Individual Investor Trading and Stock Returns

Ron Kaniel, Gideon Saar, and Sheridan Titman

First version: February 2004
This version: May 2005

Ron Kaniel is from the Faqua School of Business, One Towerview Drive, Duke University, Durham, NC 27708 (Tel: 919-660-7656, ron.kaniel@duke.edu). Gideon Saar is from the Johnson Graduate School of Management, Cornell University, 455 Sage Hall, Ithaca, NY 14853 (Tel: 607-255-7484, gs25@cornell.edu). Sheridan Titman is from the McCombs School of Business, University of Texas at Austin, Austin, TX 78712 (Tel: 512-232-2787, Sheridan.Titman@mccombs.utexas.edu). We wish to thank Shuming Liu for dedicated research assistance. We are grateful for comments from Simon Gervais, Larry Harris, Joel Hasbrouck, Roni Michaely, Maureen O'Hara, Lei Yu and seminar (or conference) participants at Cornell University, Duke University, INSEAD, London Business School, New York University, and the American Finance Association meetings. This research began while Saar was on leave from NYU and held the position of Visiting Research Economist at the New York Stock Exchange. The opinions expressed in this paper do not necessarily reflect those of the members or directors of the NYSE.

Individual Investor Trading and Stock Returns

Abstract

This paper investigates a unique dataset that enables us to determine the aggregate buy and sell volume of individual investors for a large cross-section of NYSE stocks. We find that individuals trade as if they are contrarians, and that the stocks that individuals buy exhibit positive excess returns in the following month. These patterns are consistent with the idea that risk-averse individuals provide liquidity to meet institutional demand for immediacy. We further examine the relation between net individual trading and short-horizon (weekly) return reversals that have been documented in the literature. Our results reveal that net individual trading predicts future returns, and that the information content of past trading by individuals is distinct from that of past return or past volume. Furthermore, net individual trading predicts weekly returns in the post-2000 era for stocks of all sizes, while past return seems to have lost its predictive power for all but small stocks over the same time period. Lastly, we note that net individual trading activity does not seem highly correlated across the stocks in our sample.

1. Introduction

For a variety of reasons, financial economists tend to view individuals and institutions differently. Institutions are generally much larger, are more sophisticated, and are believed to be better informed than individual investors. Individuals, on the other hand, are said to have psychological biases and are often thought of as the proverbial noise traders in the sense of Kyle (1985) or Black (1986).

This study examines the investment choices of individual investors with a unique dataset that was provided to us by the NYSE. The dataset was constructed from the NYSE's Consolidated Equity Audit Trail Data (CAUD) files that contain detailed information on all orders that execute on the exchange. For each stock on each day, we have the aggregated volume of executed buy and sell orders of individuals. We create a daily measure of net individual investor trading for each stock by subtracting the sell volume of individuals from their buy volume and dividing by the average daily volume of the stock in the previous year.

Our focus is on the dynamic relation between individual investor trading and returns over relatively short horizons (e.g., weekly and monthly). We examine the extent to which purchases and sales of shares by individuals are influenced by past returns and the extent to which individual trades predict future returns. Consistent with earlier studies, we find that individuals tend to be contrarians, at least in the short-run. The mean market-adjusted return in the 20 days prior to a week of intense individual selling is 3.51%, while prior to a week of intense individual buying it is −2.12%.[1] More interestingly, we find that the trades of individuals can be used to forecast future returns. Specifically, we find that stocks experience excess returns of 1.49% in the 20 days following a week of intense buying by individuals.

[1] In contrast, there are a number of studies suggesting that institutions tend to be momentum traders (e.g., Grinblatt, Titman, and Wermers (1995); Nofsinger and Sias (1999); Wermers (1999); Sias (2003); Sias, Starks, and Titman (2003)).

This positive relation between individual trades and future returns can be interpreted in a number of ways. One possibility is that the individuals who trade through the NYSE possess private information about fundamentals. This is possible, since the trades of the more sophisticated individuals tend to be executed on the NYSE. However, we think it is unlikely that the trades of individual investors on the NYSE can be viewed as informed relative to the trades of institutions, which are likely to be the counterparties to the individual investor trades.

An alternative explanation that we find more appealing is that the contrarian tendency of individuals leads them to act as liquidity providers to other investors (e.g., institutions) that require immediacy. Following Stoll (1978), Grossman and Miller (1988), and Campbell, Grossman, and Wang (1993), one can argue that investors who require immediacy must offer price concessions to induce risk-averse individuals to take the other side of their trades, and that this, in turn, results in subsequent return reversals. These return reversals show up as short horizon excess returns following concentrated individual buying. Hence, over short intervals, individuals may outperform institutions, even when they are at an information disadvantage. Indeed, we find that the excess returns that individuals earn are greater when they buy less liquid stocks, which is consistent with this explanation.

In addition, since individuals tend to be contrarians, their profits may also relate to the short-horizon return reversals first observed by Jegadeesh (1990) and Lehmann (1990). In principle, these reversals can be due to either investor overreaction or to illiquidity.[2] If the return reversals are due to overreaction, then it may be the case that the

[2] Jegadeesh (1990) and Lehmann (1990) both discuss the possibility of overreaction. Lehmann (1990) also suggests that frictions in liquidity provision may explain the weekly reversals and Jegadeesh and Titman (1995), who examine the relation between return reversals and bid-ask spreads, provide evidence that is consistent with a liquidity explanation for daily reversals. More recently, Subrahmanyam (2005) develops a model to distinguish between illiquidity and overreaction. He tests the model using the Lee and Ready (1991) algorithm to indirectly measure whether buyers (and equivalently sellers) are providers or demanders of liquidity. The results of his tests are inconsistent with the liquidity provision explanation.

short-horizon profits of individuals come from their contrarian tendencies. If this is the case, then we might expect the abnormal returns of individuals to diminish after controlling for past returns. Alternatively, if return reversals arise because of illiquidity, and if the aggregate net trading of individuals provides a better measure of the demand for immediacy than past price changes, then one might expect individual trades to be a better predictor of short-horizon returns than past returns.

In addition to past returns, our analysis controls for trading volume in light of evidence that indicates that volume is related to short-horizon returns (e.g., Conrad, Hameed, and Niden (1994); Gervais, Kaniel, and Mingelgrin (2001); Llorente, Michaely, Saar, and Wang (2002)). Volume can arise from shocks to investor hedging needs, private information, or trader interest in a given stock. Since such shocks can give rise to demand by individuals, it is possible that volume and net individual investor trading contain the same information about future returns.

To evaluate these different possibilities, we examine the returns of portfolios constructed from independent sorts on net individual trading, trading volume and returns. In a double-sort on net individual trading and past returns we find a relation between individual trading and future returns but no evidence of an independent past returns effect. Sorting on net individual trading and volume shows that both variables predict returns but seem to contain different information. We also run multivariate regressions of weekly returns on past returns, volume, and net individual trading. The results of these regressions indicate that trading by individuals is a powerful predictor of future returns that is not subsumed by either past returns or past volume. Correcting for bid-ask bounce and nonsynchronous trading causes past return to lose its predictive power, but net individual trading remains a significant predictor.

Finally, we also consider the possibility that the higher returns associated with individual trades arise because individuals induce excess volatility. Such a connection was suggested by the theoretical work of De Long, Shleifer, Summers, and Waldmann

(1990a) on noise traders (most often associated in the limits-to-arbitrage literature with individual investors). We investigate this and find only temporary shifts in volatility. We also look at the question of whether the actions of individuals are “systematic” in the sense that they affect all stocks at the same time. We conduct a principal component analysis of net individual trading and find very little correlated actions of individuals across stocks: the first principal component of this variable explains only 1.70% of the variance over and above a simulated benchmark created from independent data.

We are not, of course, the first to consider the investment behavior of individual investors. Most existing empirical work, however, investigates individual investor trading abroad due to the scarcity of suitable U.S. data. As we mentioned earlier, our finding that individual investors tend to be contrarians is consistent with earlier evidence which includes, Choe, Kho, and Stulz (1999), who find evidence of contrarian choices using Korean data, Grinblatt and Keloharju (2000), who find similar results in a study of Finish individuals, Jackson (2003), who documents this behavior in a study of Australian individuals, and Richards (2005), who shows contrarian tendencies of individuals in six Asian markets. In addition, in a study of U.S. individuals who invest in an index fund, Goetzmann and Massa (2002) find that countrarians outnumber momentum traders two to one, and in studies of individual investors who trade using one of the major U.S. discount brokers, Odean (1998, 1999) find that individuals tend to hold on to their losers and sell their winners, which is somewhat different but consistent with the idea that individuals are contrarians.[3]

Many of the above studies also examine the investment performance of individual investors and, in contrast to our evidence, most find that individuals do poorly. In particular, Odean (1999) and Grinblatt and Keloharju (2000), looking at longer horizons, find that individual investors make poor investment choices. Like us, Odean studies U.S.

---

[3] Bailey, Kumar, and Ng (2004) use the same dataset as Odean and find that U.S. individuals who invest abroad also exhibit contrarian behavior (relative to the foreign country’s stock index).

stocks, but the broker that provides his data executes most of its trades off the NYSE, so his sample does not overlap with ours. Barber, Lee, Liu, and Odean (2004a) examine the performance of individuals in Taiwan, and find losses at short as well as long horizons.[4] Our results also contrast with Griffin, Harris, and Topaloglu (2003), who find no significant relation between the trading imbalances taken from brokers who predominantly serve individuals and the future daily returns of NASDAQ stocks, and with Barber, Odean, and Zhu (2003), who find that stocks bought by clients of two U.S. brokerage firms do not reliably underperform or overperform the stocks they sold. However, our finding of excess returns following purchases by individuals is similar to the Australian evidence in Jackson (2003) that individuals perform well over shorter horizons.[5]

The rest of the paper is organized as follows. The next section presents the sample and the unique dataset we use. Section 3 presents analysis of the dynamic relation between net individual trading and returns. The investigation of short-horizon return predictability and its relation to net individual trading is carried out in Section 4. Section 5 discusses interpretations of the results and provides additional evidence on competing explanations. Section 6 concludes.

2. Data and Sample

We study the trading of individuals using a special dataset that was provided to us by the New York Stock Exchange (NYSE). The dataset contains four years of aggregated daily buy and sell volume of individual investor orders for a large cross section of NYSE stocks. The dataset was constructed from the NYSE's Consolidated Equity Audit Trail

---

[4] The behavior of individuals in Taiwan seems to be somewhat different from the behavior of their U.S. counterparts. Many individuals in Taiwan engage in active trading (including day trading, see Barber, Lee, Liu, and Odean (2004b)), and annual turnover on the Taiwan Stock Exchange averaged 292% over their sample period (1995-1999), compared with 69% on the NYSE.

[5] San (2004) uses quarterly data on institutional and insider holdings to identify stocks with more individual investor trading. She finds that after adjusting for risk, stocks individuals buy outperform those they sell.

Data (CAUD) files that contain detailed information on all orders that execute on the exchange, both electronic and manual (those handled by floor brokers). One of the fields associated with each order, called Account Type, specifies whether the order comes from an individual investor.

The Account Type designation of individual investor orders has its origins in the aftermath of October 1987. The NYSE introduced the Individual Investor Express Delivery Service that provides priority delivery of orders that have been identified as individual investor orders.[6] The goal of the service is to ensure that individual investors are not disadvantaged relative to professional investors in periods of extreme market conditions. In order to implement the system, new Account Type categories that identify individual investors were created in October 1988, and orders coming from individual investors are now marked as such by their brokers (Account Type is a mandatory field a broker has to fill for each order that is sent to the NYSE).

The Account Type field is not audited by the NYSE on an order-by-order basis. It is reasonable to assume, however, that individual investor orders are marked as such because designating an order as coming from an individual investor has some advantages. At the same time, NYSE officials monitor the use of this field by brokers. Any abnormal use of the individual investor designation in the Account Type field by a brokerage firm is likely to draw attention, which prevents abuse of the system. We therefore believe that the Account Type designation of individual investor orders is fairly accurate.

Our sample contains all common, domestic stocks that were traded on the NYSE any time between January 1, 2000 and December 31, 2003.[7] We use the CRSP database to construct the sample, and match the stocks to the NYSE dataset by means of ticker symbol and CUSIP. This procedure results in a sample of 2,034 stocks. An important

---

[6] The service is activated when the Dow Jones Industrial Average moves more than a certain amount up or down from the previous day's close. When the Individual Investor Express Delivery Service was introduced in October 1988, the threshold was a 25-point move from the previous day’s close.

[7] The NYSE does not store CAUD data for the period prior to January 2000.

advantage of this dataset is that the information about daily buy and sell volume of individual investors was created by aggregating executed *orders*, rather than trades. In other words, the classification into buy and sell volume in our dataset is exact, and we do not have to rely on classification algorithms such as the one proposed by Lee and Ready (1991). Table 1 presents summary statistics for the entire sample and for three size groups.

We should note that some brokers either sell some of their order flow (in NYSE-listed stocks) to wholesalers for execution or internalize a certain portion of their clients' orders by trading as principal against them. Since such pre-arranged trading practices cannot be carried out on the NYSE, these trades take place on one of the regional exchanges (or alternatively reported to the NASD) and are therefore not in our sample of NYSE executions. For example, Schwab internalized 66% of its orders in the fourth quarter of 2003, while Fidelity sent about 38% of its volume in NYSE-listed stocks to the Boston Stock Exchange to be executed by its own specialist.[8] However, it is very likely that the fraction of volume these brokers send to the NYSE consists of orders that create an imbalance not easily matched internally. This means that imbalances in the orders of individuals find their way to the NYSE even if some of the more balanced individual volume is executed elsewhere. Therefore, our net individual trading measure (detailed below) that captures imbalances in individuals' executed orders on the NYSE probably reflects (even if not perfectly) the individuals' imbalances in the market as a whole.

We construct a daily measure of net individual investor trading by subtracting the value of the shares sold by individuals from the value of shares bought, and standardize the measure by the average daily dollar volume. Specifically, we define Net Individual Trading (NIT) for stock *i* on day *t* as:

[8] These figures are taken from an article by Kate Kelly in the Wall Street Journal ("SEC Overhaul Could Topple Best-Price Rule," March 5, 2004).

$$NIT_{i,t} = \frac{\text{Individual buy dollar volume}_{i,t} - \text{Individual sell dollar volume}_{i,t}}{\text{Average daily dollar volume in previous year}_{i}}$$

where the denominator is the stock's average daily dollar volume (from CRSP) for the year ending on day *t-1*.[9]

3. Dynamic Relation between Net Individual Trading and Returns

We start by aggregating daily net individual trading to create a weekly NIT measure and identify those weeks where either positive or negative NIT is most pronounced. This is done by comparing each stock's NIT value in a given week (the formation week) with the values of NIT in the previous 9 weeks. Based on this comparison we place the stocks in decile portfolios, where decile 1, the "intense selling portfolio," contains stocks for which NIT in the formation week is the most negative relative to their NITs in the previous 9 weeks. Similarly, decile 10, the "intense buying portfolio," contains stocks with the most positive NIT relative to the previous 9 weeks. For robustness, we also look at the results for somewhat less intense trading by forming a selling portfolio from the stocks in deciles 1 and 2, and a buying portfolio from the stocks in deciles 9 and 10.

We look at the more extreme deciles because our goal is to identify periods in which individuals in the aggregate are net buyers or net sellers. This portfolio formation procedure, similar in spirit to the methodology in Gervais, Kaniel, and Mingelgrin (2001), has the advantage that it uses a moving average of nine weeks and therefore is robust to a potential trend in the measure. The reason we adopt the methodology of forming deciles by comparing a stock's NIT in the formation week relative to its own past NIT is because the impact of trading imbalances on future prices should be related to each stock's ability to absorb order flow.[10]

---

[9] For example, to compute the denominator for February 3, 2000 (for a certain stock) we average the daily dollar volume over all trading days from February 3, 1999 to February 2, 2000.

[10] Subrahmanyam (2005) makes a similar point stating that inventory control effects predict a downward pressure on the price of a stock in the absolute rather than the relative (cross-sectional) sense.

Table 2 presents the cumulative market-adjusted returns for these four net individual trading portfolios.[11] These cumulative returns are calculated for 20, 15, 10 and 5 days before the first day or after the last day of the formation week. The cells in the table contain the time-series means and t-statistics for each of the cumulative return measures. The first line of the table shows that intense individual selling (decile 1) follows an increase in the prices of stocks. The mean excess return in the 20 days prior to the selling week is 3.51%, and the mean excess return in the five days prior to that week is 1.68%. These returns are highly statistically significant. The last line of the table describes the returns in the week prior to intense individual buying activity (decile 10). The excess return in the 20 days prior to intense buying is –2.12%, and is highly statistically significant. We get similar results with the less extreme portfolios (deciles 1 and 2 for selling, and deciles 9 and 10 for buying), suggesting that our findings are not driven by outliers.

The results in Table 2 indicate that U.S. individual investors can be characterized as contrarians, which is consistent with the findings regarding individual investors in Australia, Finland, and several Asian markets. The table also reveals that there are positive excess returns following weeks in which individuals accumulate shares. The portfolio of stocks in decile 10 earns 0.38% market-adjusted returns in the week after intense buying and 1.49% in the 20 days following portfolio formation (both statistically significant). On the other hand, market-adjusted returns following intense selling by individuals are not significantly different from zero.

To examine the robustness of these results we also formed NIT deciles based on weekly cross-sectional sorting and replicated the analysis in Table 2. The results were similar, and both contrarian tendencies and the return predictability on buying were statistically significant.

---

[11] We use the value-weighted portfolio of all stocks in the sample as a proxy for the market portfolio.

We also examined the robustness of our results to different definitions of the net individual trading measure. We used a non-standardized measure (without dividing by the average volume), a measure standardized by average volume over the entire sample period, and one standardized by predicted volume from a regression model. We also used several definitions of the deviations from net individual trading by subtracting the mean over the sample period, a moving average over the previous year, or the predicted value from a regression model. The results using all measures were very similar, showing significant contrarian patterns and return predictability following purchases by individuals.

We also examined the robustness of our results to different definitions of returns. Specifically, we repeated the analysis with excess returns from a market model regression, with industry-adjusted returns, with raw returns, and with returns generated from end-of-day quote midpoints (constructed using the TAQ database).[12] All return definitions generate similar and statistically significant results.

4. Short-Horizon (Weekly) Predictability of Returns

This section examines how our evidence relates to the Jegadeesh (1990) and Lehmann (1990) findings on short-horizon return reversals. Given that individuals tend to be contrarians, it is possible that the short-horizon excess returns associated with individual buys simply reflect the Jegadeesh and Lehmann return reversals.

To examine this issue we form 25 portfolios by independently sorting stocks into five quintiles based on their past week's return and five quintiles based on their NIT

[12] For industry-adjusted returns we used a classification into ten industry portfolios (based on four-digit SIC codes) made available by Kenneth French. The exact specification of the ten industry portfolios can be obtained from:
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_10_ind_port.html

measures relative to their past nine week average. For each of the portfolios we compute the market-adjusted return in the week following the formation week.[13]

Panel A of Table 3, which reports the time-series averages of the weekly market-adjusted returns for the 25 portfolios, reveals no apparent evidence of return reversals in our sample period when conditioning on NIT. The last two columns of the table look at the payoffs to a trading strategy that buys quintile 5 and sells quintile 1. If the return reversal strategy that buys the portfolio with last week's most negative return and sells the one with last week's most positive return can be used to generate profits, the payoffs in the column Q5–Q1 should be negative and significant. The table shows that the payoffs to this strategy are not statistically different from zero in any of the NIT quintiles.[14]

On the other hand, there is a pronounced pattern within each quintile of past returns going from past individual selling (NIT quintile 1) to past individual buying (NIT quintile 5). The market-adjusted return in each column of the table becomes more positive as we go from the stocks that individuals sold the previous week to those individuals bought. The bottom two lines of the panel provide information about the payoffs to buying a portfolio that is comprised of stocks that experience more intense individual buying in the previous week (NIT quintile 5) and selling those stocks experiencing intense individual selling (NIT quintile 1) in each return quintile. All these portfolios realize statistically significant positive payoffs, ranging from 0.24% to 0.60% per week.[15]

---

[13] We examined the robustness of our findings to different definitions of returns by repeating the analysis using market-model-adjusted returns, industry-adjusted returns, raw returns, and returns computed from end-of-day quote midpoints (as in Section 3). Our conclusions from all these return definitions were the same.

[14] We use the Newey-West correction in the computation of the t-statistics.

[15] The payoffs are in terms of percentage of dollar invested in the long position of this zero-investment strategy.

We also consider the possibility that net individual trading predicts returns because it serves as a proxy for volume, which was shown by Gervais, Kaniel, and Mingelgrin (2001) to predict future returns. To examine whether the NIT return relationship is independent of volume we repeat the analysis sorting the stocks each week into five quintiles of weekly NIT and five quintiles of turnover. The assignment of a stock into a turnover quintile on a given week follows the methodology in Gervias *et al.* and is similar in nature to the way we assign stocks each week into NIT quintiles (the turnover of a stock on a certain week is compared to the turnover of the same stock in the previous nine weeks). Based on these 5X5 sorts, 25 portfolios are formed as the intersection of the five turnover quintiles and five NIT quintiles, and their returns in the following week are calculated.

Panel B of Table 3 reveals that the information in the NIT measure is distinct from that in turnover, and both provide independent information about future returns. In particular, the strategy of buying the stocks in NIT quintile 5 and selling the stocks in NIT quintile 1 produces statistically significant payoffs in each turnover column, and the strategy of buying the stocks in turnover quintile 5 and selling those in turnover quintile 1 generates statistically significant payoffs in each NIT row.

The finding that both net trading of individual investors and turnover predict the subsequent week's return is especially interesting. Gervais, Kaniel, and Mingelgrin (2001) suggest that the high-volume return premium, or the tendency of prices to increase after periods with high turnover, is due to shocks in trader interest. If high volume attracts investor attention to the stock, the investor recognition hypothesis (e.g., Merton, 1987) argues that the stock value would increase due to better risk sharing. A reasonable candidate for a class of investors who do not follow all the stocks all the time but may be attracted to a certain stock after a volume shock brings media attention to it are individual investors. This reasoning suggests that conditioning on a variable that specifically measures individual investor trading could potentially explain the high-volume return premium, leaving no role for turnover. Our findings, however, suggest that turnover and NIT contain different information and neither of them subsumes the other.

To examine turnover, NIT, and past returns simultaneously we estimate a series of Fama and MacBeth (1973) regressions.[16] Table 4 presents the estimates of Fama and MacBeth regressions of returns in week *t* on a set of dummy variables that represent week *t*-1 return quintiles, turnover quintiles, and NIT quintiles. The results from these regressions are consistent with the findings from the portfolio sorting approach in Table 3. In particular, we find that NIT and turnover are both significant predictors of future returns in these multiple regressions. The relation between NIT and returns is quite strong for small, mid-cap, and large stocks, while turnover strongly predicts returns for the small and mid-cap stocks, but is only weakly related to returns for large stocks. After controlling for NIT and turnover, we find no evidence of return reversals for the entire sample and only weak evidence of return reversals for small stocks, (but not for mid-cap and large stocks).[17]

We also estimate a Fama-MacBeth regression specification where the current week's return is regressed on the past week's return rather than on dummy variables for past return quintiles. We do this for two reasons. First, one could argue that there is some loss of information associated with the transformation of returns into quintile dummy variables, and that this may bias our tests against finding a past return effect. Second, this

[16] Specifically, a cross-sectional regression is performed for each week in the sample period. Then, we construct test statistics based on the time-series of the estimated coefficients (using the Newey-West correction for the standard errors).

[17] While the Fama-MacBeth t-statistic on the mean coefficient of each of the four past return dummy variables is not different from zero, we also wanted to test the joint hypothesis that the coefficients on all four dummy variables are equal to zero. Unlike the situation in a regular regression framework where the joint hypothesis can be easily tested, the Fama-MacBeth specification does not satisfy the conditions necessary for an F-test. We therefore treated each set of coefficients on a single dummy variable (e.g., past return of quintile 2) from the cross-sectional regressions as a sample. This created four possibly related samples. We then tested the joint hypothesis that the means of the four samples are all equal to zero using a Friedman nonparametric test that allows for related samples. The test statistic could not reject the hypothesis that the mean coefficients on the dummy variables are equal to zero.

specification is comparable with past literature (e.g., Jegadeesh, 1990; Subrahmanyam, 2005) that documents a significant past return predictability effect.

We use a transformation of NIT into decile ranks to be consistent with our analysis in section 3. In other words, each stock is put into one of the ten deciles in a certain week according to its NIT value that week relative to the NIT of that same stock in the previous nine weeks, where decile 1 (10) contains stocks with the most negative (positive) NIT. We then use the decile rank of each stock on each week (the NITDecile variable) as an independent variable in the regressions.[18] Similarly, we use a transformation of turnover into decile ranks (as we do for NIT) because Gervais, Kaniel, and Mingelgrin (2001) found such a transformation of volume useful in predicting returns.

In Panel A of Table 5 we use CRSP returns to be consistent with most of the papers in the return predictability literature. In both the univariate and multivariate regressions, the coefficients on NITDecile and TurnoverDecile are positive and highly statistically significant, which is consistent with the findings in the last table, but the coefficient on past return is negative and significant, which is consistent with the past literature, but is inconsistent with the results in the prior table.[19] In the separate regressions on small, mid-cap and large stocks we observe that the significant relation between past returns and future returns is generated entirely from the smaller stocks.

The significant showing of past returns in the sample of small stocks prompted us to examine the robustness of these results to two issues: bid-ask bounce and

---

[18] For robustness, we also ran the regressions using NIT, rather then the NIT decile ranks, as the independent variable. This specification is similar in spirit to the cross-sectional robustness tests that we conducted in Section 3. The results were similar in that the mean coefficient on NIT was positive and statistically significant in all the models (univariate and multivariate).

[19] While the mean coefficient on past return is much larger than the mean coefficients on NITDecile and TurnoverDecile, the past return effect is in fact much smaller than the NIT or volume effects. To see this note that the magnitude of a typical weekly return is in the order of 10^{-2}, which means that its effect on future returns is in the order of 10^{-4} (the mean coefficient on past return in the multivariate equation is −0.0215). In contrast, the mean of the decile rank variable used for NITDecile (or TurnoverDecile) is about 5.5, which means that the effects of NIT and volume on future returns are in the order of 10^{-3}.

nonsynchronous trading.[20] To eliminate the effect of bid-ask bound we use the TAQ database to create a return series from end-of-day quote midpoints.[21] The closing TAQ midpoint may also mitigate the problem of non-synchronous trading. Since the specialist keeps a binding quote in each stock and can change the quote even when there is no trading, the quote prevailing at the close of the market presumably contains updated pricing information even if the last trade occurred long before the close.

Panel B of Table 5 presents the results of the regressions with the midquote returns. While both NIT and Turnover are strongly related to future returns in the entire sample and all subsamples, the past returns effect is weaker with midquote returns. Here, past return is not significant in the regression on the entire sample and it comes out significant only in the small cap subsample, with a significance level that is weaker than we observed for the regressions using CRSP returns.

The finding of no return reversals, even in a univariate specification, for mid-cap and large stocks seem surprising given the evidence in previous studies of short-horizon return dynamics. Since the four-year sample period we consider does not overlap with the sample periods examined in the previous studies of weekly return reversals, we use the same methodology to examine return reversals over four-year periods starting in 1964. This exercise is intended to provide some insight on whether this phenomenon has changed over time, and whether the period we study is unusual relative to the time periods considered in earlier studies.

The results in Table 6 indicate that the return reversal phenomenon has been changing. The second column of Table 6 shows a very clear trend in the estimated mean

[20] Conrad, Gultekin, and Kaul (1997) claim that a large portion of the documented weekly return reversal can be explained by bid-ask bounce. Lo and MacKinlay (1990) present a framework where non-trading induces negative serial correlation in the returns of individual stocks. While their simulations show that the impact of non-trading on short-horizon returns of individual stocks is negligible, it can still contribute to the significant coefficient that we find on past returns.

[21] Since the quality of intraday data in TAQ may not be as high as the quality of the CRSP data, if the absolute value of the difference between the TAQ return and the CRSP return is greater than 15%, we set the TAQ return to a missing value for the purpose of the regressions.

coefficients over the past decade or so since the publication of the work by Lehmann (1990) and Jegadeesh (1990) on the predictability of short-horizon returns. While the magnitude of the mean coefficient on past return fluctuates throughout the decades, it monotonically decreases from the 1988–1991 period (–0.0909) to the 2000–2003 period (–0.0229). In fact, the magnitude has been at an all-time low since 1996. The analysis of size groups shows that the decline in the magnitude and significance of the mean coefficient over the past decade can be found in stocks of all sizes. Since small stocks demonstrate a higher degree of weekly return reversal than mid-cap or large stocks, the declining trend still leaves a statistically significant mean coefficient during our sample period, 2000–2003. The smaller magnitude of reversals in larger stocks coupled with the declining trend over the past decade result in non-significant mean coefficients for the mid-cap and large groups in the most recent four-year period.

5. Potential Explanations

In this section we explore explanations for the finding of positive excess return following individual buying. The most straightforward explanation is that the individuals whose trades are executed on the NYSE have private information about the fundamentals of stocks. While plausible, we find this explanation less appealing since it is unclear how individuals, who have far fewer resources than institutions, could gain the upper hand in discovering private fundamental information and trade on it profitably in such a wide-spread fashion.[22]

Another interpretation of these results is that individuals provide liquidity to institutions that require immediacy. This explanation is consistent with both the

[22] It is possible that the individual orders that are executed on the NYSE come from relatively sophisticated, and possibly informed, individual investors. However, Jones and Lipson (2004) use NYSE proprietary order level data and find that orders coming from individuals have smaller permanent price impacts relative to institutional orders, suggesting that individuals have less private information than institutions about stocks' fundamentals.

contrarian patterns we found and the positive excess return after buying. What may be happening is that individuals sell shares when the buying pressure from institutions pushes prices up and buy shares when the selling pressure from institutions pushes prices down. We do not claim that individuals provide liquidity by trading actively like dealers making two-sided markets. Rather, it could be that when institutions trade large positions in a certain direction and start moving prices, individuals end up taking the other side of these positions.

We know from models of risk-averse liquidity provision like Grossman and Miller (1988) and Campbell, Grossman, and Wang (1993) that investors who require immediacy (e.g., institutions) must offer price concessions to induce other risk-averse investors, in this case individuals, to take the other side of their trades. These price concessions result in subsequent return reversals because the future cash flows of the stock do not change, and these could be the short-horizon excess returns we find following concentrated individual buying. Our evidence is therefore consistent with the hypothesis that individuals who trade on the NYSE tend to at least implicitly react to the liquidity needs of institutions, and at least in the short run, earn abnormal returns by exploiting their counterparties demand for immediacy.[23] While the effect should be symmetric in the sense that liquidity provision could be profitable for individuals both when they buy and when they sell, the information content of institutional trading may also affect the pattern of returns we observe. Institutional buying activity is more likely to be motivated by information then their selling activity (see Saar (2001) and references

[23] Campbell, Ramadorai, and Vuolteenaho (2004) use institutional 13-F filings and trade information from TAQ to identify institutional trading. Their results suggest that institutions demand rather than provide liquidity, and seem particularly likely to demand liquidity when they sell stocks, consistent with the excess return we find when individuals buy and provide liquidity to the institutions. Andrade, Chang, and Seasholes (2005) find that Taiwanese individuals lose on average when demanding liquidity. Barber, Lee, Liu, and Odean (2004a) find that while Taiwanese individuals on average lose when trading, they gain from liquidity providing trades at short horizons (10 and 25 days).

therein), which may explain why individuals fail to profit when they take the other side of institutional buys.

If the excess returns individuals earn when buying represent compensation for providing liquidity to institutional sellers, we should expect to find higher compensation (larger excess returns) when individuals buy less liquid stocks. To test this hypothesis we use the percentage effective spread (the distance from the transaction price to the quote midpoint divided by the quote midpoint) as a proxy for the liquidity of a stock.[24] The larger the effective spread, the greater the price movement on trades and therefore the less liquid the stock. In Table 7 we sort stocks each week according to the average percentage effective spread and put them into three groups: small, medium, and large.[25] We then form the intense buying portfolio of individuals (decile 10) separately for each spread group. We observe that individuals realize greater excess returns when buying less liquid stocks: 0.72% in the 20 days following portfolio formation in the small spread group, 1.42% in the medium spread group, and 2.46% in the large spread group. These results are consistent with the hypothesis that individuals generate excess returns by accommodating the liquidity needs of institutions.

In light of the literature on the relation between liquidity and excepted returns, one could argue that sorting on spread is basically sorting on a stock characteristic that could be priced. As such, the result of higher excess returns following purchases by individuals in the high effective spread group could mean that individuals buy riskier stocks. We therefore computed another measure of excess return by subtracting the return on a portfolio of stocks with similar spreads. For example, the excess return on the intense buying portfolio in the small spread group is its raw return minus the return on the entire small spread group. The magnitudes of the excess returns 20 days after intense

[24] The percentage effective spread measure is constructed using the TAQ database.

[25] Our weekly sorting into spread groups has the advantage that a stock may be classified not just according to its average liquidity properties but also according to the state of liquidity of the stock on that week.

buying by individuals are smaller using this definition of excess returns, but are still positive and highly significant. More importantly, they maintain the pattern whereby individuals earn higher excess returns when buying less liquid stocks: 0.42% in the 20 days following portfolio formation in the small spread group compared with 1.40% in the large spread group.

5.2 Individual Investors and Behavioral Finance

Individual investors play a central role in behavioral finance. In particular, the literature often associates individuals with "noise" traders who follow positive-feedback strategies (De Long, Shleifer, Summers and Waldmann (1990b)), lose money on average (Black (1986)), and create excess volatility (De Long, Shleifer, Summers and Waldmann (1990a) and Shleifer and Summers (1990)). De Long, Shleifer, Summers and Waldmann (1990a) suggest that the activities of these noise traders make stocks more volatile or riskier. The excess return we find following increases in individual investment could, therefore, be compensation for this added risk.

To examine in more detail volatility patterns around intense trading by individuals we follow the same basic procedures that generated the numbers in Table 2, but calculate volatility rather than mean returns. We compute for each stock in each of the four portfolios the standard deviation of daily returns in 9-day windows centered on k = –20, –15, –10, –5, 0, +5, +10, +15, and +20 days (where day 0 is the middle of the formation week). Since we are interested in abnormal volatility around intense individual trading activity, we subtract from these numbers the "normal" 9-day return standard deviation (which we compute as the average of daily return standard deviations on all non-overlapping 9-day windows in the sample period). Table 8, which presents the mean of these abnormal volatility measures in each NIT portfolio, tells us how volatility of returns evolves around the trading of individuals.

A clear pattern emerges from the table: volatility increases prior to intense individual activity and subsequently decreases. Take for example the volatility of returns around intense individual selling (first line of the table, going across the columns): it is –0.0001 below average volatility at k = –20, then increases to 0.0012 above average volatility at k = –5, reaches 0.0018 at k = 0, and then decreases to –0.0008 by k = +20. The next two columns test the increase of volatility from k = –20 to k = 0, which is 0.0020 and statistically significant, and the decrease of –0.0027 from k = 0 to k = +20, again statistically significant. The last column of the table tests the more "permanent" change in volatility, from k = –20 to k = +20, and finds no significant change. An even greater increase in volatility (0.0033) is observed from –20 to 0 before intense buying activity (decile 10), and most of it is subsequently reversed (–0.0023) from 0 to +20. Therefore, it seems that the increase in volatility we observe is temporary in nature and disappears after the abnormal trading period, making the limits-to-arbitrage explanation of the excess return we document less attractive.

Volatility is not the only measure of risk one could examine. Theoretical models that postulated increased risk due to the activity of noise traders typically featured a single risky asset, and therefore used volatility as the risk measure. In a multi-security economy, other risk measures that take into account the correlations across stocks may be appropriate. Our use of volatility seems reasonable considering the short windows we examine around individual trading. However, a natural question to ask is whether the actions of individual investors have "systematic" effects in the sense that they affect all stocks at the same time. We therefore looked at whether the dynamic relations we identified for stocks (the contrarian pattern and predictability) exist between the value-weighted market return and a value-weighted measure of net individual trading. We found no statistically significant patterns, suggesting that the behavior of individuals may not be correlated across stocks.

The lack of dynamic patterns at the market portfolio level prompted us to carry out additional analysis. The importance of this issue rests in part on the suggestion of the behavioral finance literature that, if indeed individual investors are “noise” traders, a systematic variation in their behavior would affect expected returns. This argument is succinctly made by Lee, Shleifer, and Thaler (1991): “If different noise traders traded randomly across assets, the risk their sentiment would create would be diversifiable, just as the idiosyncratic fundamental risk is diversifiable in conventional pricing models. However, if fluctuations in the same noise trader sentiment affect many assets and are correlated across noise traders, then the risk that these fluctuations create cannot be diversified. Like fundamental risk, noise trader risk will be priced in equilibrium.”

To examine this question, we conduct a principal component analysis of the daily net individual trading measure and look at the percentage of variance of NIT that is explained by the first ten principal components. We construct 1,000 random sub-samples of 180 stocks each from among the stocks that have a complete set of daily returns, and look at the mean and standard deviation of the percentage of variance across the 1,000 random sub-samples.[26] We use simulations to generate principal components for independent random matrices, and use these as a benchmark for evaluating the percentage of variance explained by the principal components in the real data (details of the methodology are provided in the Appendix).[27]

Panel A of Table 9 shows the results of the principal component analysis of the net individual trading measure and also of daily returns. The daily return analysis is

---

[26] We chose 180 stocks as the size of a sub-sample because it is approximately a tenth of the number of stocks, and is therefore roughly comparable to the number of stocks in a size decile. We present the principal component analysis of size deciles later in this section.

[27] We use simulations to create a benchmark because any arbitrary decision on the size of the sub-samples affects the estimates. For example, the percentage of the variance explained by the first principal component is at least 1% in a 100-stock sub-sample because each stock contributes one unit of variance to the analysis. The simulated benchmark helps us determine whether the structure observed in the data is really there, as opposed to being generated by our particular choices or simply by chance (see Freedman and Lane (1983)).

shown to provide a sense of the magnitude of co-movement observed in the cross section of stocks. For example, 21.25% of the daily variation in returns of stocks in our sample is explained by the first five principal components. However, the third line of the panel shows that the percentage of variance explained by the first five principal components of the simulated independent data is 5.33%, and therefore the difference between these two numbers, roughly 15.92%, is a better measure of the structure in the real data. The analysis of NIT reveals very little evidence of correlated actions of individual investors across stocks. Indeed, the first (and largest) principal component of NIT explains only 1.70% of the variance (adjusted using the simulated data) compared with 12.07% for returns.

Since some papers (e.g., Lee, Shleifer, and Thaler (1991); Kumar and Lee (2002)) claim that “noise” trading of individuals is potentially stronger in small stocks, we sort the sample into ten deciles according to each stock’s average market capitalization over the sample period. Each decile contains less than 200 stocks, and therefore we do not need to draw random sub-samples to analyze the real data. Nonetheless, we still need to adjust the estimates using simulations of independent, normally-distributed data (details are provided in the Appendix). Panel B of Table 9 presents the results. Contrary to what one might have expected based on the above papers, the percentage of the NIT variance explained by the first five principal components (adjusted using the simulations) is lower for small stocks (3.13% for decile 1) than for large stocks (10.31% for decile 10).

Our findings contrast with those of Kumar and Lee (2002) who examine correlations among order flow imbalances of stocks traded by clients of a single U.S. discount broker. They find that their measure of order flow imbalance is moderately correlated across stocks, concluding that there is evidence of a systematic component in retail investor trading.[28] Our analysis shows little by means of correlated actions of individuals who trade on the NYSE. These results may suggest that finding a systematic influence of individual investors on expected returns may be difficult.

6. Conclusions

Our analysis of the trading of individual investors on the NYSE provides results that seem surprising in light of the extant literature. First, we document that individual buying predicts positive excess returns, which contrasts with Odean (1999) who finds that individual investors perform poorly. Odean uses data on the trading of individuals through one discount broker and analyzes returns over a longer horizon. One explanation for the conflicting results is that our focus on short-horizon behavior picks up the returns individuals earn from liquidity provision, and possibly obscures the individuals' informational disadvantage relative to institutions that is likely to show up at longer horizons. Alternatively, it may be the case that the individual investors who execute their trades on the NYSE differ from those in the Odean data set, who trade with a discount broker that sends only a portion of its order flow to execute on the NYSE. It could be that individuals whose orders execute on the NYSE are more sophisticated and better informed than those who invest with this discount broker.

Second, we find a very significant relation between returns and the past order flow imbalances of individual investors, which seem to contrast with Subrahmanyam (2005) who finds that net trade imbalances in general do not seem to predict returns. Perhaps the net order flow of individuals that we consider is a better measure of the demand for liquidity than the net trade imbalances measure of Subrahmanyam, who uses the Lee and Ready (1991) algorithm to indirectly infer whether trades are initiated by

[28] Barber, Odean, and Zhu (2003) do not focus on the correlation in individual trading across many stocks, but they show that clients of two different brokers tend to trade the same stocks at the same time. They also show temporal persistence in that if individuals are buying a stock one month they are more likely to be buying it the following month as well.

buyers or sellers. The Lee and Ready (1991) algorithm establishes which party to a trade used a market order (by comparing the transaction price to the quote midpoint), and classifies that party as a liquidity demander. In contrast, we classify individuals as liquidity providers regardless of how they execute their orders, which allows for very different interpretations of the data. For example, institutions that want to move large positions might use dynamic limit order strategies and their demand for immediacy might be accommodated by contrarian individuals who would offer their shares with market orders. In this example, the Lee and Ready algorithm would classify the institutions as liquidity providers and the individuals as liquidity demanders, while we would make the opposite classification.

In general, the contrarian behavior we document of individual investors on the NYSE seems important for understanding short-horizon return predictability. The underlying reason for why individuals act in such a way is not well understood, and one can find arguments in the behavioral literature supporting both contrarian tendencies (e.g., loss aversion in Odean (1998)) as well as a tendency to buy winners (e.g., positive feedback trading in De Long, Shleifer, Summers, and Waldmann (1990b); attribution bias in Daniel, Hirshleifer, and Subrahmanyam (1998)). Whatever the reason, the contrarian choices of individuals lead them to implicitly provide liquidity to other market participants who demand immediacy.

In theory, the extent to which price reversals are observed depends on the risk aversion of the liquidity providers and the amount of capital available for liquidity provision. Suppose that individual investors are the only ones providing liquidity in the market. If contrarian individual investors are in some sense too active relative to the demand for immediacy, there will be an excess supply of liquidity in the market. If this is the case, then the contrarian individuals who implicitly provide liquidity will tend to lose money by trading with more informed investors at unfavorable terms. On the other hand, if there are too few contrarian investors relative to the demand for immediacy, then those individuals who implicitly provide liquidity will realize excess returns.

In reality, liquidity is provided by professional traders (e.g., specialists) as well as contrarian individuals. One would expect that the amount of capital that these professionals devote to their market making activity is determined by the aggregate demand for liquidity as well as the amount of liquidity implicitly supplied by individual investors. In equilibrium, these professional traders will supply liquidity up to the point where their trading profits just cover their costs. Over the past 20 years institutional trading has increased and the importance of individual investors has declined, suggesting that there may have been a positive shift in the demand for immediacy and a negative shift in the supply of liquidity. If this is indeed the case, and if the amount of capital devoted to liquidity provision is slow to adjust, then this shift could create a potential short-term profit opportunity for those traders that provide liquidity.

The evidence in this paper is consistent with the view that a short-term liquidity provider could have generated profits by mimicking the trades of individual investors during our sample period. There is also anecdotal evidence suggesting that in response to this opportunity, there has been an increase in the number of professional investors who specialize in short-term contrarian trading strategies, and thus indirectly provide such services.[29] Indeed, the presence of these traders may be responsible for the reduction we document in the return reversals first observed by Jegadeesh (1990) and Lehmann (1990).

---

[29] For example, Automated Trading Desk (ATD) is one of the firms that pioneered the use of computerized expert systems applied to liquidity provision. While currently they also work on an agency basis for institutional investors, their core competency has been proprietary limit-order strategies that provide liquidity to the market and profit from short-term price movements. ATD trading in 2003 accounted for about 5% of Nasdaq volume and more than 2% of the volume in listed stocks. It is also interesting to note that there has been a tremendous drive for consolidation among NYSE specialist firms in the past 15 years. The number of specialist firms trading NYSE common stocks declined from 52 in 1989 to seven in 2004. One argument made to support these consolidations was that liquidity will be enhanced by having better-capitalized market making firms.

Then why don't the strategies implemented by these short-term traders eliminate the excess returns associated with the trading of individuals? This is a difficult question that clearly warrants additional research. The most natural explanation is that these high frequency strategies are quite costly to implement, so we expect to observe high pre-transaction costs returns. It is also possible that the remaining return is needed to compensate those firms for the risk associated with undertaking the liquidity-supplying trading strategies. Moreover, it may be the case that mechanical strategies are unable to implement the strategies implicitly implemented by individual investors. While the trades of all market participants (including individuals) are public information, the Account Type information identifying the orders of individual investors cannot be used to implement a trading strategy in real time because it is not publicly available (it is not available even to the specialists who oversee trading on the NYSE floor). Therefore, institutions could not simply use NIT to formulate their strategies, but rather would have to base a strategy on a proxy for net individual trading, increasing the risk associated with such a strategy.

The evidence we present seems to suggest that understanding short-horizon return predictability requires understanding the implicit liquidity provision of individuals as well as the explicit liquidity provision of professional investors. In particular, liquidity provision may be viewed as the interplay between different types of investors who populate the market. At the very least, our work suggests that understanding the behavior of one investor type, individuals, holds some promise for explaining observed return patterns.

Appendix

Our sample consists of 2,034 stocks and 1,004 trading days. For the analysis in Panel A of Table 9 we first construct 1,000 random sub-samples of 180 stocks each from among the stocks that have a complete set of daily returns. We perform a principal component analysis using the Principal Axis method for each sub sample, and then compute the mean and standard deviation across the 1,000 sub-samples of the percentage of the variance explained by the first ten principal components. These summary statistics are reported in the panel as “Real Mean” and “Real Std”.

The adjustment using simulations is done as follows. We construct another set of 1,000 random sub-samples of 180 stocks each. We calculate the mean and standard deviation of the variable analyzed (say the net trading of individual investors) for each stock in a sub-sample. We then generate an artificial time-series for each stock drawn from a normal distribution with the same mean and standard deviation. We conduct a principal component analysis on the 180 independent time-series and note the percentage of the variance explained by the first ten principal components. We repeat this process for each sub-sample ten times and average the percentage of the variance explained by each principal component in order to get estimates that are less noisy. We end up with 1,000 estimates for sub-samples of simulated, independent data (reported in the table as Sim. Mean), and look at the differences (Diff.) between the real and simulated means.

The results demonstrate the importance of considering a simulated benchmark. For example, the first principal component in Panel A explains on average 1.11% of the variance of the simulated, independent data. The fact that the first eigenvalue explains considerably more than 1/180 of the variance of a 180-stock sample of randomly generated returns is not entirely surprising. It is well known that the distribution of the spacing *x* between adjacent eigenvalues of a random matrix whose elements are i.i.d Gaussian is closely approximated by the “Wigner surmise” $P(x) \approx Axe^{-Bx^2}$ (see, for example, Porter (1965)). Furthermore, numerical experiments have shown that the surmise holds for a wide range of distributions (e.g., Lehman (2001)). Therefore, the use of a simulated benchmark aids in evaluating the strength of the structure found in the real data.

For the analysis in Panel B of Table 9 we sort the sample into ten deciles according to each stock's average market capitalization over the sample period. We perform a principal component analysis on each decile separately. To create the simulated benchmark for these estimates we start by using the mean and standard deviation of each stock to generate 500 artificial time-series drawn from the normal distribution. We then use these simulated data to run 500 separate principal components analyses for each decile, and we report in the table the difference between the estimate of the percentage of variance in the real data and the mean of the 500 estimates of the simulated data.

References

Andrade, S.C., Chang, C., Seasholes, M.S., 2005. Trading Shocks, Asset Prices, and the Limits of Arbitrage. Working paper, U.C. Berkeley.

Bailey, W., Kumar, A., and Ng, D., 2004. Venturing abroad: Foreign investments of U.S individual investors. Working paper, Cornell University.

Barber, B. M., Lee, Y., Liu, Y., and Odean, T., 2004a. Who gains from trade? Evidence from Taiwan. Working paper, University of California, Davis.

Barber, B. M., Lee, Y., Liu, Y., and Odean, T., 2004b. Do individual day traders make money? Evidence from Taiwan. Working paper, University of California, Davis.

Barber, B. M., and Odean, T., and Zhu, N., 2003. Systematic noise. Working paper, University of California, Davis.

Black, F., 1986. Noise. *Journal of Finance* 41, 529-543.

Choe, H., Kho, B., and Stulz, R., 1999. Do foreign investors destabilize stock markets? The Korean experience in 1997. *Journal of Financial Economics* 54, 227-264.

Campbell, J. Y., Grossman, S. J., and Wang, J., 1993. Trading volume and serial correlation in stock returns. *Quarterly Journal of Economics* 108, 905-939.

Campbell, J. Y., Ramadorai, T., and Vuolteenaho, T. O., 2004. Caught on tape: Predicting institutional ownership with order flow. Working paper, Harvard University.

Conrad, J. S., Gultekin, M. N., and Kaul, G., 1997. Predictability of short-term contrarian strategies: Implications for market efficiency. *Journal of Business & Economic Statistics* 15, 379-386.

Conrad, J. S., Hameed, A., and Niden, C., 1994. Volume and autocovariances in short-horizon individual security returns. *Journal of Finance* 49, 1305-1329.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A., 1998. Investor Psychology and Security Market Under- and Over-reactions. *Journal of Finance* 53, 1839-1886.

De Long, J. B., Shleifer A., Summers, L. H., and Waldmann, R. J., 1990a. Noise trader risk in financial markets. *Journal of Political Economy* 98, 703-738.

De Long, J. B., Shleifer A., Summers, L. H., and Waldmann, R. J., 1990b. Positive feedback investment strategies and destabilizing rational speculation. *Journal of Finance* 45, 379-395.

Fama, E., and MacBeth, J., 1973. Risk, return and equilibrium: Empirical tests. *Journal of Political Economy* 71, 607-636.

Freedman, D. A., and Lane, D., 1983. Significance testing in a nonstochastic setting. In *A Festschrift for Erich L. Lehmann*. Bickel, R. J., Doksum, K. A., and Hodges, J. L., editors. Wadsworth, Inc., Belmont, California.

Gervais, S., Kaniel, R., and Mingelgrin, D. H., 2001. The high-volume return premium. *Journal of Finance* 56, 877-919.

Goetzmann, W., and Massa, M., 2002. Daily momentum and contrarian behavior of index fund investors. *Journal of Financial and Quantitative Analysis* 37, 375-390.

Griffin, J. M., Harris, J., and Topaloglu, S., 2003. The dynamics of institutional and individual trading. *Journal of Finance* 58, 2285-2320.

Grinblatt, M., and Keloharju, M., 2000. The investment behavior and performance of various investor types: A study of Finland's unique data set. *Journal of Financial Economics* 55, 43-67.

Grinblatt, M., Titman, S., and Wermers, R., 1995. Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *American Economic Review* 85, 1088-1105.

Grossman, S. J., and Miller, M. H., 1988. Liquidity and market structure. *Journal of Finance* 43, 617-633.

Jackson, A., 2003. The aggregate behavior of individual investors. Working paper, London Business School.

Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. *Journal of Finance* 45, 881-898.

Jegadeesh, N., and Titman, S., 1995. Short-horizon return reversal and the bid-ask spread. *Journal of Financial Intermediation* 4, 116-132.

Jones, C. M., and Lipson, M. L., 2004. Are retail orders different? Working paper, Columbia University.

Kumar, A., and Lee, C. M. C., 2002. Individual investor sentiment and comovement in small stock returns. Working paper, Department of Economics, Cornell University.

Kyle, A. S., 1985. Continuous auctions and insider trading. *Econometrica* 53, 1315-1336.

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

Lee, C. M. C., Shleifer, A., and Thaler, R. H., 1991. Investor sentiment and the closed-end fund puzzle. *Journal of Finance* 46, 75-109.

Lehmann, B. N., 1990. Fads, martingales, and market efficiency. *Quarterly Journal of Economics* 105, 1-28.

Lehman, R. C., 2001. First-order spacing of random matrix eigenvalues. Department of Mathematics mimeo, Princeton University.

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

Lo, A. W., and MacKinlay, A. C., 1990. An econometric analysis of nonsychronous trading. *Journal of Econometrics* 45, 181-211.

Mase, B., 1999. The predictability of short-horizon stock returns. *European Finance Review* 3, 161-173.

Merton, R. C., 1987. A simple model of capital market equilibrium with incomplete information. *Journal of Finance* 42, 483-510.

Nofsinger, J. R., and Sias, R. W., 1999. Herding and feedback trading by institutional and individual investors. *Journal of Finance* 54, 2263-2295.

Odean, T., 1998. Are investors reluctant to realize their losses? *Journal of Finance* 53, 1775-1798.

Odean, T., 1999. Do investors trade too much? *American Economic Review* 89, 1279-1298.

Porter, C., 1965. *Statistical Theories of Spectral Fluctuations.* Academic Press, New York and London.

Richards, A., 2005. Big fish in small ponds: The trading behavior of foreign investors in Asian emerging equity markets. *Journal of Financial and Quantitative Analysis* 40, 1-27.

Saar, G., 2001. Price impact asymmetry of block trades: An institutional trading explanation. *Review of Financial Studies* 14, 1153-1181.

San, G., 2004. Who gained more from the late 1990s bubble — individuals or institutions? Working paper, Tel Aviv University.

Sias, R. W., 2003. Reconcilable differences: Momentum trading by institutions. Working paper, Washington State University.

Sias, R. W., Starks, L., and Titman, S., 2003. The price impact of institutional trading. Working paper, Washington State University and The University of Texas.

Shleifer, A., and Summers, L. H., 1990. The noise trader approach to finance. *Journal of Economic Perspectives* 4, 19-23.

Stoll, H., 1978. The supply of dealer services in securities markets. *Journal of Finance* 33, 1133-1151.

Subrahmanyam, A., 2005. Distinguishing between rationales for short-horizon predictability of stock returns. *Financial Review* 40, 11-35.

Wermers, R., 1999. Mutual fund trading and the impact on stock prices. *Journal of Finance* 54, 581-622.

Table 1
Summary Statistics

The sample of stocks for the study consists of all common, domestic stocks that were traded on the NYSE at any time between January 1, 2000 and December 31, 2003 with records in the CRSP database. We use ticker symbol and CUSIP to match the stocks to a special dataset containing daily aggregated buying and selling volume of individuals that was provided to us by the NYSE. There are 2,034 stocks in our sample. From the CRSP database, AvgCap is the average monthly market capitalization of a stock over the sample period; AvgPrc is the average daily closing price; AvgTurn is the average weekly turnover (number of shares traded divided by the number of shares outstanding); and StdRet is the standard deviation of weekly returns. From the NYSE dataset we report the weekly Dollar Volume, defined as the sum of executed buy and sell orders, and the Executed Order Size in shares of individual investors. We sort the stocks by market capitalization into ten deciles, and form three size groups: small stocks (deciles 1, 2, 3, and 4), mid-cap stocks (deciles 5, 6, and 7), and large stocks (deciles 8, 9, and 10). The summary statistics are presented for the entire sample and separately for the three size groups.

		AvgCap (in million $)	**AvgPrc** (in $)	**AvgTurn** (in %)	**StdRet** (in %)	**Individuals Dollar Volume** (1000s $)	**Individuals Executed Order Size** (shares)
All stocks	Mean	5,303.2	59.80	2.52	0.0700	4,304.1	770.9
	Median	943.7	21.98	2.05	0.0589	1,131.0	644.6
Small stocks	Mean	317.1	13.57	2.37	0.0836	716.3	904.7
	Median	308.7	11.60	1.63	0.0697	377.1	722.5
Mid-Cap stocks	Mean	1,311.8	25.94	3.22	0.0667	2,144.2	711.5
	Median	1,230.3	23.94	2.51	0.0591	1,417.1	613.5
Large stocks	Mean	14,054.0	136.87	3.10	0.0598	11,147.6	675.0
	Median	5,018.0	37.15	2.56	0.0532	4,991.8	618.2

Table 2
Returns around Individual Trading

This table presents analysis of market-adjusted returns around intense buying and selling activity of individuals as given by the net individual trading measure (NIT). For each week in the sample period, we use the previous nine weeks to form NIT deciles. Each stock is put into one of ten deciles according to the value of NIT in the current week relative to its value in the previous nine weeks. Decile 1 contains the stocks with the most intense selling (negative NIT) while decile 10 contains the stocks with the most intense buying (positive NIT). We present the results for four portfolios: (i) decile 1, (ii) deciles 1 and 2, (iii) deciles 9 and 10, and (iv) decile 10. Let k be the number of days prior to or following portfolio formation each week. We calculate eight cumulative return numbers for each of the stocks in a portfolio: CR(t-k,t-1) where $k \in \{20, 15, 10, 5\}$ days and t is the first day of the formation week, and CR(t+1,t+k) where $k \in \{5, 10, 15, 20\}$ days and t is the last day of the week. The return on each portfolio is then adjusted by subtracting the return on a market proxy (the value-weighted portfolio of all stocks in the sample). We present the time-series mean and t-statistic for each market-adjusted cumulative return measure. ** indicates significance at the 1% level and * indicates significance at the 5% level (both against a two-sided alternative).

Portfolio		**k=–20**	**k=–15**	**k=–10**	**k=–5**	**k=+5**	**k=+10**	**k=+15**	**k=+20**
Intense Selling	Mean	0.0351**	0.0325**	0.0265**	0.0168**	-0.0006	0.0002	0.0013	0.0036
(decile 1)	t-stat.	(15.04)	(16.55)	(17.64)	(16.80)	(-0.64)	(0.11)	(0.70)	(1.64)
Selling	Mean	0.0338**	0.0306**	0.0248**	0.0152**	-0.0007	0.0002	0.0013	0.0033
(deciles 1&2)	t-stat.	(15.68)	(16.99)	(18.21)	(17.14)	(-0.77)	(0.14)	(0.71)	(1.60)
Buying	Mean	-0.0215**	-0.0209**	-0.0175**	-0.0118**	0.0031**	0.0064**	0.0096**	0.0134**
(deciles 9&10)	t-stat.	(-9.92)	(-11.32)	(-11.48)	(-11.39)	(3.45)	(4.69)	(5.62)	(6.57)
Intense Buying	Mean	-0.0212**	-0.0211**	-0.0183**	-0.0125**	0.0038**	0.0070**	0.0107**	0.0149**
(decile 10)	t-stat.	(-9.52)	(-11.13)	(-11.75)	(-11.44)	(3.94)	(4.95)	(6.12)	(7.16)

Table 3
Return Predictability: Portfolio Sorting Approach

This table presents analysis of weekly return predictability conditional on the previous week's return (Panel A) or turnover (Panel B) and the net individual trading measure (NIT). For each week in the sample period, we use the previous nine weeks to form NIT quintiles. Each stock is put into one of the five quintiles according to the value of NIT in the current week relative to its value in the previous nine weeks (where quintile 1 has stocks with more negative NIT, or more selling, and quintile 5 has stocks with more positive NIT, or more buying). In Panel A, each week in the sample period stocks are also sorted on return and put into five quintiles (quintile 1 has stocks with the most negative return and quintile 5 has stocks with the most positive return). We then form 25 portfolios as the intersection of the five return quintiles and five NIT quintiles, and compute for each portfolio the market-adjusted return in the week following the formation week. We present the time-series mean return for each of the 25 portfolios sorted by return and net individual trading. The last two rows of the panel give the payoff to the strategy of buying NIT quintile 5 and selling NIT quintile 1, and the last two columns of the panel give the payoff to the strategy of buying return quintile 5 and selling return quintile 1. Panel B present similar analysis except that we sort on past turnover (rather than past return) and past NIT. The construction of the 25 portfolios is analogous to the one in Panel A, and the last two columns of the panel give the payoff to the strategy of buying turnover quintile 5 and selling turnover quintile 1. ** indicates significance at the 1% level and * indicates significance at the 5% level (both against a two-sided alternative). The t-statistic is computed using the Newey-West correction.

Panel A: Weekly Return Predictability using Past Return and NIT

		Return(t)						
		Q1 (<0)	Q2	Q3	Q4	Q5 (>0)	Q5–Q1	t-statistic
NIT(t)	Q1 (<0)	-0.0015	-0.0012	0.0000	0.0000	-0.0009	0.0006	(0.35)
	Q2	-0.0003	-0.0005	-0.0004	-0.0006	-0.0019	-0.0016	(-0.90)
	Q3	0.0006	-0.0010	0.0000	-0.0003	-0.0006	-0.0011	(-0.63)
	Q4	0.0031	0.0004	0.0006	0.0012	0.0008	-0.0023	(-1.54)
	Q5 (>0)	0.0045	0.0015	0.0024	0.0025	0.0029	-0.0016	(-1.00)
	Q5–Q1	0.0060**	0.0027**	0.0024**	0.0025**	0.0038**		
	t-statistic	(4.87)	(2.92)	(3.38)	(3.26)	(3.43)		

Panel B: Weekly Return Predictability using Past Turnover and NIT

		Turnover(t)						
		Q1 (low)	Q2	Q3	Q4	Q5 (high)	Q5–Q1	t-statistic
NIT(t)	Q1 (<0)	-0.0026	-0.0025	-0.0014	0.0002	0.0011	0.0037**	(2.78)
	Q2	-0.0033	-0.0028	-0.0009	0.0003	0.0022	0.0055**	(3.94)
	Q3	-0.0034	-0.0015	-0.0005	0.0007	0.0019	0.0054**	(3.89)
	Q4	-0.0019	-0.0005	0.0019	0.0035	0.0051	0.0070**	(5.41)
	Q5 (>0)	0.0013	0.0010	0.0015	0.0036	0.0053	0.0040**	(2.98)
	Q5–Q1	0.0039**	0.0034**	0.0029**	0.0035**	0.0042**		
	t-statistic	(2.95)	(3.58)	(3.44)	(3.20)	(3.64)		

Table 4
Fama-MacBeth Approach with Dummy Variables for Past Return Quintiles

This table presents a regression analysis of short-horizon (weekly) return predictability. The dependent variable is weekly return (from CRSP), Return(t+1), and the independent variables are an intercept, and a three sets of dummy variables. The first set is formed by sorting Return(t) into quintiles and using four dummy variables for quintiles 1 through 4. The second and third sets are quintile dummy variables for NIT(t) and Turnover(t). Construction of the net individual trading measure (NIT) is described in Section 2. We implement a Fama-MacBeth methodology for the regressions: (i) a cross-sectional regression is performed for each week in the sample period, and (ii) test statistics are based on the time-series of the coefficient estimates. We present the mean coefficient from the weekly regressions, and use the Newey-West correction for the standard errors to compute the t-statistics. We present results separately for all stocks and for three size groups. ** indicates significance at the 1% level and * indicates significance at the 5% level (both against a two-sided alternative).

Variable	All Stocks	Small Stocks	Mid-Cap Stocks	Large Stocks
Intercept	0.0080**	0.0077**	0.0091**	0.0069**
(t-statistic)	(3.97)	(3.33)	(4.40)	(3.37)
NIT Q1	-0.0037**	-0.0041**	-0.0037**	-0.0031**
(t-statistic)	(-6.78)	(-4.86)	(-5.94)	(-4.35)
NIT Q2	-0.0032**	-0.0042**	-0.0026**	-0.0025**
(t-statistic)	(-5.65)	(-4.31)	(-4.22)	(-3.69)
NIT Q3	-0.0026**	-0.0037**	-0.0024**	-0.0016*
(t-statistic)	(-4.74)	(-4.37)	(-3.58)	(-2.49)
NIT Q4	-0.0012**	-0.0011	-0.0019**	-0.0008
(t-statistic)	(-2.63)	(-1.37)	(-3.01)	(-1.29)
Turnover Q1	-0.0054**	-0.0069**	-0.0027**	-0.0015
(t-statistic)	(-7.12)	(-6.24)	(-3.16)	(-1.50)
Turnover Q2	-0.0047**	-0.0066**	-0.0024**	-0.0018*
(t-statistic)	(-6.87)	(-6.51)	(-3.14)	(-2.15)
Turnover Q3	-0.0031**	-0.0038**	-0.0015*	-0.0020*
(t-statistic)	(-5.09)	(-4.02)	(-2.12)	(-2.57)
Turnover Q4	-0.0018**	-0.0020*	-0.0014*	-0.0008
(t-statistic)	(-3.30)	(-2.14)	(-2.10)	(-1.17)
Return Q1	0.0012	0.0030*	-0.0010	-0.0004
(t-statistic)	(1.01)	(2.57)	(-0.68)	(-0.23)
Return Q2	0.0000	0.0006	-0.0008	-0.0016
(t-statistic)	(0.00)	(0.54)	(-0.77)	(-1.27)
Return Q3	0.0002	0.0013	-0.0009	-0.0013
(t-statistic)	(0.22)	(1.49)	(-0.92)	(-1.23)
Return Q4	-0.0003	0.0013	-0.0023*	-0.0008
(t-statistic)	(-0.46)	(1.62)	(-2.57)	(-1.01)

Table 5
Fama-MacBeth Approach with Continuous Past Return Variable

This table presents a regression analysis of short-horizon (weekly) return predictability. The dependent variable is weekly return (from CRSP), Return(t+1), and the independent variables are an intercept, Return(t), NITDecile(t), and TurnoverDecile(t). The TurnoverDecile variable is from Gervais, Kaniel, and Mingelgrin (2001). It classifies the weekly turnover (number of shares traded over the number of shares outstanding) into ten deciles by comparing it to the same stock's turnover in the previous nine weeks. The net individual trading (NIT) measure is described in section 2, and the NITDecile variable is constructed in a similar fashion to TurnoverDecile. We implement a Fama-MacBeth methodology for the regressions: (i) a cross-sectional regression is performed for each week in the sample period, and (ii) test statistics are based on the time-series of the coefficient estimates. We present the mean coefficient from the weekly regressions, and use the Newey-West correction for the standard errors to compute the t-statistics. In Panel A we use CRSP returns, while in Panel B we compute returns using end-of-day quote midpoints from the TAQ database. ** indicates significance at the 1% level and * indicates significance at the 5% level (both against a two-sided alternative).

Panel A: CRSP Returns

Size Groups	Intercept (t-statistic)	Return(t) (t-statistic)	NITDecile(t) (t-statistic)	Turnover Decile(t) (t-statistic)
All Stocks	0.0035	-0.0225**		
	(1.93)	(-3.10)		
	0.0008		0.0005**	
	(0.45)		(7.28)	
	-0.0006			0.0007**
	(-0.31)			(6.87)
	-0.0030	-0.0215**	0.0004**	0.0007**
	(-1.53)	(-2.97)	(6.65)	(7.71)
Small Stocks	0.0028	-0.0312**		
	(1.32)	(-3.93)		
	-0.0008		0.0006**	
	(-0.39)		(6.50)	
	-0.0028			0.0010**
	(-1.28)			(7.97)
	-0.0057*	-0.0316**	0.0005**	0.0010**
	(-2.56)	(-3.96)	(5.64)	(8.39)
Mid-Cap Stocks	0.0042*	-0.0053		
	(2.34)	(-0.54)		
	0.0024		0.0004**	
	(1.30)		(5.36)	
	0.0028			0.0003**
	(1.40)			(3.18)
	0.0004	-0.0033	0.0004**	0.0003**
	(0.19)	(-0.34)	(5.24)	(3.34)
Large Stocks	0.0035*	-0.0203		
	(1.99)	(-1.68)		
	0.0015		0.0004**	
	(0.82)		(4.33)	
	0.0016			0.0003**
	(0.77)			(2.77)
	-0.0004	-0.0158	0.0003**	0.0003**
	(-0.17)	(-1.30)	(4.24)	(3.35)

Panel B: Midquote Returns from the TAQ Database

Size Groups	Intercept (t-statistic)	Return(t) (t-statistic)	NITDecile(t) (t-statistic)	Turnover Decile(t) (t-statistic)
All Stocks	0.0031	-0.0132		
	(1.73)	(-1.91)		
	0.0006		0.0005**	
	(0.33)		(7.03)	
	-0.0007			0.0007**
	(-0.37)			(6.76)
	-0.0031	-0.0122	0.0004**	0.0007**
	(-1.60)	(-1.76)	(6.84)	(7.44)
Small Stocks	0.0024	-0.0177*		
	(1.15)	(-2.51)		
	-0.0011		0.0006**	
	(-0.52)		(6.19)	
	-0.0029			0.0009**
	(-1.30)			(8.35)
	-0.0057*	-0.0180*	0.0005**	0.0009**
	(-2.57)	(-2.54)	(5.71)	(8.67)
Mid-Cap Stocks	0.0039*	-0.0016		
	(2.18)	(-0.16)		
	0.0022		0.0004**	
	(1.21)		(5.06)	
	0.0026			0.0003**
	(1.29)			(2.97)
	0.0003	0.0005	0.0004**	0.0003**
	(0.15)	(0.05)	(5.06)	(3.02)
Large Stocks	0.0032	-0.0167		
	(1.79)	(-1.38)		
	0.0012		0.0004**	
	(0.66)		(4.28)	
	0.0013			0.0003**
	(0.62)			(2.70)
	-0.0007	-0.0121	0.0003**	0.0003**
	(-0.36)	(-1.00)	(4.32)	(3.24)

Table 6
Return Predictability: Historical Trends

This table presents an investigation of historical trends in short-horizon (weekly) return predictability with past return as the predictive variable. The dependent variable is weekly return (from CRSP), Return(t+1), and the independent variables are an intercept and Return(t). We implement a Fama-MacBeth methodology for the regressions: (i) a cross-sectional regression is performed for each week in the sample period, and (ii) test statistics are based on the time-series of the coefficient estimates. We present the mean coefficient from the weekly regressions, and use the Newey-West correction for the standard errors to compute the t-statistics. Since our main analysis (e.g., Table 5) uses four years of data (2000-2003), we examine historical trends by running the regressions on non-overlapping four-year periods going back from 2003 to the beginning of data availability in CRSP. The table presents regression results for all stocks and by size groups. We sort stocks according to market capitalization into ten deciles, and define deciles 1, 2, 3, and 4 as small stocks, deciles 5, 6, and 7 as mid-cap stocks, and deciles 8, 9, and 10 as large stocks. ** indicates significance at the 1% level and * indicates significance at the 5% level (both against a two-sided alternative).

	All Stocks		**Small Stocks**		**Mid-Cap Stocks**		**Large Stocks**	
	Intercept	Return(t)	Intercept	Return(t)	Intercept	Return(t)	Intercept	Return(t)
1964 – 1967	0.0039**	-0.0765**	0.0054**	-0.0925**	0.0036**	-0.0695**	0.0024*	-0.0561**
	(3.21)	(-11.33)	(3.77)	(-12.32)	(2.95)	(-8.31)	(2.23)	(-7.27)
1968 – 1971	0.0013	-0.0920**	0.0013	-0.1084**	0.0013	-0.0848**	0.0012	-0.0786**
	(0.63)	(-12.63)	(0.58)	(-12.83)	(0.64)	(-9.67)	(0.72)	(-10.05)
1972 – 1975	0.0004	-0.0973**	0.0006	-0.1263**	0.0004	-0.0814**	0.0003	-0.0635**
	(0.16)	(-14.59)	(0.22)	(-17.86)	(0.16)	(-10.24)	(0.13)	(-7.64)
1976 – 1979	0.0046**	-0.0797**	0.0062**	-0.0930**	0.0046**	-0.0804**	0.0023	-0.0658**
	(3.04)	(-12.58)	(3.33)	(-13.98)	(3.06)	(-10.88)	(1.78)	(-9.06)
1980 – 1983	0.0051**	-0.0698**	0.0061**	-0.0765**	0.0050**	-0.0715**	0.0042*	-0.0657**
	(3.04)	(-13.34)	(3.38)	(-13.49)	(2.99)	(-10.67)	(2.52)	(-7.85)
1984 – 1987	0.0023	-0.0688**	0.0013	-0.0758**	0.0026	-0.0720**	0.0035	-0.0710**
	(1.10)	(-10.84)	(0.58)	(-10.50)	(1.26)	(-9.16)	(1.83)	(-7.80)
1988 – 1991	0.0036*	-0.0909**	0.0033	-0.1114**	0.0033*	-0.0358**	0.0036*	-0.0471**
	(2.16)	(-7.83)	(1.64)	(-7.06)	(2.19)	(-4.37)	(2.51)	(-5.31)
1992 – 1995	0.0031**	-0.0730**	0.0035**	-0.0936**	0.0026**	-0.0331**	0.0029**	-0.0446**
	(3.37)	(-12.63)	(3.14)	(-11.59)	(2.92)	(-4.50)	(3.57)	(-6.42)
1996 – 1999	0.0028	-0.0376**	0.0022	-0.0448**	0.0029	-0.0182	0.0033*	-0.0302**
	(1.74)	(-5.69)	(1.27)	(-6.75)	(1.72)	(-1.48)	(2.22)	(-3.52)
2000 – 2003	0.0031	-0.0229**	0.0038*	-0.0383**	0.0033	0.0099	0.0023	-0.0126
	(1.78)	(-3.27)	(1.98)	(-4.94)	(1.86)	(1.09)	(1.30)	(-0.99)

Table 7
Returns around Individual Buying by Effective Spread Groups

This table presents analysis of market-adjusted returns around intense buying by individuals separately for three groups of stocks sorted by percentage effective spreads. We partition the sample each week into three groups according to the average percentage effective spread (from the TAQ database) of the stocks. For each week in the sample period, we use the previous nine weeks to form net individual trading deciles for each spread group. Each stock is put into one of ten deciles according to the value of NIT in the current week relative to its value in the previous nine weeks. Decile 10 contains the stocks with the most intense buying (positive NIT). Let k be the number of days prior to or following portfolio formation each week. We calculate eight cumulative return numbers for each of the stocks in the portfolio: CR(t-k,t-1) where $k \in \{20, 15, 10, 5\}$ days and t is the first day of the formation week, and CR(t+1,t+k) where $k \in \{5, 10, 15, 20\}$ days and t is the last day of the week. The return on the portfolio is then adjusted by subtracting the return on a market proxy (the value-weighted portfolio of all stocks in the sample). We present the time-series mean and t-statistic for each market-adjusted cumulative return measure. ** indicates significance at the 1% level and * indicates significance at the 5% level (both against a two-sided alternative).

Spread Group		**k=–20**	**k=–15**	**k=–10**	**k=–5**	**k=+5**	**k=+10**	**k=+15**	**k=+20**
Small %EffSprd	Mean	-0.0066**	-0.0085**	-0.0087**	-0.0069**	0.0011	0.0024*	0.0043**	0.0072**
	t-stat.	(-4.67)	(-6.55)	(-7.94)	(-8.28)	(1.51)	(2.13)	(3.59)	(5.21)
Medium %EffSprd	Mean	-0.0196**	-0.0198**	-0.0169**	-0.0118**	0.0038**	0.0078**	0.0108**	0.0142**
	t-stat.	(-8.63)	(-9.94)	(-10.12)	(-10.40)	(3.38)	(4.85)	(5.37)	(6.30)
Large %EffSprd	Mean	-0.0399**	-0.0371**	-0.0310**	-0.0201**	0.0070**	0.0122**	0.0186**	0.0246**
	t-stat.	(-8.80)	(-9.72)	(-10.08)	(-9.72)	(3.68)	(4.24)	(4.98)	(5.65)

Table 8

Return Volatility around Individual Trading

This table presents analysis of daily standard deviation of returns around intense buying and selling activity of individuals as given by the net individual trading measure (NIT). For each week in the sample period, we use the previous nine weeks to form NIT deciles. Each stock is put into one of ten deciles according to the value of NIT in the current week relative to its value in the previous nine weeks. Decile 1 contains the stocks with the most intense selling (negative NIT) while decile 10 contains the stocks with the most intense buying (positive NIT). We present the results for four portfolios: (i) decile 1, (ii) deciles 1 and 2, (iii) deciles 9 and 10, and (iv) decile 10. For each stock and each week, we calculate the standard deviation of daily returns in a 9-day window centered on day $k \in \{-20, -15, -10, -5, 0, 5, 10, 15, 20\}$, where $k = 0$ is the middle of the formation week. We subtract from these numbers the "normal" 9-day return standard deviation (which we compute as the average of daily return standard deviations on all non-overlapping 9-day windows in the sample period). Every week we calculate the average of these standard deviations across all the stocks in each of the four portfolios. Each cell in the table contains the time-series mean for each portfolio and a t-statistic testing the hypothesis of a zero mean. The last three columns provide the differences in standard deviations from $k = –20$ to $k = 0$, $k = 0$ to $k = +20$, and $k = –20$ to $k = +20$, with t-statistics testing the hypothesis of zero differences. ** indicates significance at the 1% level and * indicates significance at the 5% level (both against a two-sided alternative).

Portfolio	**k = –20**	**k = –15**	**k = –10**	**k = –5**	**k = 0**	**k = +5**	**k = +10**	**k = +15**	**k = +20**	**k = –20 to k = 0**	**k = 0 to k = +20**	**k = –20 to k=+20**
Intense Selling (decile 1)	-0.0001 (-0.29)	-0.0002 (-0.38)	-0.0001 (-0.29)	0.0012** (2.60)	0.0018** (4.22)	0.0002 (0.45)	-0.0006 (-1.30)	-0.0007 (-1.73)	-0.0008 (-1.91)	0.0020** (4.41)	-0.0027** (-5.53)	-0.0007 (-1.34)
Selling (deciles 1&2)	0.0001 (0.22)	0.0001 (0.22)	0.0002 (0.42)	0.0008 (1.88)	0.0009* (2.07)	-0.0002 (-0.44)	-0.0006 (-1.35)	-0.0007 (-1.78)	-0.0008 (-1.79)	0.0008 (1.75)	-0.0016** (-3.49)	-0.0009 (-1.68)
Buying (deciles 9&10)	-0.0008 (-1.89)	-0.0008* (-2.05)	-0.0004 (-1.09)	0.0005 (1.22)	0.0013** (2.66)	0.0007 (1.57)	0.0001 (0.18)	0.0000 (0.04)	-0.0001 (-0.28)	0.0020** (4.05)	-0.0014* (-2.91)	0.0006 (1.23)
Intense Buying (decile 10)	-0.0010** (-2.62)	-0.0011** (-2.83)	-0.0008* (-2.00)	0.0008 (1.83)	0.0022** (4.50)	0.0013** (2.68)	0.0003 (0.62)	0.0001 (0.17)	-0.0001 (-0.15)	0.0033** (6.44)	-0.0023** (-4.71)	0.0010 (1.88)

Table 9
Principal Component Analysis

This table presents a principal component analysis of returns and the net individual trading measure (NIT) at the daily frequency. Panel A reports the results of a principal component analysis of 1,000 sub-samples of 180 stocks each (since we have more stocks in our sample than days in the sample period). We perform a principal component analysis on each sub-sample, and report the mean (Real Mean) and standard deviation (Real Std.) across sub-samples of the percentage of the variance explained by the first 10 principal components. We then construct 1,000 additional 180-stock random sub-samples. We compute for each stock the mean and standard deviation of the variable of interest (say NIT) and generate an artificial time-series for each stock drawn from a normal distribution with the same mean and standard deviation. We perform a principal component analysis on the simulated data of each sub-sample, and report the mean (Sim. Mean) across sub-samples of the percentage of the variance explained by the first 10 principal components. We then report the difference in the percentage of the variance explained by the different principal components (PC1, PC2, sum of PC1-5, sum of PC1-10) between the real data and the simulated data. Panel B reports the results of a principal component analysis done separately on each size decile for NIT. We sort the stocks according to average market capitalization over the sample period into 10 deciles. We perform a principal component analysis on each decile and report the percentage of the variance explained by both the first 5 and the first 10 principal components (PC1-5 and PC1-10, respectively). We then use the mean and standard deviation of each stock to generate 500 artificial time-series drawn from the normal distribution to form 500 independent sub-samples for each decile. We perform a principal component analysis on each sub-sample and save the mean across the sub-samples of the percentage of the variance explained by the first 5 and 10 principal components. We then report the difference in the percentage of the variance explained by the principal components between the real data and the simulated data.

Panel A: Percentage of Variance Explained by Principal Components (1000 random samples of 180 stocks)

		PC1	**PC2**	**PC1-5**	**PC1-10**
Returns	Real Mean	0.1317	0.0267	0.2125	0.2709
	Real Std.	0.0079	0.0027	0.0097	0.0100
	Sim. Mean	0.0111	0.0108	0.0533	0.1033
	Diff.	0.1207	0.0159	0.1592	0.1676
NIT	Real Mean	0.0280	0.0239	0.1007	0.1641
	Real Std.	0.0022	0.0017	0.0044	0.0052
	Sim. Mean	0.0111	0.0108	0.0533	0.1033
	Diff.	0.0170	0.0131	0.0474	0.0608

Panel B: Percentage of Variance of NIT Explained by Principal Components (size deciles)

		Decile 1 (small)	**Decile 2**	**Decile 3**	**Decile 4**	**Decile 5**	**Decile 6**	**Decile 7**	**Decile 8**	**Decile 9**	**Decile 10 (large)**
PC	Real	0.0805	0.0945	0.0872	0.0941	0.0946	0.0994	0.0995	0.1096	0.1202	0.1521
1-5	Diff.	0.0313	0.0454	0.0382	0.0450	0.0456	0.0503	0.0504	0.0606	0.0710	0.1031
PC	Real	0.1427	0.1536	0.1456	0.1552	0.1526	0.1558	0.1573	0.1678	0.1846	0.2191
1-10	Diff.	0.0474	0.0583	0.0506	0.0599	0.0576	0.0605	0.0619	0.0727	0.0893	0.1241