Small Trader Reactions to Consecutive Earnings Surprises

Devin Shanthikumar[+]
Harvard Business School

First Draft: October, 2002
This Version: June 4, 2006

Abstract

This paper documents small and large trader preferences for earnings growth consistency by examining reactions to the time-series of earnings surprises, focusing on patterns of positive or negative earnings surprises. The results show that small traders react more strongly as a string of earnings increases continues while large traders react similarly throughout the string. Focusing on series of strong positive and strong negative earnings surprises, small traders react more strongly to earnings information if a surprise occurs later in a series of similar surprises, with significant increases in reaction strength between the first, second, and third surprises, while large traders react similarly to each surprise in the series. The results are robust to a variety of controls, including controls for prior returns and the use of alternate earnings expectations models such as a random walk with or without drift and expectations based on analyst forecasts.

[+] Harvard Business School, Morgan Hall 377, Boston, MA 02163, phone: (617)495-6856, fax: (617)496-7363, email: dshanthikumar@hbs.edu. This paper is based on a portion of my thesis at Stanford University. I would especially like to thank my principal advisor, Ming Huang, for all of his advice throughout the process of writing this paper. I would also like to thank Darrel Duffie and Lukasz Pomorski for helpful discussion, and an anonymous reviewer, John Griffin, and seminar participants at the Harvard Business School, Northwestern, Stanford University, the University of California at Irvine, the University of Illinois at Urbana-Champaign, the University of Texas at Austin, the 2004 Financial Management Association Annual Meeting, the 2005 American Accounting Association Annual Meeting and the 2003 London Business School Doctoral Conference for many helpful suggestions.

1. Introduction

Traditional models suggest that market prices of a firm's securities will be formed based on discounted future cash flows, and that current earnings are signals of those future cash flows. Earnings surprises affect expectations for the future, but the specific price effects depend on the earnings expectation models that investors use. A large body of empirical papers documents the significant effect of earnings surprises on security prices (Kothari, 2001). While reactions to individual earnings surprises have been widely explored, reactions to the time series pattern of earnings have been largely neglected. Yet investors use information in both new and prior earnings surprises to form their expectations for the future. Barth, Elliott and Finn (1999) show that firms with patterns of increasing earnings have higher return reactions to the particular earnings surprise than other firms, and have higher price-earnings multiples than other firms. Lakonishok, Shleifer and Vishny (1994) provide evidence that investors over-extrapolate past growth in firm's earnings, sales and cash flows and argue that returns to value strategies are due largely to this tendency of the typical investor to over-extrapolate, rather than risk. Finally, Kasnik and McNichols (2002) find that there is a market premium for firms which consistently meet analyst earnings expectations; although they do not explicitly test whether the premium disappears once the string of meeting expectations is broken. The price and returns evidence clearly points to the importance of earnings patterns in investors' earnings expectation models.

While the price evidence points to an interesting pattern, it is not clear why this pricing effect occurs. Barth, Elliott and Finn (1999) comment that they cannot identify the cause of the price/returns pattern. They suggest several possible explanations. Earnings patterns may capture dimensions of growth or risk that their proxies for these variables do not capture, or there may be an element of mispricing due to "earnings momentum" trading. One particularly important aspect of this price and returns pattern is the following: It is not clear whether investors are underreacting, overreacting or reacting correctly to the pattern of similar earnings surprises. It is even more difficult to use future returns to test whether the reaction is correct or not in this context than with post-earnings-announcement drift, as the logical risk explanations and mispricing would both predict the same future returns. In addition, firm risk (or perceived risk) would supposedly be changing on a quarterly basis rather than at a single event.

Myers, Myers and Skinner (2005) follow up on the Barth, Elliott and Finn study. They attempt to identify whether the type of price multiples and returns documented in Barth Elliott and Finn (1999) are due to risk or mispricing by comparing returns with a matched sample. They compare the returns to firms with 20 quarters of consecutive non-decreases in quarterly seasonally-adjusted earnings per share with a sample of firms matched on size and performance over the same period. They find that the firms with consistent earnings non-decreases earn returns that are significantly higher than the matched firms, and this premium disappears rapidly once the string ends. While this suggests that the firms with consistent performance are overvalued, the creation of a matched sample is by nature imperfect.

A complementary approach is to directly examine the trading reactions of differentially sophisticated investors. Rather than comparing firms with a control sample, we can compare different investors' reactions to the same firms. There are two primary reasons to expect different behavior from small and large traders; information access and investor sophistication. Investors making larger trades will invest more in information acquisition and information processing, as the fixed costs are employed to earn a larger dollar return and an investor with higher quality information will tend to make larger trades based on that information. Large trades are also more likely to be made by institutional investors and professional investment advisors, whose training and resources will allow them to make more sophisticated trading decisions. Experimental researchers have tested for differences between nonprofessional and professional judgments in several accounting and finance contexts, often finding that professionals are more successful in avoiding behavioral biases. Finally, a growing body of empirical work documents differences between small and large trader behavior in several settings.

If the return differences are due to risk or growth factors, there is no reason to expect differences between small and large traders. If prices do not adjust immediately to this information, we might expect larger traders who are more sophisticated and more informed to react more strongly. If, on the other hand, less sophisticated investors are reacting most strongly to these patterns, it is more likely that the response is in fact an overreaction due to investor preference for consistency. Examining the differences between small and large traders in this setting makes two contributions – first, providing greater insight into the information content of a series of similar earnings surprises and second, showing how investors of each type react to the time-series of earnings information.

To measure investor reactions I use New York Stock Exchange Trades and Quotations (TAQ) data. I split trades based on the dollar value of the trade, and calculate measures of trade-initiation which capture which side of a trade demands more immediacy. This measure captures the direction (buy vs. sell) of the reaction[1]. I examine two types of earnings patterns. While Myers, Myers and Skinner (2005) look at cases with 20 quarters of consecutive non-decreases in earnings, only 746 firms exhibit this pattern over the entire period between 1963 and 2004 for the entire market. I focus on the higher frequency patterns of 1-2 years of consistent quarterly performance, to capture investor reactions in a more typical setting. To begin, I look at each trader group's level of reaction to "strings" of consecutive increases in seasonally-adjusted quarterly earnings. To incorporate earnings expectations and eliminate very common, and often expected, small "surprises" such as one cent earnings increases, I focus on sequences of "strong positive" and "strong negative" surprises. A strong positive surprise is defined to be one that is in the top 30% of earnings surprises and a strong negative surprise is one that is in the bottom 30%. In addition, to capture the reaction to earnings information irrespective of whether earnings surprises tend to be more or less extreme as a sequence continues, I focus on a slope variable, which captures the investors' sensitivity to earnings information.

I find that small traders react significantly more positively as a string of earnings increases continues, looking at almost two years of quarterly increases, while large traders do not exhibit a pattern in their trading. Small traders react more negatively during the first year of quarterly decreases, but the pattern breaks down after that, possibly due to the rarity of negative strings. Again, large traders exhibit no clear pattern. Looking at sequences of quarterly earnings surprises, of up to one year in length, small traders react more strongly as a sequence continues while large traders react roughly the same to each surprise in the sequence. This pattern, the focus of this study, is robust to a variety of controls, such as controls for past returns and the use of alternate earnings expectations models. These results suggest, first, that small traders have a preference for consistent performance, while large traders do not value consistency itself. Second, the results suggest that the return patterns documented in Barth, Elliott and Finn (1999) are more likely to be capturing an overvaluation than risk and growth based differences in returns.

The paper proceeds as follows. Section 2 describes the hypothesis development and additional related literature. Section 3 explains the empirical methods and data used in this paper, including the data and methods used to measure earnings surprises and trade reactions. Section 4

[1] Lee and Ready (1992) develops the method and Lee and Radhakrishna (2000) and Odders-White (2000) test the accuracy of the buy/sell classification and the separation of individual and institutional investors.

presents the main trade-reaction results. Section 5 presents additional robustness checks, including controls for prior returns and alternate earnings expectations. Section 6 concludes.

2. Hypothesis development

This paper aims to fill a gap in the literature created by the lack of evidence on investor reactions to the time-series of earnings surprises. It is important to note that, unlike with post-earnings-announcement drift, it is not clear whether the market is reacting correctly or incorrectly to the series of earnings surprises. If a series of positive earnings surprises indicates lower risk, we would predict that the securities with longer strings would exhibit lower future returns, the same pattern we would expect if there is mispricing due to investor overreaction. This makes it necessary to use an alternate method, such as the one used in this paper, to analyze investor responses.

Although strings of earnings increases are most relevant to the Barth, Elliott and Finn (1999) paper, strings of decreases are also relevant to the more general question of how investors use the time-series of earnings to form their expectations. In particular, if a series of consistent earnings increases suggests to investors that a firm has lower risk or better growth prospects, a series of consistent earnings decreases may suggest higher risk and worse growth prospects. Similarly, if investors overreact to consistent increases, they may overreact to consistent decreases as well. Because of this, I will include both positive and negative earnings surprises in the analysis throughout. In addition, I will focus on investor reactions to true earnings surprises rather than to increases and decreases in seasonally-adjusted earnings.

2.1 Small/Large Traders

There are two theoretical reasons to expect different behavior from small and large traders; information access and investor sophistication. In addition, a large body of empirical work has documented differences in small and large trader behavior. This diverse set of evidence all suggests that large traders will make superior trading decisions to small traders.

Larger trades are likely to incorporate more information than smaller trades for two primary reasons. Investors making larger trades will invest more in information acquisition and information processing, as the fixed costs are employed to earn a larger dollar return and an investor with higher quality information will tend to make larger trades based on that information. Cready (1988) and Easley and O'Hara (1987) develop theoretical models predicting that larger trades will incorporate more information and Hasbrouck (1988, 1991) provides empirical evidence that larger trades contain more information.

Large trades are also more likely to be made by institutional investors and professional investment advisors, whose training and resources will allow them to make more sophisticated trading decisions. Tools available to the expert and expertise in the task may increase the expert's ability to avoid biases. Researchers have tested for differences between nonprofessional and professional judgments in several accounting and finance contexts, in addition to many non-market settings, often finding that professionals are more successful in avoiding behavioral biases. In a recent study, Frederickson and Miller (2004) specifically test for differences in analysts' and nonprofessional investors' firm valuations, when given pro-forma earnings disclosures. They find that nonprofessionals are much more strongly influenced by the pro-forma earnings number. Through debriefing questions, they find that roughly half of the nonprofessionals rely on simple heuristics to value the firms, while the professional analysts rely on well-defined valuation models. Smith and Kida (1991) survey the heuristics and biases literature, focusing on auditor judgments. They cite several papers showing that, although certain biases do persist in professional auditors, biases are not as strong with auditor subjects as with student subjects.

Prior empirical studies on small and large trader reactions have shown that there are significant differences between the two groups of investors in several specific settings, consistently finding that small traders make less sophisticated trading decisions. Bhattacharya (2001), Battalio and Mendenhall (2005) and Lee (1992) find differences in small and large trading surrounding earnings announcements, while Hirshleifer, Myers, Myers and Teoh (2003) look at individual brokerage account trading. All four of these papers find that small traders buy for earnings announcements, even if the earnings surprise is negative. Bhattacharya (2001) and Battalio and Mendenhall (2005) find that small traders rely more strongly on a seasonal random-walk model to form their earnings expectations than on analyst expectations, even though analyst expectations will impound much more information than the time-series model. Bhattacharya, Black, Christensen and Mergenthaler (2006) and Bhattacharya, Black, Christensen and Allee (2005) find differences between small and large trading surrounding pro-forma earnings announcements. Malmendier and Shanthikumar (2006a and 2006b) find that investors react more naively to analyst buy/sell recommendations, earnings forecasts and long-term-growth forecasts, taking buy/sell recommendations literally and reacting to the occurrence of forecasts rather than their information content.

In all of these settings, the findings are consistent with small traders displaying less sophistication and using less information in their trading decisions. However, none of these prior studies examine investor reactions to the time-series of information. This is an important distinction as this paper examines investor reactions to a pattern of information, rather than a single new information arrival as in the previous work.

2.2. Earnings Strings/Series

To begin, I examine strings of quarters with increases (or decreases) in seasonally-adjusted earnings, a pattern similar to the one examined in Barth, Elliott and Finn (1999) and Myers, Myers and Skinner (2005). This gives us a first look at small and large trader reactions to the time-series of earnings information. Ke, Huddart and Petroni (2003) and Ke and Petroni (2004) exploit breaks in these types of earnings strings to test the use and dissemination of insider information among insiders and active institutional investors prior to the break, but neither examine trading during the formation of the string. I do not make a specific prediction for the comparison of small and large trader reactions. Instead I test the null hypothesis:

> **N1:** Small and large traders will react similarly to a sequence of increases or a sequence of decreases in seasonally-adjusted quarterly earnings.

Related to the returns patterns documented in the prior work, if the price and return patterns documented in these papers are due to risk characteristics and firm growth prospects, there is no reason to expect a difference between small and large trader reactions, particularly if prices adjust immediately. If prices do not adjust immediately, we would expect the better informed and more sophisticated large traders to react more strongly to a string, earning abnormal returns until prices reach equilibrium levels. If, on the other hand, the premium is due to mispricing, we might expect the less informed and less sophisticated small traders to have a greater preference for these firms than the more sophisticated large traders.

An investor preference for a string of positive earnings increases may be due either to investors' failure to understand the nature of a "true surprise," i.e. investors' reacting too strongly to small positive surprises, or to a more general preference for sequences of positive performance. For example, if we find that small traders react more strongly to a string than large traders, it could be that both groups prefer consistent performance, but small traders view a one cent earnings increase as a surprise and large traders do not. Similarly, it may be that on average the two groups have similar reactions, but one group is more sensitive the information content of the earnings surprise (how large is it?) than the other. To address this issue I focus on a subset of earnings surprises that are likely to be viewed as true earnings surprises by both investor groups. In particular, I look at reactions to a series of strongly positive (top 30% of surprises) or strongly negative (bottom 30%) earnings surprises, using both a random-walk earnings expectations model and analyst-based expectations. I focus on the "slope" of the reaction – i.e. the strength of the trader reaction to the information content of the earnings surprise – rather than to the level of the reaction. The predictions are similar: if the price premium on consistent earnings performers is due to investor overreaction, small traders will react more strongly as a series of strongly positive surprises continues, as compared to large traders. If the premium is due to risk or growth differences, we should find either a similar reaction or a stronger increasing reaction pattern from large traders. This leads to a second null hypothesis:

> **N2:** Small and large traders will react similarly to a sequence of strong earnings surprises.

This second approach, focusing on strongly positive and negative surprises and reaction slope, has two advantages over the first, and so the paper will focus on N2. First, the second approach eliminates the alternate explanation that small and large traders react similarly to the time-series of surprises, but interpret different earnings numbers as surprises. Second, the approach allows us to capture sensitivity to earnings information as a series continues, rather than a level of reaction. Because of this, reactions will not seem stronger (weaker) simply because earnings surprises at a given point in a series tend to be more (less) extreme, or because there is a different level of "attention effect" as the series continues, causing purchasing that is unrelated to the information content of the earnings announcement[2].

3. Research Design and Data

The sample is restricted to securities trading on the New York Stock Exchange between January 1, 1993 and December 31, 2002. Only ordinary common shares, excluding certificates and depository receipts, are included. Foreign companies, Americus trust components, closed-end fund shares and Real Estate Investment Trusts are also excluded, corresponding to including securities with a Center for Research in Security Prices (CRSP) share code of 10 or 11. The final sample includes 2,801 securities for 2,723 firms. Returns data are obtained from CRSP. Earnings announcements and firm characteristics are taken from Compustat, and analyst earnings forecasts are taken from the Institutional Brokers Estimates System (IBES). Trading measures are calculated from the New York Stock Exchange Trades and Quotations (TAQ) database. Data on individual accounts with a large discount brokerage firm are used to evaluate the TAQ-based data.

3.1. Measuring earnings surprise events

This paper is based on the event-study methodology, and thus the most basic data needed are the event dates. Quarterly earnings announcement dates are taken from Compustat. Hughes and Ricks (1987) report that, with their sample of 677 annual earnings announcements for the years 1979-1981, the Compustat date is accurate 75% of the time. They do not describe the errors for the other 25% of the cases. To ensure that the Compustat report date will give accurate event dates for our sample, I randomly select 125 stock-quarter combinations between 1993 and 2001. I compare the Compustat earnings announcement date with the first report date found in

[2] See Barber and Odean (2003) for a more general study of attention effects. The purchasing behavior of small traders in response to negative surprises documented in Lee (1992) is one possible example of an attention effect, and suggests that controlling for this effect is important when examining investor reactions to earnings announcements.

the Dow Jones News Service (DJNS) and Wall Street Journal (WSJ). In each case, the Dow Jones News Service reported before the Wall Street Journal. In the random sub-sample, there are eight earnings announcements that do not appear in the DJNS Index or the WSJ. Four of these lack earnings announcement dates in Compustat. There are six announcements lacking dates in Compustat, and which are available in the DJNS Index. Among the remaining 117 announcements, the dates are the same in 79.5% of the cases. In the cases where the dates differ the Compustat date is usually one day after the DJNS report date, corresponding with the Wall Street Journal report date. There are only 3 cases, 2.6% of the sample, in which the dates differ by more than 1 day, and only 1 case where they differ by more than 2 business days. Using the Compustat date allows for a sample of tens of thousands of announcements, while hand-collected dates would limit the sample to a few hundred announcements. The minimal cost from inaccuracies is certainly worthwhile.

If an announcement is made on a holiday or weekend, the first trading day after the announcement date is used as the event date. The sample is limited to earnings announcements falling in the years 1993 through 2002 as TAQ data is only available beginning in 1993. The final sample of earnings surprises for which I have all necessary data contains 59,658 earnings announcements, for the main earnings surprise measure.

There are three primary alternatives for measuring earnings expectations and earnings surprises that are used in the literature. The typical measures are based on prior earnings, analyst forecasts of earnings, or stock returns around the announcement date. Alternatively, rather than looking at earnings "surprises", some prior work focuses on earnings increases and decreases. As a robustness check, I use increases and decreases in Section 5. The primary measure for this paper is earnings surprise based on expectations built from prior earnings announcements. I use the standardized unexpected earnings measure commonly used in the post-earnings-announcement drift literature. In order to calculate standardized unexpected earnings (SUE), I assume that earnings expectations are based on a seasonal random walk model. The primary measure uses a seasonal random walk with drift, but I also repeat the study using a seasonal random walk without drift, and the results are robust to this variation. Expected earnings are described by

$$E\left(e_t^i\right) = e_{t-4}^i + \delta^i \quad , \tag{1}$$

where δ^i is the earnings drift for firm *i*. For each stock, drift is estimated using up to twenty quarters of previous data, as

$$\hat{\delta}^i = \frac{1}{n}\sum_{j=1}^{n}\left(e^i_{t-j} - e^i_{t-j-4}\right) , \qquad (2)$$

where n ≤ 16. Less data is used if the full period is not available, although at least one year's worth of data is required. This introduces a slight survivorship bias into the sample, but it eliminates only 5.34% percent of the earnings announcements and 2.97% of firms. I then standardize the unexpected earnings measure by dividing each firm surprise by the standard deviation of that firm's earnings, as measured by the available subset of the preceding 20 announcements. As a robustness check, I also normalize the unexpected earnings measure by the standard deviation of earnings changes rather than the standard deviation of earnings. As reported in Section 5, results are similar. The primary measure is as follows:

$$SUE^i_t = \frac{e^i_t - e^i_{t-4} - \hat{\delta}^i}{\sqrt{Var(e^i_t)}} , \qquad (3)$$

where $Var(e_t)$ is estimated using the previous twenty announcements. Earnings announcements are then ranked by SUE within each year, and placed into deciles 0-9, where the most negative surprises are in decile 0 and the most positive in decile 9. Earnings announcements in deciles 4 and 5 are not strong surprises.

The second measure of earnings surprises is based on analyst forecasts. The surprise is the difference between announced earnings-per-share and the analyst forecast, normalized by stock price. The IBES analyst forecast database provides a sample size in the tens of thousands, but the limits on analyst coverage still reduce the sample size by almost one half. The analyst-based surprise measure is used as a robustness check. I use two measures of the consensus forecast. The first measure is based on the IBES summary file. The consensus forecast is defined as the most recent monthly median forecast before the earnings announcement, when there are at least four earnings forecasts for the firm. The consensus forecasts occur a mean of 15.8 days and median of 13 days before the earnings announcement dates, so they tend to be approximately two weeks old. The mean and median forecasts tend to be very similar, with a correlation coefficient of 0.998. The second measure of consensus is constructed from the IBES detail files, and uses the median forecast occurring at least one week before the earnings announcement and no more than two months before the announcement, when there are at least four earnings forecasts during that period. With both consensus measures, the earnings surprise is defined as the difference between announced earnings-per-share and the consensus, divided by the price on the date of the consensus forecast. The sample contains 29,649 earnings announcements using the monthly consensus measure and 14,505 earnings announcements using the daily measure. Results using these measures are reported in Section 5.

The third measure of earnings surprise used in the literature is the stock return on the days surrounding the earnings announcement. Rather than measuring the difference between expected and realized earnings, this method measures the reaction to this difference, making it unsuitable for this study in which we are interested in the reactions to surprises. Although the return-based measure is not ideal for the analysis in this paper, I repeat the primary tests using this surprise measure. I define the surprise to be the cumulative abnormal return over trading days –1 through 1 in event-time. This yields mixed results, with some evidence that small traders exhibit an increasing reaction, and some evidence that the trading reactions are virtually identical to the surprises regardless of prior surprise history. These results do not speak directly to the behavioral question at hand.

Based on the earnings surprise measure, I assign an earnings surprise a value of N=0 if it is a mild surprise, in deciles 3, 4, 5 or 6. The surprise is assigned a value of N=1 if it is a strongly negative surprise (deciles 0, 1 or 2) and the preceding surprise for that firm was not strongly negative, or if it is strongly positive (deciles 7,8 or 9) and the preceding surprise was not strongly positive. The surprise has a value of N=2 if it is the second surprise of the same type, strongly negative or strongly positive, N=3 if it is the third and so on.

Table 1 reports sample statistics for the firms with complete data. The table displays firm size, book-to-market ratio, and standardized unexpected earnings statistics for the whole sample and divided by earnings surprise N value. One possible concern with the sample is that if the earnings surprise measure induces positive autocorrelation in quarterly earnings surprises, then an N=2 surprise is not really a surprise. It is possible that almost all the N=2 surprises are followed by an N=3 surprise, for example. The table shows that only about one half of the N=1 surprises are followed by an N=2 surprise. Similarly, about one half of the N=2 surprises are followed by an N=3 surprise, so that each surprise in the series is still a surprise. The magnitudes of the standardized unexpected earnings variables do not seem to vary systematically as the series continues. In addition, the use of several alternate earnings surprise measures mitigates this concern. In particular, analyst expectations should take into account any changes in expectations due to the prior series. Finally, one pattern that emerges from Table 1 is an apparent decrease in firm size with increasing N, but as the high standard deviation suggests, these decreases are not significant. Regression results show that the t-statistics on decreases in size are consistently less than 1. For example, while the magnitude of the firm-size decrease from N=1 to N=7 is $812 million, the standard error on that decrease is $1.16 billion.

3.2. Measuring Trade-Initiation

The raw trading data is collected from the New York Stock Exchange Trades and Quotations database (TAQ). This database reports every round-lot trade and every quote from 1993 onwards on the New York Stock Exchange, American Stock Exchange and NASDAQ. In order to measure trading reactions, I first classify NYSE trades as buyer or seller initiated, and then aggregate on a daily basis for each stock. I use the modified Lee and Ready (1991) algorithm, recommended in Odders-White (2000), to determine which side initiated a given trade: the buy side or sell side. This algorithm is commonly used in the empirical market microstructure literature (see Odders-White 2000 for a list of papers using the Lee-Ready algorithm). The algorithm involves matching a trade to the most recent quote preceding the trade by at least 5 seconds. If a price is nearer the bid price it is classified as seller initiated and if it is closer to the ask price it is classified as buyer initiated. If a trade is at the midpoint of the bid-ask spread, it is classified based on the previous price using a "tick test"– if the trade occurs at a price that is higher (lower) than the price of the previous trade it is classified as buyer (seller) initiated. The original Lee-Ready algorithm also employs a "zero-tick" in the case that a trade is at the bid-ask midpoint and the same price as the previous trade, but this aspect of the algorithm is the most problematic. Odders-White (2000) reports that the "zero-tick" is only about 60% accurate in general. Because of this, I use the modified Lee-Ready algorithm without the "zero-tick" test, and the few trades falling into this category remain unclassified, with a total of approximately 10% of trades remaining unclassified. Using the original Lee-Ready algorithm for a large subsample of the data yields similar final results as the modified algorithm.

To separate small and large trades I use a set of four cutoffs: $5,000, $10,000, $20,000 and $50,000. The five size-based categories are used to test whether there is a clear division between small and large trader behavior or a gradual change. Since trades less than $5,000 may not have the market power to explain post-earnings-announcement-drift, it is important to test whether the results hold using a higher small trade cutoff. I then aggregate the trade-by-trade data to find daily trading measures for each stock. The final dataset contains data from over 640 million trades over ten years, with each category containing a similar number of trades. 20.8% of the classified trades occur below $5,000, 16.4% between $5,000 and $10,000, 16.8% between $10,000 and $20,000, 20.4% between $20,000 and $50,000 and 25.6% above $50,000.

The analysis in the paper will focus on "small" trades of less than $5,000 or less than $20,000 and "large" trades of at least $50,000. The choice of primary cutoffs in this paper is strongly based on the evidence of Lee and Radhakrishna (2000), who analyze these sorts of methods. There are also several other papers which use the NYSE TAQ data to judge reactions to information, with similar small/large cutoffs.[3] Lee and Radhakrishna (2000) show that dollar based cutoffs create less noise in separating individuals from institutions than share-based cutoffs. They also find that, for their three-month sample from 1990-1991, a very low cutoff such as $5,000 or less is most effective in separating out individuals. Similarly a high cutoff of $50,000, or even $100,000, is most effective in separating out institutions. In more recent work, Griffin, Harris and Topaloglu (2003) show, with their more recent May 2000 through February 2001 sample of NASDAQ firms, that there is still an extremely strong relationship between trade size and trader identity. While this paper's aim is not specifically to discriminate between individuals and institutions, this interpretation is useful.

3.3. Calculating abnormal trading measures

In order to aggregate across firms, and to be able to make clearer conclusions regarding the comparison of event-time trading and non-event time trading, I calculate abnormal trading measures. The primary variable of interest is a measure of trade imbalance. Intuitively, if every trade after an announcement were being initiated by the buy side, then the trading reaction to that announcement is extremely positive. Similarly, if all trades were being initiated by the sell side,

---

[3] Other existing papers use the NYSE TAQ data to examine small and large trader behavior in a similar way, judging reactions to information. Lee (1992) uses a single cutoff of $10,000, with $5,000 and $20,000 as robustness checks. Hvidkjaer (2001), in his dissertation, uses dollar cutoffs of $20,000 and $50,000 with a buffer zone in between. He uses roughly corresponding share-based cutoffs as well. Battacharya (2001) uses dollar cutoffs of $5,000 and $50,000. The buy/sell classification is suited to determining reactions to information. One possible confounding effect is the desire to minimize the revelation of private information by using limit orders or splitting trades, but this should be less of a problem in the case of public information such as earnings announcements. In addition, the primary cutoff between medium and large trades is not extremely high, in this study or its predecessors, so that strategizing to minimize market impact should not be a major factor in trade-size category.

then the reaction is strongly negative. To capture this concept, the raw trade imbalance measure is calculated as follows, for firm i, investor type x, and date t:

$$IMB_{i,x,t} = \frac{buys_{i,x,t} - sells_{i,x,t}}{buys_{i,x,t} + sells_{i,x,t}} \quad . \tag{4}$$

I then normalize this trade imbalance measure by subtracting off the non-event-time firm-year sample mean, and dividing by the non-event-time firm-year sample standard deviation, with

$$IMB_{i,x,t}^{abnormal} = \frac{IMB_{i,x,t} - E\left(IMB_{i,x,year(t)}\right)}{\sqrt{Var(IMB_{i,x,year(t)})}} . \tag{5}$$

This controls for systematic differences in trading behavior. The event period that is excluded in calculating $E(IMB_{i,x,year(t)})$ and $Var(IMB_{i,x,year(t)})$ consists of days -5 through 5 in event time; the eleven trading days centered on any earnings announcement date. The exclusion is important to maintain differences in reactions to the earnings announcements as part of our abnormal imbalance measure. The normalization should correct for systematic differences in trading behavior for different stocks, trading groups, and times, but not correct for differences in trading reactions to the earnings surprises.

Adjusting raw trade imbalance in this way allows me to aggregate across firms, years, and trader groups without concerns for differences in the general trading behavior associated with them. Normalizing the measures by the standard deviation allows for comparisons of estimated values that would be impossible to make if the values were not normalized. For example, without the normalization a seemingly more extreme small trader reaction could be the result of higher small trade imbalance volatility overall. Dividing by the standard deviation controls for systematic differences in the volatility of large trades and small trades or in the volatility of the stocks in which large and small traders invest.

In order to evaluate the abnormal trade imbalance measures calculated from TAQ-based trade initiation, I compare these values with trading data from 50,000 individual accounts at a large discount brokerage firm, for the period of 1991-1996. These data describe the actual trades made by the brokerage account clients, including trades made through both market and limit orders, along with characteristics of these trades and the individuals owning the accounts. The overlapping period and firm sample reduces the brokerage firm trade data to 462,310 trades. Correlations between imbalance measures for these brokerage account trades and the TAQ-based abnormal trade imbalance, both aggregated to the monthly level, are displayed in Table 2. The correlations between small trade imbalance and individual trade imbalance are significantly positive, for number of trades, number of shares and dollar value traded, while the correlations between large trade imbalance and individual trade imbalance are significantly negative. This means that the stronger the large trade "buy pressure" the more individual ownership among these brokerage accounts is decreasing. If we believe individuals in the brokerage firm database are representative of small traders in general, then this simply reflects a shift in ownership from small traders to large traders: The more strongly large traders buy, the more small-trader ownership decreases.

An additional adjustment is made to the abnormal trade imbalance, used as a robustness check. I control for prior returns over varying horizons of up to three months with a return-adjusted abnormal trade imbalance. This control is particularly important in ensuring that lagged trading behavior is due to the information in the earnings surprise, and not a naïve response to the drift in the period prior to the earnings surprise. In calculating returns, I use cumulative abnormal returns (CARs), which are the sums of daily abnormal returns, defined as

$$CAR^{i}_{t_0,t_1} = \sum_{t=t_0}^{t_1} \left(AR^{i}_{t} \right), \tag{6}$$

where AR^i_t is the CRSP beta-adjusted abnormal return for security *i* on day *t.* Fama (1998) summarizes theoretical and statistical reasons CARs are preferable to buy-and-hold returns. In order to calculate the adjusted measure, I estimate the equation

$$IMB^{abnormal}_{i,t} = \alpha_0 + \alpha_1 AR^{i}_{t-1} + \alpha_2 CAR^{i}_{t-5,t-2} + \alpha_3 CAR^{i}_{t-20,t-6} + \alpha_4 CAR^{i}_{t-60,t-21} + \varepsilon_{i,t}, \tag{7}$$

for each trade-size-group's trade-imbalance. This equation essentially groups prior returns into the prior day, week, month and quarter. Return adjusted abnormal trade imbalance is the residual from the above equation, that is, the abnormal trade imbalance that is not accounted for by the previous day, week, month and quarter returns. Details of investor reactions to prior returns and the test results using this return-adjusted measure are reported in Section 5.

An alternate normalization procedure is used as an additional robustness check. Based on the evidence in Chordia, Roll and Subrahmanyam (2002), I perform a normalization controlling for calendar-effects, serial correlation in the trade imbalance variable, and dependence of trade imbalance on prior returns, similarly to Frieder (2004). The first step involves regressing raw trade imbalance on indicators for month (January, February…) and day-of-week (Monday, Tuesday…). This regression is run for each security separately, using the entire sample period. The residual is used in the second step, where the calendar-adjusted imbalance is regressed on the previous fifteen trading-days' calendar-adjusted imbalance and security return. Again, these regressions are run for each stock separately. The residual from these regressions is used in the final step. In order to ensure that the values are comparable across trade size groups, in the third step, the mean for that security and trade-size group is subtracted and the mean-adjusted imbalance is divided by the standard deviation for that security and trade-size group. This final step is similar to the primary normalization, removing the mean and standard deviation effects for each firm and trade-size group. Results using this normalization are similar and are reported in Section 5. Details of the procedure and insights for calendar-effects in returns are described in Shanthikumar (2004a).

4. Trade Reaction Results

In this section I present results for the trading behavior of the different trade-size groups, beginning with earnings increase/decrease strings and then focusing on sequences of strong positive and strong negative surprises.

An earnings increase (decrease) occurs if earnings are higher (lower) than in the same quarter of the prior year. Let INC_1 indicate an earnings value which is an increase in seasonally-adjusted earnings following a quarter with either a decrease or no change. INC_2 indicates an increase following an INC_1 increase, indicating that it is the second quarter in a row with an increase in seasonally-adjusted earnings. INC_3 indicates an increase following an INC_2 increase, and so on. $INC_{\geq 7}$ indicates earnings increases which are the seventh, eight, ninth or higher in a string of increases. DEC_X is defined in the same way, for earnings decreases. To measure investor reactions I estimate the following regression:

$$IMB_{s,t}^{abnormal} = \alpha_1^t INC_1 + \alpha_2^t INC_2 + ... + \alpha_6^t INC_6 + \alpha_7^t INC_{\geq 7} +$$

$$\beta_1^t DEC_1 + \beta_2^t DEC_2 + ... + \beta_6^t DEC_6 + \beta_7^t DEC_{\geq 7} + \varepsilon_{s,t}\,, \qquad (8)$$

where t is the trading day in event-time and s is the specific earnings surprise. In untabulated regressions I estimate equation 8 for a range of individual trading days surrounding the earnings announcement date. To control for potential differences in the timing of small and large trader reactions, I use the sum of abnormal trade imbalance over days -5 through 5 in event time as the dependant variable for much of the analysis. Results are similar using alternate event windows. In order to determine difference t-statistics across trader groups, the regression is modified in various ways.

Reactions to a string of such increases or decreases are reported in Table 3. For a string of increasing earnings, there is a significant increasing reaction on the part of small traders, that is, their reactions become more positive as this series of increases continues. The coefficients increase almost monotonically. The increase between the smallest trader reaction to the first positive surprise and the reaction to all surprises from the third on is statistically significant at the 1% level, as is the increase between the second surprise and the fifth, sixth and seventh on. There is no such pattern with large traders. For a string of decreasing earnings, the pattern is not consistent for either group. Small traders react significantly more negatively to a DEC_2 or DEC_3 surprise than to a DEC_1 surprise, and significantly more negatively to a DEC_4 surprise than to a DEC_1, DEC_2 or DEC_3 surprise, but for decreases after that the small trader reaction becomes more positive. The sample sizes are far smaller for earnings decreases than for earnings increases, and the difference is larger later in the string. These reactions seem consistent with the mispricing stories told in Barth, Elliott and Finn (1999) and Myers, Myers and Skinner (2005). In particular, small traders seem to value strings of earnings increases, possibly leading to the managerial focus on achieving these strings.

Turning our attention to sequences of strong positive and strong negative earnings surprises, I estimate the following regression:

$$IMB_{s,t}^{abnormal} = \alpha_1^t \mathrm{I}(N_s = 1) + \alpha_2^t \mathrm{I}(N_s = 2) + \alpha_3^t \mathrm{I}(N_s = 3) + \alpha_4^t \mathrm{I}(N_s \geq 4) +$$

$$\beta_1^t \mathrm{I}(N_s = 1) SurpDec_s + \beta_2^t \mathrm{I}(N_s = 2) SurpDec_s + \beta_3^t \mathrm{I}(N_s = 3) SurpDec_s$$

$$\beta_4^t \mathrm{I}(N_s \geq 4) SurpDec_s + \varepsilon_{s,t} \quad , \tag{9}$$

where t is the trading day in event-time, s is the specific earnings surprise, $SurpDec_s$ is the surprise decile for surprise s and $\mathrm{I}(N_s = X)$ is the indicator that the N value, as defined in section 3 as the position of a surprise in a series, for surprise s has value X. I focus on the first one year of earnings surprise sequence as the sample sizes are much smaller for these sequences than for earnings increase/decrease strings. I again estimate equation 9 for a range of individual trading days surrounding the earnings announcement date, but present results for the sum of abnormal trade imbalance over days -5 through 5. In order to make the behavior clearer over time for the purpose of figures, I use a moving average method for a moving 5-day window. Results are similar using alternate event windows. In order to determine difference t-statistics across trader groups, the regression is modified in various ways. In these regressions, α measures the intercept for the reaction to the earnings announcement, which is roughly the reaction to an extremely negative surprise (decile 0). β reflects the sensitivity of abnormal trade imbalance to the earnings surprise decile, the measure of the information content of the surprise. Essentially, β measures the strength of the reaction. I also estimate a variant on the basic regression which is more similar to equation 8, to examine behavior for positive surprises (deciles 7,8,9) and negative surprises (deciles 0,1,2) directly.

Table 4 displays the coefficient estimates for the sum of abnormal trade imbalance over days -5 through 5, for trades less than \$5,000, between \$5,000 and \$10,000, between \$10,000 and \$20,000, between \$20,000 and \$50,000 and finally for trades of at least \$50,000, as well as differences in the β estimates. Figure 1 displays the results for β for N=1, N=2 and N=3, over a longer period of time surrounding the announcement date, and using the moving average method. Figure 1, Panel A displays the results for the smallest trade-size category, and shows that small trades display the hypothesized trading behavior. Results are similar if we define small trades as those less than \$10,000 or less than \$20,000. The slope is higher for N=2 than for N=1, and higher still for N=3 than for N=2. Figure 1, Panel B displays the results for large trades. Their trading does not depend more on surprise decile for N=2 or N=3 firms than for N=1 firms. Hypothesis testing shows that the increase in β exhibited by the small traders, and displayed in Figure 1, is statistically significant. For trades below the \$5,000 cutoff, N=2 dependence is significantly higher than N=1 at the 10% level for the entire period from day -10 through day 7. The reaction to N=3 is significantly higher than to N=1 for days -10 through 23. The increase from N=2 to N=3 is also significant at the 10% level from days -7 through 1, days 3-5 and day 7.

On days 4 and 6 the difference is significant at an 11% level. Small traders are reacting more strongly to each successive earnings surprise for the first three surprises in the sequence, strongly supporting our hypothesis.

In the discussion below, I will focus on the two extreme groups displayed in Figure 1, but results are similar if alternate cutoffs such as $10,000 or $20,000 are used. Comparing the smallest and largest trades, the scale of coefficients is noticeably different. Small traders are not only increasing their absolute reaction, but are increasing their benchmarked reaction as well, benchmarked against large traders. Small and large traders react similarly to the first surprise in the series, but the small trader reaction more than triples by the third surprise while the large trader reaction remains at roughly the same level. Small traders react significantly more weakly to the first surprise than large traders, except for four days around the event. They react more strongly than large traders to the subsequent surprises, with the differences between the small and large reaction being significant at the 10% level for days 0 through 6 for N=2, days 0 through 10 for N=3 and 0 through 7 for N≥4, along with additional days before and after these periods.

The increasing reaction phenomenon is not driven solely by reactions to positive surprises. Evidence on the intercept term in Table 4, and untabulated results for reactions to positive and negative surprises separately, show that the small trade reaction becomes more extreme for both positive and negative surprise series, and the large trade reaction does not become more extreme for either. Generally, I focus on slope rather than absolute reactions to the two extremes due to the potential confounding factor of attention effects (see Barber and Odean 2003), but results for each extreme show that the small trade reaction to the most negative decile becomes lower as N increases. Similarly the reaction to the most positive decile becomes higher as N increases. The reactions to the intermediate deciles follow a similar pattern, although the change in reactions is not quite as clear for these less extreme surprises. For large traders, there is no clear increasing or decreasing reaction pattern at either end.

While there is a significant difference between the reactions of the very smallest trades and the very largest trades, it is also interesting to examine the trading in the intermediate range. The results displayed in Table 4 show that as trade size grows, the reaction to the first surprise increases in magnitude. Similarly, the reactions to the later surprises seem to drop as trade size grows, although this relationship is not monotonic. There is a gradual progression between the behavior of the very small traders and the very large traders. A stronger reaction to later surprises appears to be more persistent as trade size increases. Most importantly, the increasing reaction pattern is clearly present for trades in the $5,000 to $10,000 and $10,000 to $20,000 ranges, although not all of the increases are significant. The significance of the increasing pattern for trades in the $10,000 to $20,000 range does vary with additional controls and method variations, but the pattern remains consistent, and the increasing reaction for the $5,000 to $10,000 range is robust.

5. Robustness Tests

Controls for Prior Returns, and Alternate Trade-Data Methods. In general, a firm that has had prior positive earnings surprises will also have had prior positive returns. It is possible that a naïve reaction to past returns could be driving the difference in reactions between small and large traders or the differences among small trader reactions to surprises in a sequence. In order to test for this, I control for prior returns as described in Section 3.3. Results from estimating equation 7 are displayed in Table 5. From the regression results in Table 5, large traders appear to be "momentum" traders – buying more strongly when prior returns have been higher – while small traders appear to be "contrarians". These results are consistent with evidence on mutual fund portfolios (Grinblatt, Titman and Wermers 1995) and indirect evidence on individual trading (Barber and Odean 2000). Because of these different responses to prior returns, the results are stronger with the controls for prior returns. Table 6 reports regression results. Small traders have a strongly increasing reaction while large traders have about the same reaction to the consecutive surprises. The regression results confirm that the increases are significant.

Table 6 also reports the regression results using the alternate normalization procedure which accounts for calendar-time effects, past returns for the fifteen prior trading-days, and past trade imbalance for the fifteen prior trading-days. The results using this alternate measure are also significant, and slightly stronger than the primary results.

I perform additional method sensitivity checks on a large subsample of the data. These additional checks include using an alternate buy/sell classification method, a time-varying size cutoff to approximate a percentage cutoff, a firm-size dependent cutoff and normalizing the trade imbalance by mean only. As a final control for any market-wide reaction effects on a given day I include date fixed-effects. The untabulated results from these variations reinforce the main results.

Possible Front-Running. Large traders may be more likely to predict which earnings series will continue and which will not. For example, they may watch the firms with more extreme earnings surprises more closely, thus reacting to new information in a timelier manner before the next earnings announcement. Alternatively they may simply learn of the coming earnings announcement values in advance. To test whether this affects the results, I include the trading results for the months prior to the earnings announcement. In particular, rather than measuring the reaction as the sum of daily abnormal trade imbalance from event-time trading-days -5 through 5, I repeat the regressions using periods such as -40 through 5, -50 through 5, and even -70 through 5, which would incorporate the reaction to the preceding surprise. With all of these variations, the results remain significant: small traders display a significantly increasing reaction while large traders do not display any increase.

Fiscal-Year-Ends. There are two reasons to repeat the analysis excluding fiscal-year-ends. First, annual earnings reports may affect the results. Second, there may be some element of increased investor involvement as a series continues, perhaps because press coverage increases as a string of similar earnings surprises progresses. Such an effect may be stronger at the fiscal-year-end, when press coverage and investor attention are high. A simple and straightforward test of both of these is to discard all quarterly announcements that occur at the end of a firm's fiscal year. Table 6 displays the results, which are extremely similar to the main results, and remain significant.

Data Sub-samples. One set of checks looks at various subsamples of the data, due to the possibility that trading behavior is not consistent over time or across types of firms. The main test is repeated on a set of thirty subsamples: each of the ten years from 1993 through 2002, each of the ten size deciles and each of the ten book-to-market ratio deciles. While the strength of the results varies from group to group, the general results remain, and there are no groups that have contradictory results or that seem to be driving the full-sample results. In addition, I repeat the key regressions including fixed-effects for these characteristics, and Table 6 displays results for security fixed-effects and quarter-year fixed-effects. As the table shows, the increasing reaction among small traders remains significant.

Alternate Earnings Surprise Measures. One potential concern with the results is that I may not be properly capturing earnings expectations, due to the use of a fairly simple model. In untabulated regression results, I repeat tests using four variations of the SUE measure and find similar and significant results whether earnings surprises are calculated with or without accounting for earnings drift and whether the unexpected earnings is normalized by standard deviation of earnings or standard deviation of earnings changes. A direct test for a potential failure to capture earnings expectations is to use analyst forecasts as a proxy for earnings expectations, as analyst forecasts will be a more sophisticated measure. Analysts should account for the history of earnings for the given firm as well as other financial variables, non-financial data, industry-wide data and economy-wide information. Using this measure, the results are slightly different, but are still consistent with the increasing reactions hypotheses. Table 6 displays these results for earnings surprises based on IBES summary data. Results are similar using IBES detail data to construct an earnings forecast consensus measure closer to the date of the earnings announcement. I use two measures of analyst-based surprise, as described in Section 3.1, and results are similar using the two measures. Small traders have a slightly stronger reaction to the N=2 surprise than to the N=1 surprise, but the difference is much smaller than using the SUE measures. Their reaction to the third surprise is higher and their reaction to surprises with $N \geq 4$ is higher still. For large traders the reaction to the second surprise is actually lower than the reaction to the first, while the reactions to the third and fourth surprises are about equal. If large traders are making less biased trading decisions, this suggests that an N=2 analyst-based earnings surprise is not as meaningful as an N=1 surprise, and would account for the similar small trade reaction to both N=1 and N=2.

6. Conclusion

This paper documents investor reactions to the time-series of earnings surprises. Results show that smaller traders react more positively as a string of earnings increases progresses, while large traders exhibit no increase in their response. Focusing on series of strong positive and negative surprises, to eliminate the possible effect of small, possibly expected, "surprises", we find that small traders react more strongly to surprises later in a series, while the largest category of traders does not exhibit increasing reactions. These results are robust to many method variations, including alternate earnings surprise measures and alternate abnormal trading measures. In addition, controlling for prior returns shows that smaller traders tend to trade as contrarians and larger traders as momentum traders, independent of earning surprises. Because of this, controlling for prior returns strengthens the results.

This trading evidence suggests that small traders have a preference for consistent earnings performance, while large traders do not have a preference for consistency. This suggests that the price/returns patterns documented in Barth, Elliott and Finn (1999) may be related more to investor preferences than to risk or growth information communicated by the pattern. The small and large investor reaction to earnings sequences may also be relevant to behavioral finance models which predict that investors will react more strongly as similar information is released (Barberis, Shleifer and Vishney, 1998, and Daniel, Hirshleifer and Subrahmanyam, 1998).

This evidence complements the return evidence in Myers, Myers and Skinner (2005) and provides the first evidence on small and large investor reactions to the time-series of earnings information.

References

Barber, Brad M. and Terrance Odean, 2003. All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors. Working paper, University of California at Davis and University of California at Berkeley.

Barberis, Nicholas, Andrei Shleifer and Robert W. Vishny, 1998. A Model of Investor Sentiment. Journal of Financial Economics 49, 307-343.

Barberis, Nicholas and Richard Thaler, 2003. A Survey of Behavioral Finance. In: Handbook of the Economics of Finance, G.M. Constantinides, M. Harris and R. Stulz ed.

Barth, Mary E., John A. Elliott and Mark W. Finn, 1999. Market Rewards Associated With Patterns of Increasing Earnings. Journal of Accounting Research 37, 387-413.

Battalio, Robert H. and Richard R. Mendenhall, 2003. Earnings Expectations and Investor Clienteles. Working paper, University of Notre Dame.

Bernard, Victor L. and Jacob K. Thomas, 1989, Post-earnings-announcement Drift: Delayed Price Response or Risk Premium?, Journal of Accounting Research 27: 1-36.

Bernard, Victor L. and Jacob K. Thomas, 1990. Evidence That Stock Prices do not Fully Reflect the Implications of Current Earnings for Future Earnings. Journal of Accounting and Economics, 13, 305-340.

Bhattacharya, Nilabhra, 2001. Investors' Trade Size and Trading Responses Around Earnings Announcements: An Empirical Investigation. The Accounting Review, 76, 221-244.

Chan, Louis K. C., Narasimhan Jegadeesh and Josef Lakonishok, 1996. Momentum Strategies. Journal of Finance 51, 1681-1713.

Daniel, Kent, David Hirshleifer and Avanidhar Subrahmanyam, 1998. Investor Psychology and Security Market Under- and Over-reactions. Journal of Finance 53, 1839-1886.

Fama, Eugene F., 1998. Market Efficiency, Long-Term Returns and Behavioral Finance. Journal of Financial Economics 49, 283-306.

Frederickson, James R. and Jeffrey S. Miller, 2004. The Effects of Pro Forma Earnings Disclosures on Analysts' and Nonprofessional Investors' Equity Valuation Judgments. The Accounting Review 79, 667-686.

Frieder, Laura, 2004. Evidence on Behavioral Biases in Trading Activity. Working paper, University of California at Los Angeles.

Griffin, John M., Jeffrey H. Harris and Selim Topaloglu, 2003. The Dynamics of Institutional and Individual Trading. Journal of Finance 58, 2285-2320.

Grinblatt, Mark, Sheridan Titman and Russ Wermers, 1995. Momentum Investment Strategies, Portfolio Performance, and Herding: A Study of Mutual Fund Behavior. American Economic Review 85, 1088-1105.

Hirshleifer, David, James N. Myers, Linda A. Myers and Siew Hong Teoh, 2003. Do Individual Investors Drive Post-Earnings-Announcement Drift? Direct Evidence From Personal Trades. Working paper, Ohio State University and University of Illinois at Urbana-Champaign.

Hvidkjaer, Soeren, 2001. A Trade-based Analysis of Momentum. Working paper, University of Maryland.

Kasznik, Ron and Maureen McNichols, 2002. Does Meeting Earnings Expectations Matter? Evidence from Analyst Forecast Revisions and Share Prices. Journal of Accounting Research 40, 727-759.

Ke, Bin, Steven Huddart and Kathy Petroni, 2003. What Insiders Know About Future Earnings and How They Use It: Evidence From Insider Trades. Journal of Accounting and Economics 35, 315-346.

Ke, Bin and Kathy Petroni, 2004. How Informed are Actively Trading Institutional Investors? Evidence From Their Trading Behavior Before a Break in a String of Consecutive Earnings Increases. Working paper, Pennsylvania State University and Michigan State University.

Kothari, S. P., 2001. Capital Markets Research in Accounting. Journal of Accounting and Economics 31, 105-231.

Lakonishok, Josef, Andrei Shleifer and Robert Vishny, 1994. Contrarian Investment, Extrapolation and Risk. Journal of Finance 49, 1541-1578.

Lee, Charles M. C., 1992. Earnings News and Small Traders. Journal of Accounting and Economics 15, 265-302.

Lee, Charles M. C., and Balkrishna Radhakrishna, 2000. Inferring Investor Behavior: Evidence from TORQ Data. Journal of Financial Markets 3, 83-111.

Lee, Charles M. C., and Mark J. Ready, 1991. Inferring Trade Directions from Intraday Data. Journal of Finance 46, 733-746.

Libby, Robert, Robert Bloomfield and Mark W. Nelson, 2001. Experimental Research in Financial Accounting. Working paper, Cornell University.

Odders-White, Elizabeth R, 2000. On the Occurrence and Consequences of Inaccurate Trade Classification. Journal of Financial Markets 3, 259-286.

Richardson, Scott A., Richard G. Sloan, Mark T. Soliman and A. Irem Tuna, 2004. Accrual Reliability, Earnings Persistence and Stock Prices. Working paper, University of Pennsylvania, University of Michigan and Stanford University.

Shanthikumar, Devin, 2004. Small and Large Trader Behavior: Reactions to Information in Financial Markets. Dissertation, Stanford Graduate School of Business.

Shanthikumar, Devin, 2004. Small and Large Trades Around Earnings Announcements: Does Trading Behavior Explain Post-Earnings-Announcement Drift? Working paper, Harvard Business School.

Sloan, Richard G., 1996. Do Stock Prices Fully Reflect Information in Accruals and Cash Flows About Future Earnings? The Accounting Review 71, 289-315.

Smith, James F. and Thomas Kida, 1991. Heuristics and Biases: Expertise and Task Realism in Auditing. Psychological Bulletin 109, 472-489.

Swaminathan, Bhaskaran and Charles M. C. Lee, 2000. Do Stock Prices Overreact to Earnings News? Working paper, Cornell University.

Tversky, Amos and Daniel Kahneman, 1974. Judgment Under Uncertainty: Heuristics and Biases. Science 185, 1124-1131.

Whitecotton, Stacey M., 1996. The Effects of Experience and Confidence on Decision Aid Reliance: A Causal Model. Behavioral Research in Accounting 8, 194-216.

Table 1
Earnings Surprise Sample Statistics

This table presents sample statistics for firm-quarters for NYSE common stock from 1993 through 2002, with matching TAQ trading data. Size represents market value (shares outstanding * price-per-share) and is reported in millions of dollars. Book-to-market ratio (B/M) is the ratio of book value of equity to market value of equity (size). SUE represents the standardized unexpected earnings based on a random-walk with drift earnings expectations model. If a surprise is the first in a series of same-type (top or bottom 30%) surprises, it receives a value of N=1. If it is the second in a series, N=2 and so on. N=0 indicates a surprise in the middle 40%.

		All N	N = 0	N = 1	N = 2	N = 3	N = 4	N >= 4
Sample Size		58,945	25,797	15,923	7,012	4,065	2,795	6,148
Size	Mean	4574	4703	4658	4470	4280	4190	4127
	Median	867	901	890	836	787	796	761
	Std. Dev.	16359	16520	15613	16403	17040	17744	17042
B/M	Mean	0.511	0.455	0.481	0.623	0.622	0.624	0.627
	Median	0.490	0.468	0.509	0.513	0.505	0.508	0.492
	Std. Dev.	6.170	7.146	7.569	0.799	0.921	1.235	1.013
SUE	Mean	-0.271	0.050	-0.603	-0.336	-0.404	-0.468	-0.431
	Median	0.050	0.060	-0.190	-0.255	-0.320	-0.450	-0.260
	Std. Dev.	26.331	0.245	49.150	3.418	3.693	3.957	3.815
SUE - top 30%	Mean	1.595		1.561	1.597	1.669	1.966	1.632
	Median	1.090		1.000	1.130	1.230	1.490	1.210
	Std. Dev.	1.638		1.797	1.571	1.466	1.542	1.354
SUE - bottom 30%	Mean	-2.491		-2.726	-2.161	-2.282	-2.523	-2.412
	Median	-1.150		-1.020	-1.220	-1.320	-1.470	-1.340
	Std. Dev.	47.713		69.086	3.678	4.077	4.201	4.330

Table 2

Correlations Between TAQ-based and Brokerage Account-based Trading Variables

This table displays correlations between the monthly sum of daily TAQ-based abnormal trade imbalance, for the given trade-size category, and monthly trade imbalance measures derived from 50,000 retail brokerage accounts. The sample comprises all NYSE common stock from 1993 through 1996 with both TAQ and brokerage data available. P-values are in parentheses.

	Brokerage Trade Imbalance		
	Number of Trades	Number of Shares	Dollar Value
Small Trades, less than $5,000	0.1976 (0.0000)	0.1740 (0.0000)	0.1738 (0.0000)
Small Trades, less than $20,000	0.2069 (0.0000)	0.1863 (0.0000)	0.1861 (0.0000)
Large Trades, at least $50,000	-0.0343 (0.0000)	-0.0402 (0.0000)	-0.0401 (0.0000)

Table 3
Series of Earnings Increases and Decreases

This table presents the coefficients from a regression of abnormal trade imbalance, summed over days -5 through 5, on a series of indicators for the earnings increase or decrease relative to the same quarter in the previous year, and the position of that increase or decrease in a string of the same sign of earnings change (i.e. a string of all earnings increases). If an announcement is the first increase (decrease) in a series of same-sign earnings changes, it is defined as an INC_1 (DEC_1) surprise. If it is the second in a series, it is INC_2 (DEC_2) and so on. The earnings announcement sample contains all non-zero earnings changes for NYSE sample firms from 1993 through 2002. <$5,000, <$20,000 and >=$50,000 indicate the trade size used to measure trade reaction. Standard errors are in parentheses.

	SUE, with drift			number of
	<$5,000	<$20,000	>=$50,000	N Inc/Dec
DEC_1	0.4923	0.5279	-0.0034	8,956
	(0.0540)	(0.0547)	(0.0422)	
DEC_2	0.2451	0.2205	0.0567	4,767
	(0.0731)	(0.0752)	(0.0580)	
DEC_3	0.2645	0.3098	0.0371	3,035
	(0.0911)	(0.0942)	(0.0727)	
DEC_4	0.0482	0.0526	0.0775	2,286
	(0.1049)	(0.1084)	(0.0837)	
DEC_5	0.1029	0.1479	-0.0421	1,231
	(0.1419)	(0.1479)	(0.1141)	
DEC_6	0.4616	0.5486	-0.1083	683
	(0.1924)	(0.1994)	(0.1537)	
DEC_{7+}	0.1900	0.2610	-0.1348	814
	(0.1755)	(0.1835)	(0.1417)	
INC_1	0.4560	0.5092	0.1850	11,525
	(0.0544)	(0.0548)	(0.0422)	
INC_2	0.5131	0.5185	0.1091	8,236
	(0.0689)	(0.0687)	(0.0530)	
INC_3	0.6703	0.6477	0.2288	6,655
	(0.0815)	(0.0809)	(0.0624)	
INC_4	0.6696	0.7274	0.1828	6,030
	(0.0890)	(0.0878)	(0.0678)	
INC_5	0.8522	0.8545	0.0617	5,062
	(0.1060)	(0.1038)	(0.0801)	
INC_6	0.8058	0.7135	0.0899	4,429
	(0.1225)	(0.1192)	(0.0920)	
INC_{7+}	0.7329	0.7552	0.2209	11,814
	(0.0561)	(0.0533)	(0.0411)	
R^2	0.0132	0.0126	0.0014	
N	47,665	55,003	55,151	

Table 4
Abnormal Trade Reaction to Earnings Surprise, by Trade-Size Category

Panel A presents the coefficient estimates for a regression of abnormal trade imbalance, summed over days -5 through 5, on a constant and the earnings surprise decile (0-9), interacted with indicators for N value. If a surprise is the first in a series of same-type (top or bottom 30%) surprises, it receives a value of N=1. If it is the second in a series, N=2 and so on. The earnings surprise sample contains all extreme (top or bottom 30%) earnings surprises for NYSE sample firms from 1993 through 2002. Panel B displays differences between the coefficient estimates for earnings surprise decile interacted with N value. Standard errors are in parentheses.

Panel A. Coefficient Estimates

	Trade size category, in thousands of dollars				
	<5	5-10	10-20	20-50	>=50
N=1	0.5783	0.4452	0.2778	0.1525	-0.0221
	(0.0652)	(0.0586)	(0.0553)	(0.0520)	(0.0507)
N=2	0.3229	0.2933	0.1467	0.0384	0.1217
	(0.0946)	(0.0859)	(0.0812)	(0.0764)	(0.0744)
N=3	0.2470	0.2935	0.2281	0.0761	0.0885
	(0.1206)	(0.1103)	(0.1044)	(0.0982)	(0.0958)
N>=4	0.0235	0.0186	-0.0880	-0.0232	-0.1125
	(0.0991)	(0.0908)	(0.0860)	(0.0809)	(0.0789)
surpDec*(N=1)	-0.0079	-0.0120	0.0087	0.0158	0.0338
	(0.0116)	(0.0103)	(0.0098)	(0.0092)	(0.0089)
surpDec*(N=2)	0.0377	0.0137	0.0329	0.0187	0.0065
	(0.0169)	(0.0151)	(0.0143)	(0.0134)	(0.0131)
surpDec*(N=3)	0.0813	0.0533	0.0500	0.0505	0.0220
	(0.0218)	(0.0194)	(0.0183)	(0.0173)	(0.0168)
surpDec*(N>=4)	0.0830	0.0830	0.0687	0.0371	0.0326
	(0.0177)	(0.0158)	(0.0149)	(0.0140)	(0.0137)
R^2	0.0129	0.0090	0.0065	0.0032	0.0018
N	28,274	32,104	32,416	32,455	32,512

Panel B. Coefficient Differences, for Earnings Surprise Decile Term

	Trade size category, in thousands of dollars				
Coef. Differences Between:	<5	5-10	10-20	20-50	>=50
(N=2) and (N=1)	0.0456	0.0257	0.0242	0.0029	-0.0273
	(0.0205)	(0.0183)	(0.0173)	(0.0163)	(0.0158)
(N=3) and (N=2)	0.0437	0.0396	0.0171	0.0318	0.0154
	(0.0276)	(0.0246)	(0.0232)	(0.0219)	(0.0213)
(N>=4) and (N=3)	0.0016	0.0296	0.0187	-0.0134	0.0106
	(0.0281)	(0.0250)	(0.0236)	(0.0222)	(0.0217)
(N=3) and (N=1)	0.0893	0.0653	0.0414	0.0347	-0.0119
	(0.0247)	(0.0220)	(0.0208)	(0.0195)	(0.0190)
(N>=4) and (N=2)	0.0453	0.0692	0.0358	0.0184	0.0261
	(0.0245)	(0.0219)	(0.0207)	(0.0194)	(0.0189)

Table 5
Relationship Between Abnormal Trade Imbalance and Prior Returns

This table presents results from regressing abnormal trade imbalance on prior cumulative abnormal returns. Small trades are trades with a value of less than $5,000 or a value of less than $20,000 and large trades are at least $50,000. Cumulative abnormal returns are CRSP beta-adjusted daily returns, summed over the respective period. The sample includes all NYSE common stock with the necessary TAQ and CRSP data, for the period 1993 through 2002. White standard errors, using within-firm clustering, are in parentheses.

	Trade Size Category		
	Less Than $5,000	Less Than $20,000	At Least $50,000
R(-1)	0.1460	-0.0866	0.9991
	(0.0178)	(0.0344)	(0.0287)
R(-2,-5)	-0.1896	-0.4551	0.3831
	(0.0093)	(0.0189)	(0.0133)
R(-6,-20)	-0.2232	-0.3872	0.1115
	(0.0051)	(0.0125)	(0.0066)
R(-21,-60)	-0.1526	-0.2051	-0.0001
	(0.0033)	(0.0076)	(0.0040)
constant	0.0058	0.0072	0.0022
	(0.0006)	(0.0004)	(0.0003)
R^2	0.0013	0.0031	0.0013
N	3,261,828	3,768,623	3,768,623

Table 6

Robustness Checks: Alternate Trade Imbalance, Mid-Fiscal-Year Results, Fixed Effects, and Analyst-based Earnings Surprises

This table presents regression results for several robustness checks and variations on the primary regression. The dependant variable is trade imbalance, summed over days -5 through 5. The specific variations are described below. Panel A present results trades smaller than $5,000, and limits the sample to securities with value less than $50 so such a trade is possible. Panel B presents the results for trades less than $20,000, when stock price is less than $200. Panel C presents results for trades of at least $50,000. . If a surprise is the first in a series of same-type (top or bottom 30%) surprises, it receives a value of N=1. If it is the second in a series, N=2 and so on. The earnings surprise sample contains all extreme (top or bottom 30%) earnings surprises for NYSE sample firms from 1993 through 2002, except where described. Standard errors in parentheses.

Alternate Trade Imbalance Measures

a. Abnormal Trade Imbalance, with additional adjustments for response to prior day, week, month and quarter returns.
b. Alternate Abnormal Trade Imbalance measure, with adjustments for calendar-time effects, prior returns and prior trade-imbalances.

Mid-Fiscal-Year

Sample excludes earnings announcements pertaining to the last quarter in the firm's fiscal year.

Fixed Effects (sample includes N=0 surprises, i.e. middle 40% surprises, in constant term)

Security. Includes fixed effect for each unique security, as identified by 8-digit CUSIP.

Quarter. Includes fixed effect for each quarter-year combination.

Analyst Surprise

Earnings surprise based on IBES summary file monthly consensus forecasts, using the median forecast, with at least four forecasts outstanding, as the earnings expectation.

Panel A. Trades <$5000

	Alternate Imbalance		Mid-Fiscal-	Fixed Effects		Analyst
	a	b	Year	Security	Quarter	Surprise
N=1	0.4486	0.0658	0.5697	0.1055	0.0817	0.7520
	(0.0677)	(0.0385)	(0.0778)	(0.0759)	(0.0732)	(0.0878)
N=2	0.1502	-0.0951	0.0816	-0.1727	-0.1925	0.5532
	(0.0979)	(0.0558)	(0.1093)	(0.1034)	(0.0999)	(0.1449)
N=3	0.0395	-0.1406	0.1772	-0.3008	-0.2416	0.1224
	(0.1254)	(0.0712)	(0.1387)	(0.1289)	(0.1239)	(0.2206)
N>=4	-0.1226	-0.2928	0.0519	-0.5734	-0.4751	0.2926
	(0.1029)	(0.0585)	(0.1179)	(0.1131)	(0.1039)	(0.2133)
surpDec*(N=1)	0.0027	0.0171	0.0020	-0.0089	-0.0066	-0.0500
	(0.0120)	(0.0068)	(0.0139)	(0.0117)	(0.0115)	(0.0153)
surpDec*(N=2)	0.0564	0.0521	0.0775	0.0421	0.0468	-0.0245
	(0.0177)	(0.0100)	(0.0198)	(0.0172)	(0.0169)	(0.0245)
surpDec*(N=3)	0.1060	0.0658	0.0928	0.0918	0.0831	0.0022
	(0.0228)	(0.0129)	(0.0254)	(0.0222)	(0.0216)	(0.0364)
surpDec*(N>=4)	0.0912	0.1033	0.0889	0.1057	0.0858	0.0488
	(0.0184)	(0.0105)	(0.0211)	(0.0189)	(0.0176)	(0.0338)
const				0.4908	0.4920	
				(0.0361)	(0.0341)	
R^2	0.0098	0.0079	0.0146	0.0010	0.0010	0.0105
N	26,196	28,274	19,998			13,579

Panel B. Trades <$20000

	Alternate Imbalance a	Alternate Imbalance b	Mid-Fiscal-Year	Fixed Effects Security	Fixed Effects Quarter	Analyst Surprise
N=1	0.3767	0.0647	0.5630	0.0886	0.0677	0.8251
	(0.0680)	(0.0398)	(0.0779)	(0.0762)	(0.0735)	(0.0891)
N=2	0.1062	-0.1389	0.0407	-0.1620	-0.2073	0.6998
	(0.0997)	(0.0584)	(0.1118)	(0.1051)	(0.1018)	(0.1487)
N=3	0.1021	-0.0661	0.3279	-0.1605	-0.1580	0.3157
	(0.1287)	(0.0751)	(0.1425)	(0.1322)	(0.1273)	(0.2263)
N>=4	-0.1423	-0.3249	0.0333	-0.5201	-0.4550	0.3084
	(0.1060)	(0.0619)	(0.1213)	(0.1158)	(0.1068)	(0.2164)
surpDec*(N=1)	0.0176	0.0287	0.0030	-0.0036	-0.0034	-0.0569
	(0.0120)	(0.0070)	(0.0139)	(0.0117)	(0.0115)	(0.0156)
surpDec*(N=2)	0.0634	0.0672	0.0811	0.0359	0.0446	-0.0220
	(0.0176)	(0.0103)	(0.0198)	(0.0172)	(0.0169)	(0.0250)
surpDec*(N=3)	0.1199	0.0882	0.0855	0.0823	0.0835	0.0003
	(0.0226)	(0.0132)	(0.0253)	(0.0222)	(0.0216)	(0.0370)
surpDec*(N>=4)	0.1189	0.1206	0.1084	0.1117	0.1053	0.0505
	(0.0183)	(0.0107)	(0.0210)	(0.0188)	(0.0176)	(0.0339)
const				0.4988	0.5004	
				(0.0358)	(0.0338)	
R^2	0.0107	0.0116	0.0146	0.0010	0.0011	0.0119
N	30,266	32,416	22,913			15,700

Panel C. Trades >= $50000

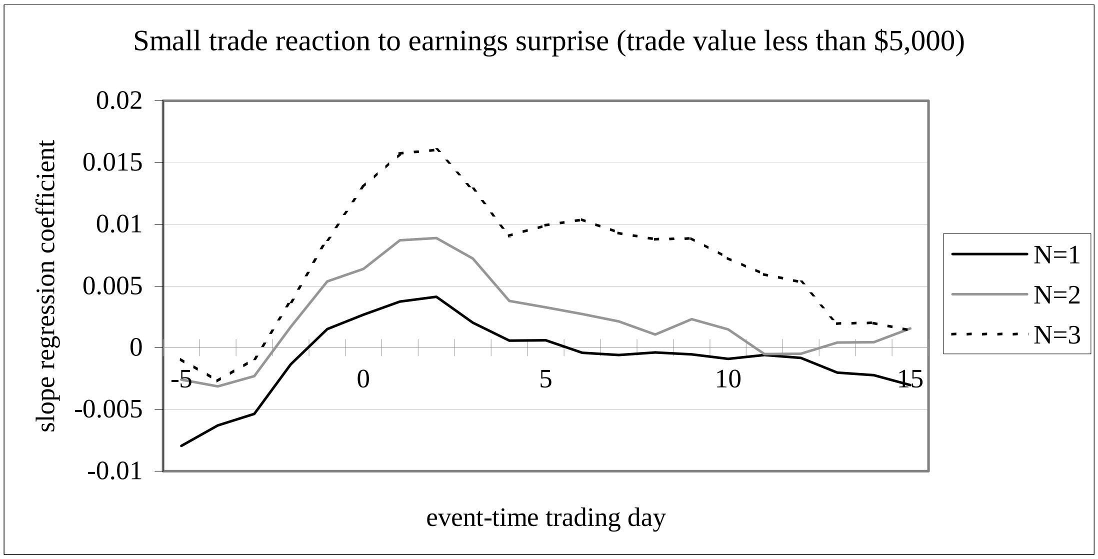
	Alternate Imbalance a	Alternate Imbalance b	Mid-Fiscal-Year	Fixed Effects Security	Fixed Effects Quarter	Analyst Surprise
N=1	0.0214	-0.0625	-0.0316	-0.1136	-0.1217	-0.1747
	(0.0517)	(0.0298)	(0.0610)	(0.0586)	(0.0570)	(0.0654)
N=2	0.1170	-0.0386	0.0982	0.0218	0.0048	-0.0617
	(0.0759)	(0.0438)	(0.0877)	(0.0808)	(0.0790)	(0.1091)
N=3	0.0867	0.0014	0.1006	-0.0045	-0.0104	0.0583
	(0.0980)	(0.0563)	(0.1117)	(0.1017)	(0.0988)	(0.1661)
N>=4	-0.0841	-0.0429	-0.0879	-0.1987	-0.2141	-0.1503
	(0.0808)	(0.0464)	(0.0952)	(0.0892)	(0.0829)	(0.1588)
surpDec*(N=1)	0.0202	0.0134	0.0300	0.0326	0.0341	0.0736
	(0.0091)	(0.0053)	(0.0109)	(0.0090)	(0.0089)	(0.0114)
surpDec*(N=2)	-0.0038	0.0048	0.0194	0.0055	0.0140	0.0626
	(0.0134)	(0.0077)	(0.0155)	(0.0133)	(0.0131)	(0.0183)
surpDec*(N=3)	0.0104	0.0078	0.0172	0.0146	0.0221	0.0244
	(0.0172)	(0.0099)	(0.0198)	(0.0171)	(0.0167)	(0.0272)
surpDec*(N>=4)	0.0203	0.0016	0.0290	0.0234	0.0337	0.0431
	(0.0139)	(0.0080)	(0.0165)	(0.0145)	(0.0136)	(0.0249)
const				0.1066	0.0990	
				(0.0275)	(0.0262)	
R^2	0.0009	0.0003	0.0016	0.0005	0.0005	0.0055
N	30,354	32,512	22,980			15,747

Figure 1
Small and Large Trade Reaction to a Sequence of Similar Earnings Surprises

If a surprise is the first in a series of same-type extreme surprise, it receives a value of N=1. If it is second in a series, N=2 and so on. A small trade is defined as one with a value of less than $5,000. A large trade is defined as one with a value of at least $50,000. For small trades, the sample is restricted to days on which a stock price is below $50, so that such a trade is possible.

The figure presents results from a regression of abnormal small trade imbalance on an intercept and the earnings surprise decile (0-9), using a 5-day moving average. The figure plots β, the slope coefficient trade reaction to each earnings announcement. The values on the time axis represent the midpoints of each five-day interval.

Panel A. Small Trades



Panel B. Large Trades