# Internet chat, disagreement, retail trading, and stock returns around earnings announcements

September 2003

Henk Berkman
School of Business and Economics
University of Auckland
Auckland, New Zealand
h.berkman@auckland.ac.nz

Paul D. Koch
School of Business
University of Kansas
Lawrence, KS 66045
pkoch@ku.edu

Preliminary. Please do not quote without permission. We thank Chris Anderson, George Bittlingmayer, Mike Bradbury, John Charnes, Dan Dahliwahl, Mike Ettredge, Catherine Shenoy, and seminar participants at Australian National University, Massey University, the University of Kansas, and the University of Sydney for their helpful comments. Koch acknowledges support from the University of Auckland, where he served as visitor while this research was conducted.

# Internet chat, disagreement, retail trading, and stock returns around earnings announcements


## Abstract

This study tests the Miller (1977) hypothesis as an explanation for stock price behavior around technology firms' earnings announcements during the late 1990s. Specifically, we examine whether the anomalous tendency for stock prices to increase (decrease) before (after) earnings announcements during this period is associated with an increase (decrease) in investor disagreement before (after) the announcement. For a sample of high-tech stocks in 1998, we use the daily number of messages posted on a firm's Internet message board (chat room) as a measure of the level of investor disagreement about the firm's prospects. Consistent with Miller, we find that stocks with a larger increase in the level of disagreement before the announcement tend to have a larger pre-announcement price increase, and a larger price reversal after the earnings announcement. We also find that both small and large investors buy before the announcement and sell afterward, although large investors begin selling sooner than retail investors. In addition, we find our disagreement measure is directly related to net initiated order flow from retail investors, but not institutional investors. Finally, we find that the relative amount of retail versus institutional trading in a stock is positively related to the pre-announcement price runup and negatively related to the post-announcement reversal, and that disagreement is significantly associated with these return patterns only for the subsample of stocks with the highest proportion of retail trading. These results are consistent with the view that retail investors are less willing or less able to (short) sell than institutional investors, and they suggest the Miller hypothesis applies more to retail investors than to institutional investors in this setting.



JEL Classification: D82, G14, G19.
Key Words: market efficiency, short sale restrictions, disagreement, earnings announcement, internet stocks, retail investors.

## 1. Introduction

Trueman et al., (2003) report a curious pattern in stock returns around quarterly earnings announcements for a sample of Internet firms during the period, January 1998 to August 2000. In the few days preceding the earnings announcements there is an average price increase of about 5%, which is followed by a similar price reversal over the following several days. This return pattern does not seem to be related to the accounting information in the earnings announcements, or to changes in risk associated with the announcement. The price pattern is similar in up and down markets, and shows up in each quarter during the period.

Trueman et al., (2003) also document an increase in buyer-initiated trades prior to the announcement and an increase in seller-initiated trades following the announcement, but they offer no explanation for this pattern of behavior. In his discussion of Trueman et al., Berger (2003, p. 274 ) observes, 'The paper thus presents an intriguing apparent anomaly, but is unable to make much progress in explaining it.' Berger, along with Trueman et al., conjectures the anomalous price and order flow patterns are related to trading by short-term retail investors, who were responsible for a relatively large proportion of trading volume in internet stocks during this period (see also Berkman et al., 2003, Ofek and Richardson, 2003).

This paper proposes an explanation for these seemingly anomalous price and order flow patterns around earnings announcements, and tests this explanation for a sample of technology stocks during 1998. Our explanation focuses on the Miller (1977) hypothesis, which considers the role of disagreement among heterogeneous investors in a market where there are short sale restrictions. According to this hypothesis, disagreement among investors results in inflated prices as optimistic investors buy the stock, whereas pessimistic investors are limited in their ability or willingness to short the stock.[1]

1 See, for example, Chen et al., (2002), Diether et al., (2001), Duffie et al., (2002), Jones and Lamont (2002), Lintner (1969), Miller (1977), and Ofek and Richardson (2003).

Recent empirical evidence indicates the Miller (1977) hypothesis offers an explanation for temporarily high stock prices or price reversals. For example, Diether et al., (2001) use the dispersion across analysts' earnings forecasts as a proxy for investor disagreement, and find stocks with higher disagreement have lower future returns over the following months. This outcome is consistent with the Miller hypothesis, suggesting that greater disagreement results in more inflated prices and lower subsequent returns.[2]

We argue that an approaching quarterly earnings announcement is an information event that is likely to trigger increased uncertainty and disagreement among investors. Consistent with the Miller (1977) hypothesis, the intensified disagreement results in upward price pressure before the earnings announcement. After the announcement, when the new information reduces uncertainty about earnings and disagreement dissipates, the initial price increase is naturally followed by a reversal.[3] If this explanation truly characterizes the behavior documented in Trueman et al., (2003), the increase in net initiated order flow and prices before the earnings announcement are both expected to be larger for stocks with a larger increase in the level of disagreement before the announcement. For these stocks we also expect a larger reversal in net initiated order flow and prices after the announcement.

Our explanation for the anomalous stock price patterns around earnings announcements focuses on changes in the level of disagreement before and after the announcement. Empirical analysis of this explanation requires a proxy for changes in the level of disagreement on the days before and after earnings announcements. We use the daily number of messages posted on a firm's internet message board (chat room) as our measure of investor disagreement in the market.

---

2 Empirical support for the Miller (1977) hypothesis is also provided in Chen, et al., (2003), Diether et al., (2001), Ofek and Richardson (2003), and Antweiler and Frank (2003b).

3 Patell and Wolfson (1979, 1981) and Donders and Vorst (1996) find that implied volatilities increase significantly in the days before earnings announcements and drop sharply after the earnings announcement to move back to their normal level. Patell and Wolfson (1981) interpret the pre-announcement implied volatility as a reflection of investors' beliefs about the range of possible stock price reactions to a forthcoming disclosure.

The amount of chat activity offers several potential advantages over analyst forecasts to measure the level of investor disagreement. First, chat data are available in a timely fashion and chat activity around earnings announcements is relatively high. Analysts, on the other hand, revise their forecasts infrequently, so that a measure of disagreement based on recent analyst forecasts displays little variation in the few days before and after earnings announcements. Second, chat data are available for many smaller firms that are not followed extensively by analysts. Finally, chat activity may be particularly useful as a proxy for disagreement among retail investors, given that small, less sophisticated traders are more likely to read and post internet messages than institutional investors.[4]

The paper makes several contributions. First, we offer an explanation for a previously documented stock market anomaly. Second, in the process we document the influence of disagreement on stock prices over a considerably shorter time frame than in previous studies. Third, we show that the level of disagreement (measured by the amount of chat activity) affects the net initiated order flow of small and large investors differently. Finally, we provide evidence that the anomalous stock price behavior around earnings announcements is significantly related to the relative amount of retail versus institutional trading in a stock.

We obtain data on the daily amount of chat activity during 1998 for a sample of 206 companies from www.siliconinvestor.com, a major message board specializing in technology stocks. For this sample of stocks, we also measure the daily buying and selling activity initiated by small and large investors, respectively, from trade and quote (TAQ) data. We find both abnormal chat activity and total trading volume increase dramatically on the days before the earnings announcement, peak on the day after the announcement, and then decline back to pre-announcement levels. We also find patterns in stock prices and trade flows over the

4 Theoretical papers that address the economic motivation for posting or reading messages on Internet chat rooms include Admati and Pfleiderer (2001), Cao, Coval and Hirschleifer (2002) and Demarzo, Vayanos and Zwiebel (2001). Empirical evidence in Antweiler and Frank (2003a,b), Berkman et al., (2003) and Das et al., (2001) supports the proposition that the number of messages posted on a firm's Internet message board (chat room) is a useful measure of the level of investor disagreement in the marketplace.

twenty-day period around earnings announcements, similar to those documented in Trueman et al., (2003). In addition we document that, while both small and large investors tend to be net buyers (sellers) during the period before (after) the earnings announcement, net buying by small traders is significantly greater than that by large traders on the day before and the day after the announcement.

Given this evidence, we conduct three sets of tests to examine whether the amount of abnormal chat activity and the relative amount of retail versus institutional trading can help to explain these patterns in stock prices and order flows. Our first set of tests relates the stock returns to the change in the level of disagreement before and after the announcement. We find the magnitude of the price runup before earnings announcements is directly related to the magnitude of changes in abnormal chat activity before the announcement. Furthermore, consistent with Miller (1977), stocks with a higher level of abnormal chat activity on the day before the announcement tend to have a larger price decrease after the announcement.

Our second set of tests analyzes the correlation between the amount of abnormal chat activity and both total trading volume and net initiated order flow by retail and institutional investors. First, we find trading volume by both retail and institutional investors is highly correlated with chat activity for all 20 days around earnings announcements. We interpret this result as an indication that the amount of chat activity is a useful proxy for disagreement among both retail and institutional investors.[5] Next, our analysis of the correlation between chat activity and net initiated order flow shows a remarkable difference between retail and institutional investors. We find a significant positive correlation between chat activity and net initiated order flow by retail investors for most days around earnings announcements. This evidence support the Miller (1977) hypothesis, and suggests that optimistic retail investors are more willing or able to initiate trades than pessimistic retail traders. However, for

5 This result is consistent with Antweiler and Frank (2003a,b), Berkman et al., (2003) and Das et al., (2001).

institutional investors, the correlation between chat activity and net initiated order flow tends to be negative or insignificant. This outcome suggests that increased chat activity prior to the announcement is not associated with asymmetric net buying by institutional investors.

These divergent results for the net buying of small versus large investors call for an explanation. Miller's (1977) theory is based on an asymmetry in the ability or willingness of investors to buy versus sell stocks. According to this theory, an increase in disagreement results in an upward price bias because of increased buying pressure from optimistic traders, whereas pessimistic traders are limited in their ability to (short) sell the stock. We argue that this asymmetry in the ability or willingness to buy and sell is likely to differ between retail and institutional traders. Retail investors typically hold only a few stocks and are thus limited in their ability to sell stocks directly from their portfolio. Moreover, the typical retail investor is averse to short selling. The risk profile and higher proportional costs of short selling make this trading strategy relatively unattractive, especially when applied in small amounts. On the other hand, institutional investors hold many stocks, and short selling is part of the daily routine for many institutional investors such as hedge funds.[6]

Our third set of tests further explores the influence of retail trading activity on this phenomenon. First, we explore the association between the magnitude of stock price changes around earnings announcements and the relative trading activity of retail versus institutional investors. We find the price runup before the earnings announcement and the price reversal after the announcement are larger for stocks with a higher proportion of retail trading relative to institutional trading. This result corroborates the work of Ofek and Richardson (2002), who report that abnormal returns following information events were particularly large for internet stocks with a high proportion of retail trading during the internet boom. Second, we

6 Recent behavioral finance work on prospect theory suggests smaller, unsophisticated investors are less willing to sell stocks they own than are sophisticated, institutional investors, when prospects for the stock look bad. Barber and Odean (2002) report that the median household maintains fewer than 3 stock positions, and that only 0.29 percent of these are short positions.

also find the relation between returns and changes in disagreement before and after earnings announcements is stronger for stocks with a relatively high amount of retail trading.

This study proceeds as follows. Section 2 reviews the relevant literature. Section 3 describes the data and our measures of stock returns, chat activity, and net order flow. Section 4 examines the patterns in stock returns and net order flow by small and large investors, over the twenty-day period around earnings announcements. Section 5 analyzes the relations among stock returns, net order flow by small and large investors, and abnormal chat activity before and after earnings announcements. The final section summarizes and concludes.

## 2. Literature Review

This paper draws on several research areas. We first review the literature regarding the impact of investor disagreement on stock prices when investors are limited in their ability to short the stock. Next, we review the growing body of research that uses chat-based data to study price formation in stock markets. Finally, we discuss empirical studies that distinguish between the order flow from retail and institutional investors to investigate differences in their investment behavior and performance.

### *2.1 Heterogenous beliefs and short-sales constraints*

Lintner (1969) and Miller (1977) introduced the idea that, when there are short-sales constraints, a stock price will be biased upward because the most optimistic investors will buy the stock whereas pessimists are kept out of the market due to the short-sales restrictions. In Miller's model, a larger increase in disagreement among investors results in a more inflated price, and is followed by a larger price decrease when the disagreement dissipates.[7] In contrast, Diamond and Verrecchia (1987) argue that short-sales constraints are recognized

7 More recent theoretical models such as Morris (1996), Chen et al., (2002) and Viswanathan (2000) arrive at similar predictions.

in the marketplace and rational traders will adjust their prices in such a way that, on average, prices are not biased.

Several recent papers provide empirical evidence that supports the basic Miller (1977) prediction, indicating that arbitrageurs do not always force prices back to equilibrium in the presence of short-sales constraints. For example, two recent papers find evidence indicating that the loosening of short sale constraints is followed by price declines. First, Ofek and Richardson (2003) show that, prior to the Internet crash in March 2000, market-adjusted prices for Internet stocks fell an average of 35 percent in the six months following the expiration of lockup provisions associated with IPO's. They attribute this dramatic price drop to the loosening of the short sale constraint after lock-up expiration.[8] Second, Chen et al., (2003) use breadth of ownership as a proxy for short sale constraints, and find empirical support for their conjecture that future returns decline when there is less breadth of ownership. As noted by the authors, their results are surprising in that there seems to be no explanation for the absence of enough short selling activity to drive prices back to their equilibrium level.[9] Two other papers focus on changes in the level of disagreement as the impetus for temporarily high prices or price reversals. First, Diether et al., (2001) use the dispersion across analysts' earnings forecasts as a proxy for investor disagreement, and find stocks with higher disagreement have lower future returns over the following months. This outcome is consistent with the Miller hypothesis, suggesting that greater disagreement results in more inflated prices and lower subsequent returns. Second, Antweiler and Frank (2003b) find that stocks that are the subject of a greater level of investor disagreement, as proxied by a higher volume of Internet chat message posting, have lower returns in the following month.

8 In IPO's for Internet stocks during the 1990's, only 15% to 20% of the shares were typically sold to the public. The lock-up period is the period of time after the IPO during which other shareholders cannot sell their shares.
9 Chen et.al. (2002) point out that even though more than 70% of mutual funds are prohibited from short selling, others could fill the void. Geczy et.al. (2002) report that the cost of short selling can be substantial, but find little support for the view that short selling restrictions made it difficult to implement a selling strategy based on the conviction that internet stocks would go down during the internet boom.

*2.2 Chat room activity and investor disagreement*

A growing literature uses stock message board data to shed light on issues related to price formation in stock markets. For example, Antweiler and Frank (2003a), Das et al., (2001), and Tumarkin and Whitelaw (2001), consider the verbal content of chat messages. They use statistical language processing techniques to measure market sentiment and the extent of disagreement prevailing in the market, and they relate these measures to stock returns, trading volume and stock price volatility.

Antweiler and Frank (2003a) and Das et al., (2001) find a strong positive correlation between the amount of posting activity and their measure of disagreement (based on the difference between the number of messages classified as bullish and bearish). Das et al., (2001) explain their finding with a rhetorical question: 'If everyone agreed, why post?'[10]

Antweiler and Frank (2003b) and Berkman et al., (2003) provide further evidence on the usefulness of chat activity as a measure of investor disagreement. Antweiler and Frank (2003b) find that stocks with a high volume of message posting have low returns in the following month. Berkman et al. test the proposition in Kandel and Pearson (1995) that differential interpretation of public information is the main reason to trade on days with minimal price changes. For a subsample of days with minimal price changes, they find a strong correlation between trading volume and the amount of abnormal chat activity, which suggests that chat activity offers a useful proxy for disagreement. Berkman et al. also report a significantly higher correlation of chat activity with retail trading volume than with institutional trading volume, indicating that chat activity might be particularly useful as a measure of disagreement among retail traders.

*2.3 Retail versus institutional net order flow*

10 Theoretical papers that address the economic motivation for posting or reading chat messages include Admati and Pfleiderer (2001), Cao, Coval and Hirschleifer (2002) and Demarzo, Vayanos and Zwiebel (2001).

Both Trueman et al., (2003) and Berger (2003) suggest the anomalous price pattern around earnings announcements for internet stocks during 1998-2000 might be related to the trading of unsophisticated, retail investors. This suggestion seems reasonable in light of the evidence that retail traders make irrational trading decisions (see, for example, Barber and Odean, 1999 and 2002) that might not always be arbitraged away (Shleifer and Vishny, 1997). Consistent with this view, Ofek and Richardson (2003) report that the first-day return for internet IPO's, and the returns around the expiration of the quiet period, are significantly larger if there is a relatively high amount of retail trading. They also document that a large proportion of trading in internet stocks during the internet boom originated from retail investors.

We use two procedures to distinguish between the trading activity of retail versus institutional investors, to analyze whether they display differential behavior in relation to the anomalous stock price pattern reported in Trueman et al., (2003). First, following Lee (1992) we use a single cutoff of $20,000 to partition all trades into small (retail) and large (institutional) trade categories. Our second procedure excludes a buffer zone of medium-sized trades, by splitting all trades into (retail) trades with value less than $10,000 versus (institutional) trades with value larger than $50,000. Lee and Radhakrishna (2000) show that excluding the buffer zone of medium-sized trades significantly reduces misclassifications.[11]

## 3. Data, Research Design and Descriptive Statistics

### *3.1 Data*

The chat data are obtained from http://www.siliconinvestor.com, one of the four major investor message boards (the others are Yahoo, the Motley Fool and Raging Bull). This web site focuses on technology companies and, unlike other chat room sites, charged users a fee

11 Using this cutoff, Lee and Radhakrishna (2000) find the probability that an individual trade is erroneously classified as an institutional trade is 2 percent, and the probability that an institutional trade is erroneously included in small trades is 19 percent. This approach follows a substantive body of prior research that suggests disaggregating volume by trade size effectively sorts investors according to investor sophistication and informedness (Atiase and Bamber 1994, Bhattacharya 2001, Chan and Lakonishok 1993, Cready 1988, Easley and O'Hara 1987, Hasbrouck 1988, 1991, Lee 1992, Lee and Radhakrishna 2000, and Utama and Cready 1997).

for *posting* chat messages during this sample period (there was no charge for *reading* messages on this site). For 331 technology companies, we obtained the daily number of messages posted on their message boards during 1998. While this sample includes several large firms aggressively followed by financial analysts and chat room activity, it is mostly comprised of smaller technology companies.[12] Our measure of chat activity for firm i on day t ($Chat_{it}$) is the number of messages posted over the 24-hour period from the previous day's market close to the current day's market close (4:00p.m. to 4:00p.m. E.S.T.). We exclude firms that average less than two chat messages per day.[13]

For each firm, for every quarter in 1998, we also collect the date and time of the earnings announcement from the *Dow Jones Newswire*. We are interested in the behavior of stock prices, net order flow, and chat activity on the days before and after each quarterly announcement. Following Trueman et al., (2003), we exclude 108 announcements that were made during regular trading hours, or for which no exact time could be obtained. These screens leave 691 quarterly earnings announcements by 206 firms in our final sample.

Analysts' earnings forecasts and actual quarterly earnings per share are obtained from the I/B/E/S database. We calculate the earnings surprise each quarter as the difference between the actual quarterly earnings announced and the most recent analyst forecast prior to the announcement, scaled by the share price 10 days before the announcement.

Measures of daily stock returns, trading volume, and net initiated order flow are generated from the Trade and Quote (TAQ) database. These data include the number of shares traded and the transaction prices for all trades during 1998. For each sample company we compute stock returns, total volume, and net initiated order flow on the days before and after each quarterly earnings announcement.

---

12 Large firms in the sample include companies such as Dell, Intel, and Microsoft.

13 Weekend chat is aggregated into Monday's chat. Hence, the value of $Chat_{it}$ on Monday includes all messages since the market close on the previous Friday. Similarly, chat during market holidays is aggregated into the following day. Screening firms with fewer than one, three, or five chat messages per day yields similar results.

Unadjusted daily stock returns ($R_{it}$) are defined as the log of the ratio of the closing price on day t to the closing price on day t-1. Our measure of daily market-adjusted "abnormal" stock returns (Adjusted $R_{it}$), is calculated as the unadjusted return minus the return on the Nasdaq Composite Index over the same period. Daily total volume for each stock ($VOL_{it}$) is obtained by aggregating the dollar value of all trades during normal trading hours (9:30a.m. - 4:00p.m., E.S.T.).

In order to generate measures of daily net initiated order flow for every stock, we first determine whether each trade is initiated by a buyer or a seller. Consistent with Trueman et al., (2003), we apply the 'tick rule' to classify trades as buyer-initiated or seller-initiated, according to whether the most recent non-zero price change was positive or negative.[14] The dollar value of individual trades initiated by buyers and sellers, respectively, is then aggregated to obtain total buyer-initiated and seller-initiated order flow during the day ($BIOF_{it}$ and $SIOF_{it}$, respectively). Total daily net initiated order flow is then defined as the difference between buyer-initiated and seller-initiated order flow, $NIOF_{it} = BIOF_{it} - SIOF_{it}$.

Since we are interested in the differential behavior of trading activity by retail versus institutional investors, we use two partitioning schemes to classify all transactions into small or large trades. First, we use a single fixed cutoff of $20,000, whereby all trades less than or equal to $20,000 are classified as retail trades, while all trades greater than $20,000 are labelled institutional trades. Second, we also use a partitioning scheme in which we exclude medium-sized trades, and split trades into retail trades with a value of $10,000 or less, versus institutional trades with a value larger than $50,000. These trade-classification schemes yield several measures of buyer- and seller-initiated order flow for small and large investors, respectively ($BIOF{\leq}10_{it}$ and $BIOF{\leq}20_{it}$ versus $BIOF{>}20_{it}$ and $BIOF{>}50_{it}$, and $SIOF{\leq}10_{it}$ and

14 If the price change from the previous trade was zero, this trade is still classified as buyer-initiated or seller-initiated based on whether the most recent non-zero price change was positive or negative (Lee and Ready 1992). The first few trades each day remain unclassified until the first non-zero price change of the day occurs. Thereafter, all trades during the day are classified as either buyer-initiated or seller-initiated. Ellis, Michaely and O'Hara (2000) document that the tick rule correctly classifies 78% of the trades for Nasdaq stocks.

SIOF≤$20_{it}$ versus SIOF>$20_{it}$ and SIOF>$50_{it}$). Finally, we define separate measures of daily net initiated order flow by size of trade, subtracting seller-initiated order flow (SIOF) from buyer-initiated order flow (BIOF) for all trade size categories. This process produces the following measures: NIOF≤$10_{it}$ for all trades of $10,000 or less, NIOF≤$20_{it}$ for all trades of $20,000 or less, NIOF>$20_{it}$ for all trades larger than $20,000, and NIOF>$50_{it}$ for all trades larger than $50,000.[15] All variables are defined in Table 1.

*3.2 Research Design*

We initially focus on the behavior of internet chat activity, stock prices, trading volume, and net order imbalance over the 20-day period extending from ten days prior to ten days after each quarterly earnings announcement during 1998 (t = -10, …, -1, 1, …, 10). We follow the convention in Trueman et al., (2003), to label the day before (after) the earnings announcement date as day –1 (+1), since all our announcements occur outside regular trading hours (i.e., after the close on day –1 and prior to the opening on day +1).

In order to make the variables comparable across firms, we standardize all variables (other than stock returns) according to the 'normal' behavior for each firm prior to all quarterly earnings announcement periods. Our standardization scheme for each firm involves first taking the difference between every variable and its 'normal' mean, and then dividing by its 'normal' standard deviation. For each firm, the normal mean and standard deviation are calculated using a sample of (up to) 40 days, comprised of the 10 days prior to all (up to) four earnings announcement periods (from day -20 through day -11). The variable names for the

15 Trueman et al., (2003) analyze the net *number* of all buyer-initiated trades, rather than the *dollar value* of net initiated order flow. Their approach is consistent with the work of Jones et al., (1994), who document that the number of trades captures all relevant information contained in transactions data regarding the association between trading activity and absolute price changes in a sample of NASD stocks. On the other hand, Chan and Fong (2000) find that the number of trades does not capture all relevant information in trade data. When we replace the dollar value of net initiated order flow with the net number of initiated trades, we obtain similar results (available on request).

standardized variables are as before, but are now preceded by "STD". For example, $STDChat_{it}$ denotes the standardized chat activity for stock i on day t.

*3.3 Descriptive Statistics*

Table 2 provides summary statistics for the daily variables used in this study. All descriptive statistics in Table 2 are calculated across all days in the announcement window (day -10 through day +10), and across all 691 quarterly earnings announcements in our sample. In Panel A of Table 2, we provide results for firm market value, stock returns, chat activity, and net initiated order flow by size of trade.

First consider the data before our standardization scheme is applied. The average firm value across our sample firms is $8.4 billion, and ranges from $10 million to a maximum of $293 billion for Microsoft. The daily close-to-close returns on the Nasdaq Composite Index in this period average 0.184%, reflecting that 1998 was a positive year for technology stocks overall. The average daily return on our sample stocks is somewhat lower at 0.126%. Chat activity ranges from a minimum of 0 to a maximum of 1067 messages per day, and averages 16 messages per day. The mean daily net initiated order flow for small trades using the $10,000 cutoff is $175 thousand dollars, and ranges from -$12.6 million to $36.2 million. The mean and range of daily net order flow tend to increase as we move to higher cutoffs.

After standardization, all variables display positive means and standard deviations greater than one. This outcome indicates that the 10-day periods of 'normal' behavior (over days –20 through –11) tend to have a lower mean and standard deviation than the 20-day period immediately around the earnings announcement.

These firm-specific variables display considerable skewness and/or kurtosis that are inconsistent with normality. For example, the daily data on firm market value and stock returns have substantive kurtosis. Before standardization, the chat and net order flow variables all display positive skewness and also exhibit substantive kurtosis. After

standardization, the positive skewness disappears for net order flow of large trades (> $20,000 or $50,000), and the substantive kurtosis remains.

In Panel B of Table 2 we present summary statistics for total daily trading volume by size of trade, as a proportion of total daily trading activity. Many of these proportional variables also display nonnormality in the form of positive skewness and/or kurtosis. It is noteworthy that the total daily volume of small trades ≤ $10,000 (or ≤ $20,000) averages 19% (or 35%) of average total daily trading volume in these stocks. These figures indicate that retail investors accounted for a substantial proportion of total trading activity in these sample stocks during 1998.

Table 3 provides an overview of the daily patterns in average standardized chat activity and trading volume, for the 20 days around earnings announcements. Both average daily chat activity and trading volume increase as the earnings announcement approaches, peaking on the day after the announcement. The t-statistic provided for each daily mean in Table 3 tests the null hypothesis that average standardized trading volume or chat activity is zero on that day. This can be interpreted as a test of whether average daily chat or volume is significantly different from that over the 'normal' period for each firm, from day -20 through -11. The positive means indicate that average standardized trading volume is higher for all days during the 20-day period immediately surrounding the earnings announcement, than it is during the 'normal' period (and significantly higher from day –9 through day +9). A similar pattern exists for standardized chat, which is significantly higher than chat during the 'normal' period from day –8 through day +5.

## 4. Stock returns and net order flow around earnings announcements

This section presents evidence on the average daily stock returns and standardized net order flow experienced over the 20 days surrounding earnings announcements, for our sample of 206 technology stocks during 1998.

*4.1 Returns around earnings announcements*

The second and third columns of Table 4 provide the mean unadjusted close-to-close returns for all days in the 20-day announcement window, along with their corresponding t-statistics. The mean return for a given day in the announcement window is obtained by averaging across all 691 quarterly announcements in the sample. The t-statistic is based on the cross-sectional standard error. Following Trueman et al., (2003) we also decompose the close-to-close return on day +1 into the return from the close on day -1 to the open on day +1 and the return from the open on day +1 to the close on that day. The last two columns in Table 4 give the analogous mean daily market-adjusted abnormal returns and their t-statistics.

Results indicate that the unadjusted returns are significantly positive for the five days immediately before the announcement, and then turn negative for three of the four days after the earnings announcement (with days +1 and +3 significantly negative). The market-adjusted returns yield a similar pattern. They are significantly positive for the two days immediately preceding the earnings announcement, and thereafter turn negative for the 4 days immediately after the announcement (with days +1, +3, and +4 significantly negative). Outside the days immediately surrounding the earnings announcement, none of the (un)adjusted returns are significantly different from zero. Figure 1 plots the cumulative raw and abnormal returns derived from the results in Table 4.

The pattern in stock returns for our sample is similar to the results presented in Trueman et al., (2003) albeit substantially smaller in magnitude. For example, Trueman et al. report that the cumulative abnormal return measured from the close on day –6 to the open on day +1 equals 4.9%, with a t-statistic of 11.4. For our sample, the cumulative abnormal return measured over the same period equals 1.8% and has a t-statistic of 3.7. Likewise, the cumulative abnormal return from the open on day +1 to the close on day +5 is reported as

-6.4% in Trueman et al., (2003) with a t-statistic of -19.2, whereas the analogous cumulative return is -2.4% for our sample, with a t-statistic of -4.8.

One noteworthy difference in our results from those documented in Trueman et al., (2003) is the abnormal return from the close on day -1 to the open on day +1. Whereas Trueman et al. report a positive abnormal return of 1.6% during this overnight period when the announcement occurred, we find a negative abnormal (i.e., adjusted) return of -0.5%. This difference accounts for a large portion of the differential cumulative abnormal return found in our study. Note that Trueman et al. use a longer sample period (from January 1998 to August 2000) and consider a sample limited to pure Internet stocks.[16]

*4.2 Net Order Flow around earnings announcements*

Table 5 provides information on the average daily standardized net initiated order flow ($STDNIOF_{it}$) for all days over the 20-day announcement window. The second and third columns present the mean standardized net order flow including all trades each day, along with the corresponding t-statistics. The next ten columns provide analogous information for our measures of net initiated order flow for small trades and large trades, respectively.

Table 5 reveals evidence of significant net buying pressure on the days leading up to the earnings announcement, followed by significant net selling pressure after the announcement. Specifically, average net order flow of all trades is positive on all ten days before the announcement, and is significantly positive (at the .05 level or better) on four of the six days immediately before the announcement. On the 10 days following the announcement average net order flow of all trades is negative on 8 out of 10 days, but is significantly negative only on day +3.

16 The average firm value in Trueman et al., (2003) is $5 billion, which is smaller than for our sample firms. Another notable difference is that one third of the announcements studied in Trueman et al. occurred during the six-month IPO lock-up period, whereas all our sample firms had been listed at least six months at the start of our sample period.

Additional insight is gained by analyzing the pattern of behavior by small and large traders separately. Columns 4 to 8 of Table 5 provide results based on the $20,000 cutoff. These results show that the average net buying pressure for all trades before the announcement is comprised of significant net buying by small traders on all six days immediately prior to the announcement (t = –6 through –1). There is also significant net buying by large traders on four of the six days immediately before the announcement (t = -6, -4, -2, and -1). Interestingly, on the day after the announcement (t = +1), small traders continue to exert significant buying pressure, on average, while the net initiated order flow from large traders becomes negative (mean NIOF>20 is significantly negative at the 10% level on day +1). The net initiated order flow from small traders subsequently becomes negative on average for six of the remaining nine days after the announcement, but is significantly negative only on day +8. The net initiated order flow from large traders is also negative on six of the remaining days following the day after the announcement, but is significantly negative only on day +3. The results of mean difference t-tests in Column 8 show that the mean net initiated order flow from small traders is significantly greater than the mean net initiated order flow from large traders on the day before and the day after the earnings announcement (-1 and +1), based on the $20,000 cutoff.

The alternative scheme for distinguishing small and large traders yields similar results. For trades smaller than $10,000 (columns 9 and 10) we find a pattern in standardized net initiated order flow similar to that of trades smaller than $20,000. However, note that the significant buying pressure using the $10,000 cutoff begins one day earlier (with NIOF≤10 significantly positive beginning on day –7). In addition, the net initiated selling by small traders after the earnings announcement begins one day later, and is generally more sporadic and smaller in magnitude when the $10,000 cutoff is used.

The pattern of net initiated order flow for institutional traders using the $50,000 cutoff (columns 11 and 12) is likewise similar to that for trades larger than $20,000. However, when we directly compare the mean daily values of NIOF≤10 and NIOF>50 (in column 13), we now find the mean net initiated order flow from small traders is significantly greater than the mean net initiated order flow from large traders on three days (-3, -1 and +1). Thus, eliminating the buffer zone of medium-size trades reveals larger differences in the net buying of small versus large traders on the days immediately surrounding the announcement.

The results in tables 4 and 5 are consistent with the evidence in Trueman et al., (2003). We find the same curious pattern in abnormal returns around quarterly earnings announcements for our sample of technology stocks during 1998. We also find significant buying pressure on the days leading up to the announcement, which turns to selling pressure on the days following the announcement. Consistent with Lee (1992), we find that small traders are significantly more active net buyers than large traders on the day before and the day after the earnings announcement.

The contrasting behavior of retail and institutional investors on the day after the earnings announcement (day +1) is of particular interest. In Table 3 we find average daily chat activity and total trading volume are greatest on day +1. In Table 4 we document that the average negative stock returns on day +1 are the largest in absolute value. In Table 5 we find retail traders continue to be significant net buyers on day +1, while institutional investors become significant net sellers on this day. A possible explanation for these divergent results is that earnings news may be digested relatively more quickly by institutional investors, who presumably have better and more timely access to the forecasts of professional analysts.[17]

## 5. Stock returns, net order flow, and abnormal chat activity

17 The sample period is prior to Regulation FD, so that some institutional investors may have had access to the earnings announcement before the earnings press release.

We propose that, according to the Miller (1977) hypothesis, if traders are unable or unwilling to (short) sell stocks, an increase in the level of disagreement (say, before an earnings announcement) should result in upward price pressure while a subsequent decline in the level of disagreement should result in a price reversal. The previous section documents patterns in average chat activity, stock returns, and net order flow over the time series of days around earnings announcements that are consistent with this proposition. We further note divergent behavior in the buying patterns of small and large traders, respectively.

This section proposes a formal, cross-sectional test of the Miller (1977) hypothesis by analyzing the interaction of chat activity with stock returns and net order flow by small and large traders, on the days around earnings announcements. Specifically, we examine whether announcements characterized by greater increases in abnormal chat before the announcement experience greater increases in stock prices and net order flow before the announcement, followed by greater reversals after the announcement. Our focus is on the two 4-day periods before and after earnings announcements, when the most significant net buying (selling) and abnormal returns are observed. We then repeat the analysis using an 8-day window.

First we analyze the cross-sectional relation between the change in abnormal chat activity and stock returns. We then focus on the cross-sectional relation from abnormal chat activity to, both, trading volume and net initiated order flow by retail and institutional investors. Finally, we investigate whether the magnitude of the anomalous pattern in stock returns around the announcement is related to the relative amounts of retail vs institutional trading volume in a firm's stock.

### *5.1 Chat Activity and Stock Price Changes*

Miller (1977) argues that, in the presence of (short) selling constraints, the price of a stock is positively related to the level of disagreement among investors about the value of that stock.

Assuming a linear relation between the level of disagreement and the natural log of prices, this implies:

$$\log(P_{i,t}) = \alpha_{i,t} + \beta \text{ Disagreement}_{i,t}; \quad (1)$$

where $P_{i,t}$ is the price of stock i at time t, $\alpha_{i,t}$ is the equilibrium price of stock i at time t in the absence of disagreement, and $\beta > 0$.

To test the Miller (1977) hypothesis, we use standardized chat activity as our proxy for disagreement on day t, and we specify an empirical model in terms of first differences:

$$\Delta\log(P_{i,t,t-1}) = a + b\, \Delta\text{STDChat}_{i,t,t-1} + e_{i,q}. \quad (2a)$$

Alternatively, we can expand the change in standardized chat on the right-hand-side of (2a) to explicitly include the level of disagreement at, both, the beginning and end of the period:

$$\Delta\log(P_{i,t,t-1}) = a + b_1\, \text{STDChat}_{i,t-1} + b_2\, \text{STDChat}_{i,t} + e_{i,q}. \quad (2b)$$

Under the null hypothesis, $H_1$: $b_1 = -b_2$, (2a) and (2b) are equivalent.

We estimate models (2a) and (2b) to assess the relation between the change in chat activity and stock returns over the 4-day periods before and after the earnings announcement, respectively. First, we define the return for the $i^{th}$ stock over the 4-day period before the $q^{th}$ announcement ($\text{RetBefore}_{i,q}$) as the log of the closing price on the day before the announcement (day -1) minus the log of the closing price 4 days earlier (day -5). According to the Miller (1977) hypothesis, we expect the magnitude of this pre-announcement return to be directly related to the magnitude of the change in abnormal chat activity from day -5 through day -1:

$$\text{RetBefore}_{i,q} = a + b\, \Delta\text{ChatBefore}_{i,q} + e_{i,q}; \quad (3a)$$

and

$$\text{RetBefore}_{i,q} = a + b_1\, \text{STDChat}(-5)_{i,q} + b_2\, \text{STDChat}(-1)_{i,q} + e_{i,q}. \quad (3b)$$

In regression (3a) we expect b to be positive; in regression (3b) we expect $b_1$ to be negative, $b_2$ to be positive, and we test $H_1$: $b_1 = -b_2$, to see if the data can discriminate between specifications (3a) and (3b).

Next, we define the return after the earnings announcement ($RetAfter_{i,q}$) as the log of the closing price at the end of the 4-day period following the announcement (day +4) minus the log of the closing price on the day before the announcement (day -1). As before, we expect the magnitude of the stock return after the announcement to be directly related to the change in the level of disagreement during the same period:

$$RetAfter_{i,q} = a + b\,\Delta ChatAfter_{,q} + e_{i,q}; \qquad (4a)$$

and $$RetAfter_{i,q} = a + b_1\,STDChat(-1)_{i,q} + b_2\,STDChat(4)_{i,q} + e_{i,q}. \qquad (4b)$$

Similar to (3a) and (3b) above, in (4a) we expect b to be positive, and in (4b) we expect $b_1$ to be negative, $b_2$ to be positive, and we test $H_1$: $b_1 = -b_2$.

To account for concurrent market movements, the four regression models above are estimated using market-adjusted returns (defined as the unadjusted return minus the return on the Nasdaq Composite Index measured over the same window). In addition, we include the quarterly earnings surprise as an explanatory variable.[18] Finally, we also present the results using an 8-day window before the earnings announcement (from day -9 to -1), and an 8-day window after the announcement (from day -1 to +8).

The results for the models that analyze stock returns and chat activity before the earnings announcement ((3a) and (3b)) are reported in Panel A of Table 6. The first model presented in Panel A uses the change in chat activity over the 4-day period before the announcement; the second model splits this change in chat activity into the levels of

18 The earnings surprise is calculated as the difference between the actual quarterly earnings per share and the most recent analyst forecast preceding the earnings announcement, scaled by the closing stock price on day –10. We have also used alternative measures of earnings surprise, such as the difference between actual quarterly earnings and the mean or median across recent analyst forecasts, and the difference between actual quarterly earnings this period and four quarters ago. The major conclusions are unaffected by this choice.

disagreement at the beginning and end of the 4-day period, respectively. The third model is similar to the second model, but incorporates an 8-day window. Panel B of Table 6 reports results for the analogous models analyzing the stock return and the change in chat activity after the earnings announcement ((4a) and (4b)).

The first model in Panel A (3a) reveals a significant positive relation between stock returns and the change in disagreement over the 4-day period before the earnings announcement. In addition, the return before the announcement is positively related to the earnings surprise. The second model in Panel A (3b) provides more detail to show that the 4-day pre-announcement return is greater if there is a lower level of disagreement (standardized chat activity) at the beginning of the period (on day –5), and a higher level of disagreement at the end of the period (on day –1). We do not reject $H_1$: $b_1 = -b_2$, indicating that the data cannot discriminate between the specification of models (3a) and (3b). These results are consistent with the Miller (1977) hypothesis, indicating that stocks with a relatively high level of disagreement just before the announcement have a relatively high price, and thus tend to have lower future returns. Similar results are found for the 8-day window.

The first model in Panel B of Table 6 (4a) reveals analogous results for the 4-day post-announcement period, in which the magnitude of the stock price decline is directly related to the magnitude of the decline in chat activity over the 4 days following the earnings announcement. Surprisingly, there is no evidence of a significant relation between the post-announcement return and the earnings surprise. The second model in Panel B (4b) once again provides more detail, to show that the magnitude of the price decline from the close on day -1 to the close on day +4 is inversely related to the level of disagreement at the beginning of this post-announcement period, and directly related to the level of disagreement at the end of the period. All coefficients in Panel B are significant at the 10% level or better, and we do not

reject H1: b1 = -b2. Once again, these results are consistent with the Miller (1977) hypothesis, and similar results are found for the 8-day post-announcement window.[19]

To summarize, this analysis demonstrates that the change in the level of disagreement among investors around earnings announcements offers a partial explanation for the anomalous stock return pattern around earnings announcements, by showing that stocks with a larger increase in disagreement before the announcement experience a larger price run-up, and a greater price reversal after the announcement. These results support the basic prediction in Miller (1977), and document the impact of disagreement on stock prices over a much shorter time frame than is examined in previous studies.

### *5.2 Chat Activity, Trading Volume, and Net Order Flow*

In this section we examine the potential sources of the anomalous price behavior documented in the previous section, by analyzing the correlations among chat activity, total trading volume, and net order flow by small and large traders, respectively. We first establish that chat activity is significantly positively related to total trading volume by both retail and institutional investors. This evidence supports the proposition that chat activity serves as a useful proxy for overall disagreement among investors (Antweiler and Frank 2003a, Berkman et al., 2003). Next, we investigate whether chat activity is differentially related to the net initiated order flow from retail versus institutional investors.

#### *5.2.1 Chat Activity and Trading Volume by Small versus Large Investors*

Consider the relation between the level of disagreement (proxied by chat activity) and total trading volume by small and large investors, respectively. Antweiler and Frank (2003a)

19 Our results are not sensitive to outliers. We estimate rank regressions, where the ranked observations for each dependent variable are regressed on the ranked observations of the other variable(s) in the model (the ranked observations for each variable range from 0 for the smallest observation to 690 for the highest observation). Results indicate the coefficients on chat activity have the same signs as in Table 6, and are all significant at the 5% level or better. In all rank regressions the earnings surprise variable is significant and positive. When we use the change in actual chat activity instead of the change in standardized chat activity, similar results are obtained.

and Berkman et al., (2003) find a significant positive relation between the amount of chat activity and total trading volume by both retail and institutional traders. These studies interpret this result as evidence that the amount of chat activity provides a useful proxy for the level of disagreement among investors. We conduct a similar analysis to assess the correlation between standardized chat activity and total trading volume by small and large investors, respectively, for each day over the 20-day period around earnings announcements.

Results are reported in Panel A of Table 7. Columns 2 and 3 give the Pearson correlation coefficients, while column 4 provides the Hotelling (1940) test of the null hypothesis that chat activity has identical correlations with retail and institutional trading volume, respectively. Columns 5 and 6 report the analogous Spearman correlations, while column 7 provides the results of the Hotelling test based on these Spearman correlations.

The most important result in Panel A is that chat activity displays a significant positive correlation with total trading volume by both retail and institutional investors, on all 20 days around earnings announcements. Both Pearson and Spearman correlations are highly significant, indicating that the strong association between chat and trading volume is not an artefact of outliers. This evidence corroborates the work in Antweiler and Frank (2003a) and Berkman et al., (2003), to further establish the usefulness of chat activity as proxy for disagreement among both retail and institutional investors.

It is noteworthy that the Pearson correlation between chat activity and retail trading volume is larger in magnitude than that between chat and institutional trading volume, for 18 of the 20 days (or for 19 of the 20 days using Spearman correlations). The Hotelling (1940) tests using Pearson correlations indicate this difference is significantly positive on 8 of the 20 days at the .10 level or better, while the analogous tests using Spearman correlations indicate

this difference is significantly positive on 5 of the 20 days.[20] These results indicate that chat activity tends to be more strongly related to trading volume by retail traders than by institutional investors. This evidence is also consistent with Berkman et al., (2003).

### *5.2.2 Chat Activity and Net Initiated Order Flow*

Next consider the relation between the level of disagreement (proxied by chat activity) and the net initiated order flow of small and large traders, respectively. In Panel B of Table 7 we provide the Pearson and Spearman correlations between standardized chat activity and net order flow by small versus large traders, respectively, as well as the Hotelling (1940) tests of the null hypotheses that these correlations are identical. This analysis shows how the level of disagreement is related differently to the net buying of small versus large investors.

Several interesting insights are obtained from the Table 7, Panel B. First, consider the correlation between chat activity and net order flow by retail investors. The Pearson correlation is significantly positive at the 10% level or better on 18 of the 20 days. The Spearman correlation is generally smaller in magnitude, but is still significantly positive on 11 of the 20 days. These results indicate that retail investors as a group have an overwhelming tendency to buy more with increases in the level of disagreement before the announcement, and sell more with decreases in disagreement after the announcement. This behavior by retail traders is consistent with the Miller (1977) hypothesis described above.

Second, consider the analogous correlations between chat activity and net order flow by institutional investors. These correlations differ dramatically from those with retail net order flow. The Pearson correlation between chat activity and institutional net order flow is significantly negative on 8 days, significantly positive on just 2 days, and is insignificant on the remaining 10 days. Similarly, the Spearman correlations are significantly negative on 5

20 In contrast, on day +9 the Pearson correlation between chat and retail volume is significantly *smaller* than that between chat and institutional volume, at the .10 level.

days, and are never significantly positive. These results indicate a tendency for institutional investors to avoid buying stocks that are the subject of greater disagreement among investors.

Third, consider the relative magnitudes of the correlations between chat activity and retail versus institutional net order flow. The correlation between retail net order flow and chat activity is greater than the correlation between institutional net order flow and chat activity on 19 of the 20 days. The Hotelling (1940) tests indicate that this difference is statistically significant on 17 days using the Pearson correlation, and on 10 days using the Spearman correlation. Altogether, the results in Panel B of Table 7 point to a strong positive association between the level of disagreement and the net buying of retail investors, but a weaker positive or negative association between disagreement and the net buying of institutional investors.

These divergent results for small versus large net order flow, documented in Panel B of Table 7, call for an explanation. Miller's (1977) theory is based on an asymmetry in the ability of investors to buy versus sell stocks. We argue that this asymmetry in the ability or willingness to buy and sell is likely to differ between retail and institutional traders. If retail investors are less willing or less able to (short) sell than institutional investors, we would expect the Miller (1977) hypothesis to describe the net buying activity of retail investors to a greater extent than that of institutional investors. In the context of this study, this means the relation between disagreement and net order flow should be stronger for retail traders than for institutional traders. The significantly higher correlation between disagreement and retail net order flow documented in Panel B of Table 7 is consistent with our conjecture that retail investors are less willing or able to trade on their negative views than institutional investors.

To summarize, this section provides evidence that chat activity is significantly correlated with total trading volume by both retail and institutional investors. This outcome corroborates other work suggesting that chat activity proxies for disagreement among both

retail and institutional investors (Antweiler and Frank 2003a, Berkman et al., 2003). On the other hand, retail and institutional investors display dramatically divergent behavior when it comes to net buying pressure in relation to chat activity. While retail net initiated order flow displays an overwhelmingly positively correlation with chat activity on the days surrounding earnings announcements, institutional net order flow tends to be negatively related to chat activity. These results indicate that institutional investors avoid buying stocks that are the subject of high chat and buying activity by retail traders. These results are also consistent with the view that retail investors are less willing or less able to (short) sell than are institutional investors, so that greater disagreement among investors results in more buying pressure from retail investors, but not from institutional investors.[21]

*5.3 Institutional versus Retail Trading*

Ofek and Richardson (2003) report that a large proportion of trading in Internet stocks during the Internet boom originated from retail investors. They also show that the first day return for Internet IPO's, and the returns around the expiration of the quiet period following IPO's, are significantly larger if there is a relatively high proportion of retail trading. Panel B of Table 2 indicates that our sample of technology stocks also experienced a relatively large proportion of trading by retail investors during 1998. Furthermore, like Ofek and Richardson, our results in Tables 3-7 point to retail traders as the main cause of the anomalous price behavior around earnings announcements. Finally, consistent with the Miller (1977) hypothesis, we find changes in disagreement are related to the anomalous price behavior around earnings announcements, but that retail and institutional investors display dramatically divergent net buying behavior in relation to the level of disagreement before and after announcements.

21 This result is also consistent with Lee (1992). Further analysis of the relation between net order flow by retail and institutional traders appears to be a fruitful area of inquiry, and should be the subject of future work.

To further investigate whether the anomalous price behavior around earnings announcements is due to trading by retail investors, we regress the pre-announcement return and the post-announcement return, respectively, on the relative amount of retail versus institutional trading activity in a stock, along with the quarterly earnings surprise, as follows:

$$\text{RetBefore}_{i,q} = a + b_1(\text{Retail})_{i,q} + b_2\,\text{Surprise}_{i,q} + e_{i,q}\,, \qquad (5)$$

$$\text{RetAfter}_{i,q} = a + b_1(\text{Retail})_{i,q} + b_2\,\text{Surprise}_{i,q} + e_{i,q}\,. \qquad (6)$$

Similar to the previous section, the pre- and post-announcement return are adjusted for market movements, and are measured over the same 4-day and 8-day windows before and after the earnings announcement. To measure the relative amount of retail versus institutional trading activity around a given stock's earnings announcements, we first compute the daily ratio of total retail trading volume for all trades ≤ \$10,000 ($VOL{\leq}10_{it}$) to institutional trading volume for all trades > \$50,000 ($VOL{>}50_{it}$). Our measure of the relative amount of retail versus institutional trading activity around the $q^{th}$ announcement for the $i^{th}$ stock ($Retail_{i,q}$) is the median value of this ratio across all 20 trading days around the announcement.

Table 8 provides the regression results. The results in Panel A show a significant (at the 10% level) positive relation between the proportion of retail trading and the magnitude of the pre-announcement price runup for that stock, for the 8-day window (the relation is positive but insignificant for the 4-day window). Likewise, in Panel B we find a significant (at the 10% level) negative relation between the relative amount of retail trading and the post-announcement price reversal for both the 4-day and 8-day windows. These results are consistent with Ofek and Richardson (2003), and show that abnormal price increases (decreases) before (after) earnings announcements for technology stocks during 1998 are greater for stocks that have a higher proportion of retail trading around the announcement.

The results in Tables 7 and 8 strongly suggest that the anomalous stock price and order flow patterns observed in Trueman et al., (2003) are related to an asymmetry in the

ability of retail investors to buy versus sell, when the level of disagreement intensifies before earnings announcements. In this light it is interesting to revisit the evidence regarding the relation between chat activity and stock returns around earnings announcements, provided in Table 6. We expect this relation to be stronger when the proportion of retail trading is higher.

To test this hypothesis, we split the sample of earnings announcements into three equal subsamples based on the proportion of retail trading ($Retail_{i,q}$). We exclude all observations in the middle group, and define a dummy variable ($Small_i$) that is equal to 1 when the $i^{th}$ stock is in the group with the highest proportion of retail investors, and zero otherwise. We then estimate regression models similar to (3a) and (4a), allowing both the intercept and the slope coefficient on the change in chat activity to differ between the groups with the highest and lowest proportion of retail trading, as follows:

$$RetBefore_{i,q} = a_0 + a_1\, Small_i + b_0\, \Delta ChatBefore_{i,q} + b_1\, Small_i * \Delta ChatBefore_{i,q} + e_{i,q}, \quad (7)$$

$$RetAfter_{i,q} = a_0 + a_1\, Small_i + b_0\, \Delta ChatAfter_{i,q} + b_1\, Small_i * \Delta ChatAfter_{i,q} + e_{i,q}. \quad (8)$$

The results of these regressions appear in Table 9. For the 4-day and 8-day periods both before and after the announcement, the coefficient of the change in chat is not significantly different from zero. However, the slope interaction term is always significantly positive. These results indicate that the relation between returns and the change in chat activity, documented previously in Table 6, is only significant for the subsample of stocks with the highest proportion of retail trading. For the group of stocks with the lowest proportion of retail trading, there is no such significant relation between returns and our measure of disagreement. This finding provides further corroboration for our proposition that the anomaly in Trueman et al., (2003) is due to changes in disagreement around earnings

announcements, combined with a sample and time period characterized by a high proportion of retail traders who are less able or less willing to (short) sell than to buy.[22]

## 6. Summary and conclusions

This paper suggests that the anomalous patterns in net-initiated order flow and stock prices around the earnings announcements of Internet stocks, documented in Trueman et al., (2003), are related to changes in the level of disagreement among investors before and after the announcement. We argue that an approaching quarterly earnings announcement is likely to exacerbate the level of disagreement among investors. Consistent with Miller (1977), the intensified disagreement results in upward price pressure before the earnings announcement, as optimistic investors buy while pessimistic investors are either unwilling or unable to sell. After the announcement, when the disagreement dissipates, the initial price increase is naturally followed by a reversal. We further suggest this behavior is more likely to characterize stocks with a relatively high proportion of retail trading activity, since retail investors are likely to be less willing or able to (short) sell than institutional investors.

We provide several cross-sectional tests of these hypotheses, using a proxy for the level of disagreement among investors based on the number of messages posted on the Internet message boards (chat rooms) of a sample of technology firms during 1998. We report a number of novel findings that shed light on the peculiar market behavior around earnings announcements for technology stocks during the Internet boom of the late 1990's. Among these new findings, we document that stocks with a larger increase in the level of disagreement before the earnings announcement tend to have a larger pre-announcement price increase, and a larger price reversal after the announcement.

22 Trueman et al., (2003) report a larger price runup and reversal compared to this paper. According to the explanation proposed in this study, their result may be due to their sample comprising only Internet stocks which are smaller on average and have an even higher proportion of retail trading activity (Berkman et al., 2003, Berger 2003, Ofek and Richardson 2003, Trueman et al.).

We also document differential buying/selling behavior by small versus large investors around earnings announcements. For example, while both small and large investors are net buyers before and net sellers after earnings announcements, large investors are quicker to switch to net sellers as the announcement approaches. Specifically, we find that small traders are significantly more active net buyers than large traders on the day before and the day after the earnings announcement, and that large traders switch to become significant net sellers on the day after the announcement, while small traders continue to be significant net buyers.

We then investigate the association between this divergent net buying activity by small and large investors, and changes in the level of disagreement (proxied by abnormal chat activity) before and after earnings announcements. Our disagreement measure is directly related to net order flow by retail investors, but tends to be negatively related to net order flow by institutional investors. Two possible interpretations of this result include the following: (i) chat may only proxy for disagreement among retail investors, not institutional investors; and (ii) the Miller (1977) hypothesis applies to a greater extent to retail investors in this setting, because they are less able or willing to (short) sell than institutional investors.

Our results do not appear to be consistent with the former interpretation (that chat activity only proxies for disagreement among retail investors), since we observe a high correlation between chat activity and total trading volume by both retail and institutional investors on all 20 days around earnings announcements. Instead, we interpret our evidence as consistent with the second interpretation, suggesting that the Miller (1977) hypothesis especially applies retail investors in this setting.

In this light, we further pursue the extent to which the anomalous price patterns around earnings announcements are due to the proportion of retail trading in a stock. We find the magnitude of the pre-announcement price runup and the post-announcement reversal is directly related to the relative amount of retail versus institutional trading volume in a firm's

stock around the announcement. Furthermore, the relation between stock returns and the change in disagreement is only significant for the group of stocks with the highest proportion of retail trading volume.

It is noteworthy that the divergent behavior of retail versus institutional investors documented in this study is likely to be especially relevant for our sample of mostly small technology firms during the technology boom of the late 1990's. Several recent studies note that retail investors were particularly active in this sector during this period (Berger 2003, Berkman et al., 2003, Ofek and Richardson 2003, and Trueman et al., 2003). Hence, this sample of firms and this time period are especially well suited to study the influence of retail versus institutional trading activity on price formation. On the other hand, the nature of this sample limits the generalizability of our results, suggesting the need for further research using a broader sample of companies over a wider time period.

## References


Admati, Anat R., and Paul C. Pfleiderer. "Disclosing Information on the Internet: Is It Noise or Is It News?" Manuscript, 2001.

Antweiler, Werner, and Murray Frank. "Is All That Talk Just Noise? The Information Content of internet Stock Message Boards," forthcoming, *Journal of* Finance, 2003a.

_____. "Internet Stock Message Boards and Stock Returns," manuscript, 2003b.

Barber, Brad, and Terence Odean. "The Courage of Misguided Convictions: the Trading Behavior of Individual Investors", *Financial Analyst Journal*, 1999, Nov./ Dec., 41-55.

Atiase, R. K., and Linda S. Bamber. "Trading Volume Reactions to Annual Accounting Earnings Announcements," *Journal of Accounting and Economics*, 1994, v17, 309-331.

_____. "All that Glitters: the Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors," manuscript, 2003.

Berger, Philip G. "Discussion of 'Anomalous Stock Returns Around Internet Firms' Earnings Announcements," *Journal of Accounting and Economics*, 34 (2003) 273–281.

Berkman, Henk, Michael Ettredge, Paul Koch, Vernon Richardson, and Susan Scholz. "Internet Chat, Retail Investor Disagreement, and Trading Volume," manuscript, 2003.

Bhattacharya, Nilabhra. "Investors' Trade Size and Trading Responses around Earnings Announcements: An Empirical Investigation," *Accounting Review*, 2001, v76 (2, April), 221-244.

Chan, L. K. C., and J. Lakonishok. "Institutional Traders and Intraday Stock Price Behavior," *Journal of Financial Economics*, 1993, v33 (April), 173-189.

Cready, W. M. "Information Value and Investor Wealth: The Case of Earnings Announcements," *Journal of Accounting Research*, 1988, v26, 1-27.

Cao, Henry H., Joshua D. Coval, and David Hirschleifer. "Sidelined Investors, Trade-Generated News, and Security Returns," *Review of Financial Studies*, 15 (2002) 615-648.

Chan, Kalok and Wai Ming Fong. “Trade Size, Order Imbalance, And The Volatility--Volume Relation,” *Journal of Financial Economics*, 2000, v57(2,Aug), 247-273.

Chen, Joseph, Harrison Hong, and Jeremy C. Stein. “Breadth of Ownership and Stock Returns,” forthcoming *Journal of Financial Economics*, 2003.

Das, Sanjiv, Asis Martinez-Jerez, and Peter Tufano. “e-Information,” manuscript, 2001.

Demarzo, Peter M., Dimitri Vayanos, and Jefferey Zwiebel, “Persuasion bias and social influence,” working paper, 2001, Stanford University.

Diamond, Douglas W. and Robert E. Verrecchia. “Constraints On Short-Selling And Asset Price Adjustment To Private Information,” *Journal of Financial Economics*, 1987, v18(2), 277-312.

Diether, Karl, Christopher Malloy and Anna Sherbina. “Differences of Opinion and the Cross-Section of Stock Returns,” *Journal of Finance*, 2002, v57(5), 2113-2141.

Duffie, Darrell, Nicolae Garleanu, and Lasse Pedersen. “Securities Lending, Shorting and Pricing,” forthcoming, *Journal of Financial Economics*., 2001.

Easley, E., and M. O’Hara. “Price, Trade Size, and Information in Securities Markets,” *Journal of Financial Economics*, 1987, v19 (September), 69-90.

Ellis, K., Michaely, R., O’Hara, M., 2000. “The Accuracy of Trade Classification Rules: Evidence from Nasdaq,” *Journal of Financial and Quantitative Analysis*, 35, 529–551.

Geczy, Musto, and Reed, 2002.

Hasbrouck, J. “Trades, Quotes, Inventories and Information,” *Journal of Financial Economics*, 1988, v22 (December), 229-252.

_____. “Measuring the Information Content of Stock Trades,” *Journal of Finance*, 1991, v46 (March), 179-207.

Hotelling, H. "The Selection of Variates for Use in Prediction with Comments on the General Problem of Nuisance Parameters," *Annals of Mathematical Statistics*, 1940, v11, 271-283.

Jones, C. M., G. Kaul, and M. L. Lipson. "Transactions, Volume and Volatility", *Review of Financial Studies*, 1994, v7(4), 631-651.

Lee, Charles, 1992. "Earnings News and Small Traders: an Intraday Analysis." *Journal of Accounting and Economics*, v15, 265–302.

Lee, Charles, and Balkrishna Radhakrishna. "Inferring Investor Behavior: Evidence from TORQ Data," *Journal of Financial Markets*, 2000, v3, 83-111.

Lee, Charles, and M. Ready. "Inferring Trade Direction from Intraday Data," *Journal of Finance*, 1992, v46, 733-746.

Lintner, John. "The Aggregation of Investor's Diverse Judgments and Preferences in Purely Competitive Strategy Markets," *Journal of Financial and Quantitative Analysis*, 1969, V4, 347-400.

Miller, Edward M. "Risk, Uncertainty, and Divergence of Opinion," *Journal of Finance*, 1977, v32, 1151-1168.

Morris, Stephen. "Speculative Investor Behavior And Learning," *Quarterly Journal of Economics*, 1996, v111(4,Nov), 1111-1133.

Ofek, Eli, and Matthew Richardson. "Dotcom Mania: The Rise and Fall of Internet Stock Prices," *Journal of Finance*, 2003a, v53, 1113-1137.

Trueman, B., Wong, M., Zhang, X. "Anomalous Stock Returns Around Internet Firms' Earnings Announcements," *Journal of Accounting and Economics*, 2003, v34, 249-271.

Tumarkin, Robert, and Robert F. Whitelaw. "News or Noise? Internet Message Board Activity and Stock Prices," *Financial Analysts Journal*, 2001, v57(3), 41-51.

Utama, S., and W. M. Cready. "Institutional Ownership, Differential Predisclosure Precision and Trading Volume at Announcement Dates," *Journal of Accounting and Economics*, 1997, v24, 129-150.

## Table 1. Variable Definitions

### Chat Activity and Stock Returns

$Chat_{it}$ = number of messages posted on company chat room for stock i on day t. Daily chat is measured from the previous day's market close to the current day's market close (4:00p.m. to 4:00p.m. E.S.T.). The chat data are from http://www.siliconinvestor.com.

Unadjusted $Return_{it}$ = $\ln(P_{i,t} / P_{i,t-1})$, where $P_{i,t}$ is the closing price on day t for stock i.

Adjusted $R_{it} = \ln(P_{i,t} / P_{i,t-1}) - \ln(NASDAQ_t / NASDAQ_{t-1})$, where $NASDAQ_t$ is the NASDAQ Composite index at the end of day t.

### Dollar Volume

$VOL_{it}$ = dollar volume of all trades in stock i on day t.

### Buyer-Initiated and Seller-Initiated Order Flow

$BIOF_{it}$ = Buyer-Initiated Order Flow across all trades for firm i on day t;
$SIOF_{it}$ = Seller-Initiated Order Flow across all trades for firm i on day t;

$BIOF{\leq}10_{it}$ = Buyer-Initiated Order Flow for all trades of $10,000 or less;
$SIOF{\leq}10_{it}$ = Seller-Initiated Order Flow for all trades of $10,000 or less;

$BIOF{\leq}20_{it.q}$ = Buyer-Initiated Order Flow for all trades of $20,000 or less;
$SIOF{\leq}20_{it}$ = Seller-Initiated Order Flow for all trades of $20,000 or less;

$BIOF{>}20_{it}$ = Buyer-Initiated Order Flow for all trades of more than $20,000;
$SIOF{>}20_{it}$ = Seller-Initiated Order Flow for all trades of more than $20,000;

$BIOF{>}50_{it}$ = Buyer-Initiated Order Flow for all trades of more than $50,000;
$SIOF{>}50_{it}$ = Seller-Initiated Order Flow for all trades of more than $50,000.

### Net Initiated Order Flow

All trades: $NIOF_{it} = BIOF_{it} - SIOF_{it}$ (all trades for firm i on day t);
Small trades of $10,000 or less: $NIOF{\leq}10_{it} = BIOF{\leq}10_{it} - SIOF{\leq}10_{it}$;
Small trades of $20,000 or less: $NIOF{\leq}20_{it} = BIOF{\leq}20_{it} - SIOF{\leq}20_{it}$;
Large trades of more than $20,000: $NIOF{>}20_{it} = BIOF{>}20_{it} - SIOF{>}20_{it}$;
Large trades of more than $50,000: $NIOF{>}50_{it} = BIOF{>}50_{it} - SIOF{>}50_{it}$.

### Standardization Scheme

We account for differences in chat and trading activity across firms, by standardizing the relevant variables according to the behavior of each firm over a 'normal' sample of days prior to the earnings announcement period. Specifically, we standardize the relevant variables by first taking the difference between every variable and its 'normal' mean, and then dividing by its 'normal' standard deviation. For each firm, the normal mean and standard deviation are calculated using a sample of (up to) 40 days, from 20 days before through 11 days before (t = -20, …, -11) all earnings announcements in 1998. Standardized variables have the same name as those defined above, but these names are preceded by "STD". For example, $STDChat_{it}$ denotes the standardized chat activity for stock i on day t.

Table 2
Descriptive statistics for daily data

The sample consists of 20-day announcement periods around 691 quarterly earnings announcements that occurred outside normal trading hours, for 206 publicly traded technology firms with an average of at least two messages per day posted in the chat rooms of www.siliconinvestor.com during1998. Daily market value of common shareholders' equity is calculated using the daily closing price multiplied by the number of shares outstanding. The daily return for our sample stocks is based on the log of relative daily closing prices, as is the daily Nasdaq return. For ease of exposition, both return measures are multiplied by 100, market value is divided by 1,000,000 and the net order flow and volume measures are divided by 1000. All other variables in this table are defined in Table 1.

Panel A. Firm Size, Stock Returns, Chat Activity, and Net Initiated Order Flow

| | N | Mean | Median | Std Dev | Skewness | Kurtosis | Min | Max |
|---|---|---|---|---|---|---|---|---|
| Market Value[a] | 13,152 | 8,404 | 864 | 26,205 | 5.73 | 41.49 | 10.00 | 293,032 |
| Nasdaq Return (%) | 13,606 | 0.184 | 0.27 | 1.55 | -.40 | 2.97 | -8.95 | 11.4 |
| Stock Return (%) | 13,606 | 0.126 | 0 | 5.94 | -1.11 | 38.04 | -108.86 | 102.38 |
| Chat | 13,636 | 16.47 | 5 | 40.35 | 7.36 | 90.33 | 0 | 1067 |
| NIOF[b] | 13,636 | 3,050 | 32 | 23,004 | 3.18 | 96.18 | -398,790 | 484,848 |
| NIOF<10[b] | 13,636 | 175 | .46 | 1,326 | 8.57 | 148.44 | -12,565 | 36,182 |
| NIOF<20[b] | 13,636 | 352 | .69 | 2,574 | 5.91 | 82.68 | -30,962 | 58,241 |
| NIOF>20[b] | 13,636 | 2,698 | 26 | 22,498 | 2.77 | 103.49 | -414,593 | 483,769 |
| NIOF>50[b] | 13,636 | 2,306 | 0 | 21,899 | 2.63 | 113.07 | -414,996 | 487,102 |
| STDChat | 13,636 | .417 | -.186 | 2.059 | 5.62 | 60.69 | -1.86 | 42.99 |
| STDNIOF | 13,636 | .038 | -.048 | 1.560 | -1.31 | 74.34 | -38.18 | 25.84 |
| STDNIOF<10 | 13,558 | .093 | -.105 | 1.600 | 5.12 | 81.61 | -10.19 | 41.24 |
| STDNIOF<20 | 13,636 | .075 | -.113 | 1.600 | 5.02 | 70.74 | -11.18 | 36.80 |
| STDNIOF>20 | 13,636 | .025 | .003 | 1.680 | -5.07 | 147.21 | -48.71 | 25.15 |
| STDNIOF>50 | 13,541 | .006 | .030 | 2.020 | -12.59 | 579.97 | -102.40 | 30.95 |

Panel B. Retail Trading Volume Relative to Total and Institutional Trading Volume

| | N | Mean | Median | Std Dev | Skewness | Kurtosis | Min | Max |
|---|---|---|---|---|---|---|---|---|
| VOL[b] | 13,636 | 88,769 | 13,173 | 262,290 | 6.86 | 66.71 | 2,031 | 5,172,007 |
| VOL<10[b] | 13,606 | 3,327 | 1,144 | 7,470 | 6.93 | 79.08 | 0 | 176,057 |
| VOL<20[b] | 13,606 | 8,754 | 2,988 | 18,859 | 6.64 | 75.63 | 2,031 | 459,481 |
| VOL>20[b] | 13,636 | 80,016 | 9,030 | 248,254 | 7.00 | 69.23 | 0 | 4,935,171 |
| VOL>50[b] | 13,636 | 65,726 | 5,632 | 218,659 | 7.21 | 72.47 | 0 | 4,283,337 |
| VOL<10 / VOL | 13,636 | .191 | .097 | .218 | 1.66 | 2.15 | 0 | 1.00 |
| VOL<20 / VOL | 13,636 | .351 | .278 | .276 | .72 | -.57 | .01 | 1.00 |
| VOL>20 / VOL | 13,636 | .649 | .722 | .276 | -.72 | .57 | 0 | .99 |
| VOL>50 / VOL | 13,636 | .435 | .441 | .272 | .00 | -1.05 | 0 | .98 |
| VOL<10 / VOL>50 | 12,256 | .984 | .175 | 3.137 | 10.30 | 165.12 | 0 | 79.21 |

[a] 1,000,000's of dollars. [b] 1,000's of dollars.

Table 3
Average daily standardized trading volume and chat activity

This table presents the mean daily standardized trading volume and chat activity across all 691 quarterly earnings announcements in the sample for each day over the 20-day announcement period (t = -10, …, -1, 1, …, 10). All earnings announcements in the sample occur outside normal trading hours, after the close on day -1 and before the open on day +1. Trading day -1 (+1) is the trading day immediately preceding (following) the earnings announcement. Trading volume, chat activity, and the standardization scheme are described in Table 1. The t-statistic associated with each mean is based on the cross-sectional standard error, and tests the null hypothesis that mean daily behavior is no different than that over the 'normal' period from day -20 through -11.

| Day | Mean STDVOL | t-statistic | Mean STDChat | t-statistic |
|---|---|---|---|---|
| -10 | .086 | 1.62 | .018 | .38 |
| -9 | .138 | 2.04 ** | .082 | 1.39 |
| -8 | .114 | 1.82 * | .148 | 2.33 ** |
| -7 | .199 | 2.58 ** | .175 | 3.34 *** |
| -6 | .313 | 2.97 *** | .184 | 3.12 *** |
| -5 | .337 | 2.14 ** | .123 | 1.96 ** |
| -4 | .145 | 2.60 ** | .119 | 2.31 ** |
| -3 | .180 | 3.04 *** | .221 | 3.60 *** |
| -2 | .477 | 4.86 *** | .352 | 5.73 *** |
| -1 | .864 | 5.49 *** | .662 | 8.09 *** |
| 1 | 2.992 | 13.94 *** | 3.550 | 20.08 *** |
| 2 | .964 | 8.70 *** | 1.222 | 12.68 *** |
| 3 | .557 | 7.31 *** | .581 | 8.50 *** |
| 4 | .360 | 5.45 *** | .282 | 4.83 *** |
| 5 | .244 | 5.12 *** | .114 | 2.31 ** |
| 6 | .305 | 4.28 *** | .035 | .68 |
| 7 | .298 | 3.64 *** | .013 | .27 |
| 8 | .368 | 2.56 ** | .065 | 1.31 |
| 9 | .282 | 3.41 *** | .046 | .99 |
| 10 | .050 | .49 | -.039 | -.40 |

* indicates significance at the.10 level; ** at the.05 level; *** at the .01 level.

Table 4
Average daily returns

This table presents the mean unadjusted return and market-adjusted return across all 691 quarterly earnings announcements in the sample, for each day over the 20-day announcement period. Trading day -1 (+1) is the trading day immediately preceding (following) the earnings announcement. The daily unadjusted return on trading day t is the (log) close-to-close return. The daily market-adjusted return is the unadjusted return minus the (log) close-to-close return on the Nasdaq Composite market index. The close-to-open return on day +1 is the return from the close on day -1 to the open on day +1 (the time interval when the announcement was released). The open-to-close return on day +1 is the return from the open to the close on the trading day after the announcement. The t-statistic associated with each mean daily return is based on the cross-sectional standard error, and tests the null hypothesis that the mean daily return equals zero.

| Day | Unadjusted Return | t-statistic | Adjusted Return | t-statistic |
|---|---|---|---|---|
| -10 | .0013 | .55 | .0002 | .10 |
| -9 | -.0019 | -.79 | -.0018 | -.81 |
| -8 | -.0009 | -.37 | -.0010 | -.41 |
| -7 | .0024 | 1.07 | -.0004 | -.18 |
| -6 | .0018 | .96 | .0001 | .08 |
| -5 | .0057 | 2.41 ** | .0028 | 1.23 |
| -4 | .0067 | 3.38 *** | .0023 | 1.30 |
| -3 | .0061 | 3.07 *** | .0027 | 1.48 |
| -2 | .0136 | 6.63 *** | .0098 | 5.04 *** |
| -1 | .0093 | 4.83 *** | .0060 | 3.36 *** |
| closetoopen | -.0038 | -1.27 | -.0054 | -1.82 * |
| opentoclose | -.0104 | -4.05 *** | -.0116 | -4.64 *** |
| 1 | -.0142 | -3.61 *** | -.0170 | -4.39 *** |
| 2 | .0006 | .28 | -.0018 | -.91 |
| 3 | -.0064 | -2.55 ** | -.0072 | -2.93 *** |
| 4 | -.0023 | -1.18 | -.0040 | -2.19 ** |
| 5 | .0021 | 1.07 | .0003 | .14 |
| 6 | .0013 | .77 | .0012 | .74 |
| 7 | .0009 | .48 | -.0003 | -.18 |
| 8 | .0003 | .20 | .0007 | .41 |
| 9 | -.0012 | -.76 | -.0022 | -1.46 |
| 10 | -.0058 | -1.06 | -.0077 | -1.42 |

* indicates significance at the .10 level; ** at the .05 level; *** at the .01 level.

Table 5
Average daily standardized net initiated order flow

This table presents the mean daily standardized net initiated order flow by trade size across all 691 quarterly earnings announcements in the sample, for each day over the 20-day announcement period. Trading day -1 (+1) is the trading day immediately preceding (following) the announcement. The mean daily standardized net initiated order flow is provided for all trades during the day, for all small trades (less than $10,000 or $20,000), and for all large trades (greater than $20,000 or $50,000). All net order flow variables and the standardization scheme are described in Table 1. The t-statistic associated with each mean and mean-difference is based on the cross-sectional standard error. Trading day –1 (1) is the trading day immediately preceding (following) the earnings announcement, and tests the null hypothesis that mean daily behaviour is no different than that over the 'normal' period from day -20 through -11.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Day | STD NIOF | t-stat | STD NIOF<20 | t-stat | STD NIOF>20 | t-stat | Difference Small-Large | STD NIOF<10 | t-stat | STD NIOF>50 | t-stat | Difference Small-Large |
| -10 | .048 | 1.02 | .035 | .73 | .037 | .79 | -.04 | .039 | .77 | .060 | 1.19 | -.31 |
| -9 | .031 | .74 | .043 | .68 | .029 | .68 | .19 | .038 | .74 | .054 | 1.17 | -.23 |
| -8 | .054 | 1.16 | .026 | .43 | .079 | 1.67 * | -.69 | .053 | .95 | .071 | 1.41 | -.24 |
| -7 | .041 | .90 | .047 | .84 | .051 | 1.07 | -.06 | .095 | 1.72 * | -.011 | -.20 | 1.38 |
| -6 | .136 | 2.33 ** | .144 | 2.30 ** | .119 | 2.10 ** | .29 | .177 | 2.56 ** | .014 | .17 | 1.49 |
| -5 | .067 | 1.37 | .109 | 2.18 ** | .034 | .68 | 1.07 | .105 | 1.93 * | .031 | .46 | .84 |
| -4 | .108 | 2.37 ** | .101 | 2.25 ** | .116 | 2.50 ** | -.24 | .109 | 2.26 ** | .087 | 1.64 * | .32 |
| -3 | .059 | 1.06 | .103 | 2.12 ** | .002 | .03 | 1.37 | .114 | 2.31 ** | -.015 | -.28 | 1.74 * |
| -2 | .309 | 4.95 *** | .320 | 4.38 *** | .275 | 4.46 *** | .47 | .282 | 5.03 *** | .185 | 3.48 *** | 1.25 |
| -1 | .226 | 4.19 *** | .314 | 4.92 *** | .153 | 2.70 *** | 1.88 * | .302 | 5.37 *** | .054 | .92 | 3.05 *** |
| 1 | -.083 | -.84 | .384 | 3.51 *** | -.211 | -1.82 * | 3.74 *** | .453 | 4.25 *** | -.249 | -1.94 * | 4.21 *** |
| 2 | -.049 | -.72 | .015 | .22 | -.009 | -.11 | .23 | .073 | .86 | .057 | .71 | .14 |
| 3 | -.161 | -2.75 *** | -.038 | -.70 | -.120 | -1.98 ** | 1.01 | .004 | .07 | -.074 | -1.06 | .87 |
| 4 | -.031 | -.71 | -.027 | -.60 | -.027 | -.59 | .00 | -.012 | -.24 | -.021 | -.44 | .12 |
| 5 | -.032 | -.68 | -.035 | -.80 | -.063 | -.89 | .33 | .009 | .22 | -.156 | -1.03 | 1.05 |
| 6 | .074 | 1.15 | .009 | .15 | .075 | 1.39 | -.84 | .043 | .71 | .072 | 1.24 | -.35 |
| 7 | .059 | 1.22 | .022 | .38 | .055 | 1.11 | -.43 | .029 | .48 | .093 | 1.76 * | -.80 |
| 8 | -.084 | -.98 | -.082 | -1.69 * | -.101 | -.95 | .16 | -.074 | -1.42 | -.111 | -.98 | .30 |
| 9 | -.020 | -.35 | -.048 | -.85 | .019 | .34 | -.85 | -.068 | -1.46 | .022 | .34 | -1.14 |
| 10 | -.141 | -1.65 * | -.016 | -.18 | -.139 | -1.56 | .99 | .017 | .16 | -.200 | -1.75 * | 1.41 |

* indicates significance at the .10 level; ** at the .10 level; *** at the.01 level.

Table 6
OLS regressions of stock returns on the change in chat activity before and after the earnings announcement

This table presents the results of OLS regressions of the pre- and post-announcement returns on the change in standardized chat activity. Results for the models relating stock returns and chat activity before (after) the announcement are provided in panel A (panel B). The 4-day (8-day) pre-announcement return in Panel A is the log of the ratio of the closing price on the day before the earnings announcement to the closing price on day -5 (day -9). This stock return is adjusted for market movements by subtracting the return on the Nasdaq Composite Index over the same period. Similarly, the 4-day (8-day) post announcement return is the log of the ratio of the closing price on day +4 (day +8) to the closing price on day -1, and is also adjusted for market movements. The change in standardized chat activity, ΔChatBefore or ΔChatAfter, is measured over the same 4-day (8-day) period. STDChat(i) is the standardized chat activity on day i relative to the earnings announcement. The earnings surprise is defined as the difference between the most recent analyst forecast and actual quarterly earnings, scaled by the closing stock price on day -10.

Panel A. Relation between Market-Adjusted Return and Chat Activity Before the Earnings Announcement

| Model | | Intercept | ΔChatBefore | STDChat(-5 or -9) | STDChat(-1) | Surprise | $F(b_1=-b_2)$ | Adj $R^2$ | Overall F |
|---|---|---|---|---|---|---|---|---|---|
| (3a) | 4-days | .020 | .008 | | | .559 | | .051 | 18.87*** |
| | | 5.66*** | 5.50*** | | | 2.56** | | | |
| (3b) | 4-days | .021 | | -.011 | .007 | .575 | 2.39 | .053 | 13.40*** |
| | | 5.86*** | | -4.92*** | 4.22*** | 2.63*** | | | |
| (3b) | 8-days | .017 | | -.011 | .013 | .378 | 0.23 | .055 | 13.83*** |
| | | 3.22*** | | -3.47*** | 5.50*** | 1.20 | | | |

Panel B. Relation between Market-Adjusted Return and Chat Activity After the Earnings Announcement

| Model | | Intercept | ΔChatAfter | STDChat(-1) | STDChat(4 or 8) | Surprise | $F(b_1=-b_2)$ | Adj $R^2$ | Overall F |
|---|---|---|---|---|---|---|---|---|---|
| (4a) | 4-days | -.025 | .006 | | | .152 | | .004 | 2.40* |
| | | -4.50*** | 2.15** | | | 0.44 | | | |
| (4b) | 4-days | -.027 | | -.005 | .008 | .135 | 0.56 | .004 | 1.78 |
| | | -4.55*** | | -1.78* | 2.01** | .39 | | | |
| (4b) | 8-days | -.024 | | -.006 | .012 | -.300 | 1.54 | .010 | 2.31* |
| | | -3.34*** | | -1.73* | 2.27** | -.70 | | | |

* indicates significance at the .10 level; ** at the.05 level; *** at the .01 level.
All regressions are estimated over the sample of N = 664 earnings announcements for which complete data on all regression variables are available.

Table 7. Correlations between standardized chat and total volume or net order flow by trade size

For each day over the 20-day announcement period, this study presents the correlations of standardized chat activity (STDChat) with total trading volume by small versus large traders (Panel A or with net initiated order flow by small versus large traders (Panel B). We also present the Hotelling (1940) T-test of the null hypothesis that standardized chat activity has identical correlations with smal versus large trading volume (Panel A) or with small versus large net initiated order flow (Panel B).

Panel A. Correlations between STDChat (z) and STDVOL<10 (y) versus STDVOL>50 (x)

| | Pearson Correlations | | | Spearman Correlations | | |
|---|---|---|---|---|---|---|
| daynumber | ρ(y,z) | ρ(x,z) | Hotelling T | ρ(y,z) | ρ(x,z) | Hotelling T |
| -10 | .427 *** | .276 *** | 2.66 *** | .300 *** | .201 *** | 1.86 * |
| -9 | .577 *** | .455 *** | 1.86 * | .260 *** | .198 *** | 1.18 |
| -8 | .471 *** | .390 *** | 1.29 | .284 *** | .244 *** | .75 |
| -7 | .532 *** | .389 *** | 2.23 ** | .332 *** | .303 *** | .53 |
| -6 | .473 *** | .331 *** | 2.05 ** | .364 *** | .265 *** | 1.82 * |
| -5 | .436 *** | .305 *** | 1.50 | .350 *** | .287 *** | 1.15 |
| -4 | .340 *** | .260 *** | 1.41 | .284 *** | .224 *** | 1.13 |
| -3 | .316 *** | .188 *** | 2.31 ** | .278 *** | .228 *** | .94 |
| -2 | .314 *** | .261 *** | .88 | .290 *** | .256 *** | .64 |
| -1 | .600 *** | .594 *** | .07 | .303 *** | .319 *** | -.30 |
| 1 | .320 *** | .347 *** | -.46 | .391 *** | .376 *** | .26 |
| 2 | .350 *** | .233 *** | 2.08 ** | .384 *** | .316 *** | 1.22 |
| 3 | .340 *** | .172 *** | 3.15 *** | .327 *** | .272 *** | 1.02 |
| 4 | .389 *** | .305 *** | 1.52 | .342 *** | .269 *** | 1.35 |
| 5 | .291 *** | .238 *** | 1.00 | .335 *** | .246 *** | 1.66 * |
| 6 | .292 *** | .180 *** | 2.11 ** | .311 *** | .205 *** | 2.00 ** |
| 7 | .261 *** | .185 *** | 1.43 | .274 *** | .172 *** | 1.95 * |
| 8 | .244 *** | .220 *** | .33 | .259 *** | .215 *** | .84 |
| 9 | .283 *** | .379 *** | -1.71 * | .240 *** | .221 *** | .36 |
| 10 | .475 *** | .444 *** | .23 | .363 *** | .264 *** | .82 |

Panel B. Correlations between STDChat (z) and NIOF<10 (y) versus NIOF>50 (x)

| | Pearson Correlations | | | Spearman Correlations | | |
|---|---|---|---|---|---|---|
| daynumber | ρ(y,z) | ρ(x,z) | Hotelling T | ρ(y,z) | ρ(x,z) | Hotelling T |
| -10 | .171 *** | -.079 ** | 4.67 *** | .122 *** | -.065 * | 3.42 *** |
| -9 | .327 *** | -.070 * | 7.28 *** | .079 ** | .030 | .90 |
| -8 | .273 *** | -.112 *** | 6.45 *** | .112 *** | .008 | 1.97 ** |
| -7 | .276 *** | -.153 *** | 7.21 *** | .098 ** | -.025 | 2.34 ** |
| -6 | .327 *** | -.105 *** | 7.85 *** | .081 ** | -.041 | 2.34 ** |
| -5 | .047 | .065 * | -.35 | .089 ** | -.063 * | 2.94 *** |
| -4 | .099 *** | .005 | 1.75 * | .065 * | -.035 | 1.85 ** |
| -3 | .159 *** | .008 | 2.83 *** | .110 *** | .004 | 2.00 ** |
| -2 | .169 *** | .098 * | 1.40 | .150 *** | .002 | 2.81 *** |
| -1 | .227 *** | -.121 *** | 5.79 *** | .023 | -.044 | 1.25 |
| 1 | .135 *** | -.108 *** | 3.58 *** | .060 | -.100 *** | 2.86 *** |
| 2 | .189 *** | .026 | 2.91 *** | .103 *** | -.009 | 2.06 ** |
| 3 | .149 *** | .017 | 2.42 ** | .058 | .004 | 1.02 |
| 4 | .087 ** | -.022 | 2.03 ** | .055 | -.029 | 1.60 |
| 5 | .011 | .002 | .17 | -.013 | -.063 * | .96 |
| 6 | .094 ** | -.023 | 2.22 ** | .079 ** | .005 | 1.43 |
| 7 | .131 *** | .019 | 2.03 ** | .044 | .003 | .78 |
| 8 | .130 *** | -.104 *** | 3.72 *** | .003 | -.041 | .86 |
| 9 | .079 ** | -.008 | 1.71 * | -.048 | -.029 | -.37 |
| 10 | .167 ** | -.024 | 1.67 * | .001 | -.160 * | 1.37 |

* indicates significance at the .10 level; ** at the .05 level; *** at the .01 level. The Hotelling T tests the null hypothesis that $\rho(y,z) = \rho(x,z)$. Hotelling's (1940) test statistic is:

T-ratio = $(\rho_{yz} - \rho_{xz}) / [(n-3)(1 + \rho_{xy}) / 2 (1 - \rho_{xy}^2 - \rho_{xz}^2 - \rho_{yz}^2 + 2 \rho_{xy} \rho_{xz} \rho_{yz})]^{1/2}$.

Table 8
OLS regressions of stock returns on the proportion of retail versus institutional trading volume

This table presents the results of OLS regressions of the pre- and post-announcement returns on the relative amount of (retail / institutional) trading volume around the announcement. The 4 day (8-day) pre-announcement return in Panel A is the log of the ratio of the closing price on the day before the earnings announcement to the closing price on day -5 (day -9) minus the return on the Nasdaq index, measured in the same way as the unadjusted return. The 4-day (8-day) post-announcement return is log of the ratio of the closing price on day +4 (day +8) to the closing price on day -1 minus the return on the Nasdaq index, measured in the same way as the unadjusted return. The earnings surprise is defined as the difference between the most recent analyst forecast and actual quarterly earnings, scaled by the closing stock price on day -10. The proportion of retail trading relative to institutional trading for a given firm is the median across all trading days in the sample of the daily total volume of all trades less than $10,000 divided by the daily total volume of all trades with a value larger than $50,000.

Panel A. Return and Proportion of Retail Trading Before Earnings Announcements

| Model | | Intercept | Proportion Retail / Inst. | Surprise | Adj $R^2$ | Overall F |
|---|---|---|---|---|---|---|
| (5) | 4-day | .022 | .005 | .724 | .011 | 4.58 ** |
| | | 5.27 *** | 1.42 | 3.01 *** | | |
| (5) | 8-day | .019 | .009 | .723 | .005 | 2.76 * |
| | | 3.17 *** | 1.79 * | 2.08 ** | | |

Panel B. Return and Proportion of Retail Trading After Earnings Announcements

| Model | | Intercept | Proportion Retail / Inst. | Surprise | Adj $R^2$ | Overall F |
|---|---|---|---|---|---|---|
| (6) | 4-day | -.021 | -.011 | -.135 | .003 | 2.06 |
| | | -3.30 *** | -1.99 ** | -.36 | | |
| (6) | 8-day | -.020 | -.013 | -.645 | .003 | 2.03 |
| | | -2.48 ** | -1.86 * | -1.40 | | |

* indicates significance at the .10 level; ** at the .05 level; *** at the .01 level.
All regressions are estimated over the sample of N = 664 earnings announcements for which complete data on all regression variables are available.

Table 9
OLS regressions of stock returns on change in chat applied to subsamples
with different proportions of retail versus institutional trading volume

This table presents results of OLS regressions of pre- and post-announcement returns on change in chat. In Panel A (Panel B) the stock return over the 4-day or 8-day window before (after) the earnings announcement is regressed on the change in chat over the same period before (after) the announcement, along with intercept and slope dummy variables. The dummy variable, Small, takes a value of 1 for the one third of observations with the highest proportion of retail versus institutional trading volume. The proportion of retail trading relative to institutional trading for a given stock is the median across all tradir days in the sample of the daily total volume of all trades less than $10,000 divided by the daily total volume of all trades with a value larger than $50,000. This model is applied to the abridged sample excluding the middle one third of observations stratified by the proportion of retail versus institutional trading volume.

Panel A. Return, Chat Activity, and Proportion of Retail Trading Before Earnings Announcements

| Model | | Intercept | Small | ΔChat Before | ΔChatBefore *Small | Surprise | Adj $R^2$ | Overall F |
|---|---|---|---|---|---|---|---|---|
| (7) | 4-day | .022 | -.003 | .000 | .013 | .503 | .039 | 5.49 *** |
| | | 3.31 *** | -.34 | -.27 | 3.11 *** | 2.19 ** | | |
| (7) | 8-day | .020 | -.002 | .000 | .013 | .296 | .016 | 2.75 ** |
| | | 2.41 ** | -.15 | .04 | 2.44 ** | .95 | | |

Panel B. Return, Chat Activity, and Proportion of Retail Trading After Earnings Announcements

| Model | | Intercept | Small | ΔChat After | ΔChatAfter *Small | Surprise | Adj $R^2$ | Overall F |
|---|---|---|---|---|---|---|---|---|
| (8) | 4-day | -.015 | -.028 | -.002 | .018 | -.029 | .033 | 4.67 *** |
| | | 1.83 * | -2.34 ** | -.47 | 3.02 *** | -.93 | | |
| (8) | 8-day | -.015 | -.025 | .000 | .021 | -.394 | .020 | 3.20 ** |
| | | -1.33 | -1.54 | .06 | 2.46 ** | -.96 | | |

* indicates significance at the .10 level; ** at the .05 level; *** at the .01 level.
All regressions are estimated over the abridged sample of N = 439 earnings announcements, excluding the middle one third of the total sample of 664 announcements with medium amounts of (retail / institutional) trading volume.