 | -0.0548*** | (-3.7508) | -0.0549*** | (-3.7707) | -0.0551*** | (-3.7893) |
| S-5 | -0.0282* | (-1.9332) | -0.0284* | (-1.9529) | -0.0287** | (-1.9704) |
| S-4 | -0.0612*** | (-4.1936) | -0.0614*** | (-4.2141) | -0.0615*** | (-4.2295) |
| S-3 | -0.0479*** | (-3.2837) | -0.0479*** | (-3.2897) | -0.0480*** | (-3.2965) |
| S-2 | 0.0000 | | | | | |
| S-1 | -0.0392*** | (-2.6828) | -0.0392*** | (-2.6929) | -0.0392*** | (-2.6934) |
| S | -0.0583*** | (-3.9892) | -0.0583*** | (-4.0015) | -0.0583*** | (-4.0042) |
| S+1 | -0.0493*** | (-3.3761) | -0.0494*** | (-3.3940) | -0.0495*** | (-3.4010) |
| S+2 | -0.0357** | (-2.4462) | -0.0357** | (-2.4538) | -0.0357** | (-2.4556) |
| S+3 | -0.0511*** | (-3.4956) | -0.0511*** | (-3.5050) | -0.0511*** | (-3.5100) |
| S+4 | -0.0367** | (-2.5131) | -0.0367** | (-2.5201) | -0.0367** | (-2.5245) |
| S+5 | -0.0468*** | (-3.2067) | -0.0470*** | (-3.2307) | -0.0471*** | (-3.2402) |
| S+6 | -0.1134*** | (-7.7616) | -0.1137*** | (-7.8046) | -0.1139*** | (-7.8239) |
| S+7 | -0.0939*** | (-6.4324) | -0.0941*** | (-6.4634) | -0.0943*** | (-6.4798) |
| S+8 | -0.0710*** | (-4.8630) | -0.0712*** | (-4.8862) | -0.0713*** | (-4.8984) |
| S+9 | -0.0510*** | (-3.4938) | -0.0511*** | (-3.5067) | -0.0511*** | (-3.5136) |

Figure I-1: Timelines for short interest announcement days

The "settlement date" is the date when market makers report the number of shares they sold short and are still outstanding. The "dissemination date" is the date information is sent to wire services. The "publication date" is the date when short positions are printed in newspapers. The short interest data collected on settlement day in month t contains information on all the outstanding short positions opened since the settlement days in month t-1.

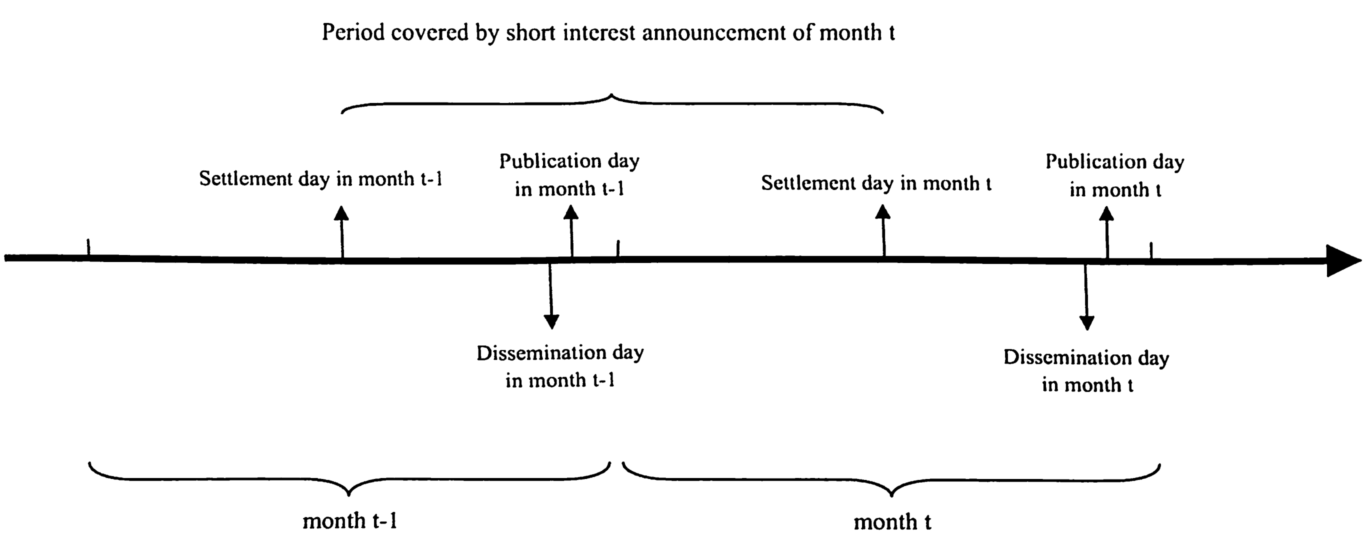

Figure V-1: Average Daily Short-to-Trade Ratio Plot

This figure shows the average Short-to-Trade Ratio across stocks and months on each labeled day. Day 0 is the settlement day, day 7 is the dissemination day, and day 8 is the publication day.

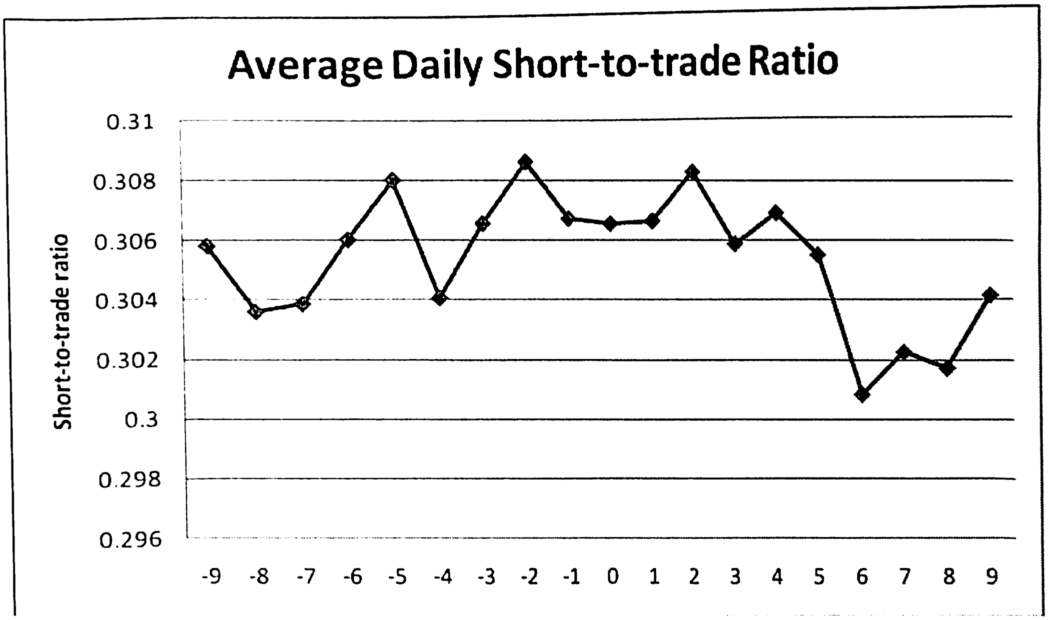

Figure V-2: Average Daily Standardized Measures Plot

This figure shows the average *DSTTRy* measure and *DSTTRpm* measure across stocks and months on each labeled day. Day 0 is the settlement day, day 7 is the dissemination day, and day 8 is the publication day.