Individual Investor Sentiment and Stock Returns

Ron Kaniel, Gideon Saar, and Sheridan Titman

First version: February 2004
This version: September 2004

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 Stern School of Business, New York University, 44 West Fourth Street, New York, NY 10012 (Tel: 212-998-0318, gsaar@stern.nyu.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, Joel Hasbrouck, and seminar participants at Duke University, INSEAD, London Business School, and New York University. 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 Sentiment 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 individual investor sentiment and short-horizon (weekly) return reversals that have been documented in the literature. Our results reveal that individual investor sentiment predicts future returns, and that the information content of investor sentiment is distinct from that of past returns or past volume. Furthermore, the trading of individuals 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 there is very little cross-sectional correlation of our individual sentiment measure 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, 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. For each stock on each day, we have the aggregated volume of executed buy and sell orders of individuals. 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. This data allows us to construct a daily measure of individual investor sentiment for each stock by subtracting the sell volume of individuals from their buy volume and dividing by the average daily volume of the stock.

Our paper focuses on the dynamic relation between individual investor sentiment and returns over relatively short horizons (e.g., weekly and monthly). The results suggest that individuals tend to be contrarians, at least in the short-run. The mean market-adjusted returns in the 20 days prior to a week of intense individual selling is 3.97%, while prior to a week of intense individual buying it is −2.54%.[1] We also examine the extent to which the accumulation of shares by individuals predicts future returns and find that stocks experience excess returns of 1.40% in the 20 days following a week of intense buying by individuals. However, we find no significant return pattern following a week of intense individual selling.

[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).

Our paper is part of a growing literature on individual investors. Because of data availability, there are a number of studies that examine non-U.S. data sets, and these studies also find that individuals exhibit contrarian investment choices. For example, Grinblatt and Keloharju (2000) find evidence of contrarian choices in a study of Finish individuals, Choe, Kho, and Stulz (1999) find similar results using Korean data, and Jackson (2003) has similar findings in a study of Australian individuals. Goetzmann and Massa (2002) examine the accounts of individual investors in a fund that follows the S&P500 index and find that countrarian outnumber momentum traders two to one. In the only other study of U.S. individual investors that we are aware of that addresses these issues, Odean (1998, 1999) finds that those who trade using one of the major U.S. discount brokers tend to hold on to their losers and sell their winners, which is somewhat different but consistent with the idea that individuals are contrarians.

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 (1998) and Grinblatt and Keloharju (2000), looking at longer horizons, find that individual investors make poor investment choices. Like us, Odean studies U.S. 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.[2] Our results also contrast with Griffin, Harris, and Topaloglu (2003), who find no significant relation between the trading imbalances of 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

[2] 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.

overperform the stocks they sold.[3] However, our findings are similar to the Australian evidence in Jackson (2003), who suggests that individuals perform well over shorter horizons.

Our evidence, which contrasts with previous U.S. findings, is consistent with the idea that the contrarian tendency of individuals leads them to act as liquidity providers to other investors (e.g., institutions) which 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.[4]

To further explore the role of individual investors as liquidity providers, we examine how our results relate to the short-horizon return reversals observed by Jegadeesh (1990) and Lehmann (1990). In theory, these reversals can be due to either investor overreaction or to illiquidity.[5] To distinguish between these alternatives, Subrahmanyam (2003) constructs a model in which a liquidity explanation requires past order flow to be a predictor of future returns together with past returns.[6] Subrahmanyam

---

[3] It is possible that Griffin *et al.* found no significant relation between individual orders and future returns because of a limitation of their dataset. Since they do not directly identify trades as coming from individuals, Griffin *et al.* categorize brokerage firms as predominantly serving either individuals or institutions, and treat all trades coming from a single brokerage firm as if they belong to a single investor type. This procedure may introduce noise into the analysis that may mask the predictability result. It is also possible that the predictability we find is more prevalent in NYSE stocks than in NASDAQ stocks.

[4] It is interesting to note that while Barber, Lee, Liu, and Odean (2004a) find that Taiwanese individuals on average lose when trading, they also find that individuals gain from liquidity providing trades at short horizons (10 and 25 days).

[5] 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.

[6] Mase (1999) considers similar issues in a study of U.K. stocks and concludes that the evidence supports overreaction.

tests his model using monthly returns and net trade imbalances signed using the Lee and Ready (1991) algorithm as a proxy for order flow. He finds no significant relation between returns and his measure of past order flow imbalance, concluding that the results do not support the notion that the provision of liquidity by risk-averse agents drives the reversals.

Subrahmanyam's (2003) tests use a measure of order flow that aggregates inferred market orders of everyone in the market rather than focusing specifically on the trading of liquidity providers. Our tests, in contrast, consider the possibility that since individuals act as contrarians, they effectively provide liquidity, irrespective of whether they trade using limit or market orders.[7] If this is indeed the case, then the trading of individuals may provide a better proxy for the relevant order flow variable than the market-wide variable used by Subrahmanyam. Our empirical tests suggest that this may in fact be the case.

In addition to examining how individual sentiment relates to these return reversals, we also look at the interplay between trading volume, investor sentiment, and returns. Our analysis is thus related to the literature documenting that volume contains information that is useful in predicting short-horizon returns (Conrad, Hameed, and Niden, 1994; Gervais, Kaniel, and Mingelgrin, 2001; Llorente, Michaely, Saar, and Wang, 2002). This literature postulates that volume arises from shocks to investor hedging needs, private information, or trader interest in a given stock. Since such shocks can give rise to a demand for immediacy by institutions and liquidity provision by

[7] The Lee and Ready (1991) algorithm used by Subrahmanyam tries to establish 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. Our measure allows for other possibilities. For example, institutions that want to move large positions could use dynamic limit order strategies and therefore their demand for immediacy would be accommodated by the individuals' market orders. Since we focus on the identity of the trader rather than the order type, we allow for a somewhat broader interpretation of liquidity provision with respect to order type.

individuals, it is possible that volume and individual investor sentiment in fact contain the same information about future returns.

To examine how liquidity provision by individuals relates to these return and volume patterns we run multivariate regressions of weekly returns on past returns, volume, and investor sentiment. The results of these regressions indicate that individual investor sentiment is a powerful predictor of future returns that is not subsumed by either past returns or past volume. We also examine the predictive effect of individual sentiment, past returns, and volume by sorting stocks into portfolios. The returns of portfolios constructed from independent sorts by individual sentiment and past volume indicate that both variables predict returns. In addition, the relation between individual sentiment and returns remains significant in portfolios sorted by past returns and individual sentiment. However, in these latter sorts, there is no evidence of an independent past returns effect. In other words, individual sentiment seems to subsume the past returns effect.

Finally, we look at the question of whether individual investor sentiment is "systematic." Our analysis is motivated by claims in the behavioral finance literature that if fluctuations in noise trader sentiment affect many assets, then the risk they create cannot be diversified and will be priced in equilibrium. We conduct a principal component analysis of the individual investor sentiment and find very little correlated actions of individuals across stocks: the first principal component of the sentiment explains only 1.33% of the variance over and above a simulated benchmark created from independent data.

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 our measure of investor sentiment and returns. The investigation of short-horizon return predictability and its relation to investor sentiment is carried out in Section

4. Section 5 looks at whether the actions of individuals are correlated across stocks, and 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 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, both electronic and manual (those handled by floor brokers). One of the fields associated with each order, called Account Type, specifies whether the order originated 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 up to 2,099 shares that have been identified as individual investor orders.[8] 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

[8] 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.

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, 2002.[9] 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 1,920 stocks. For each stock on each day, we have the aggregated volume of executed buy and sell orders of individuals. An important advantage of this dataset is that it contains information on executed *orders*, rather than trades, and therefore we are able to determine unambiguously whether an individual buys or sells shares. 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.[10]

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.[11] However, it is very likely that the fraction of volume these brokers send to the NYSE consists of orders that create

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

[10] To construct the size groups, we sort the stocks in the sample according to average market capitalization over the sample period and form ten deciles. Small stocks are defined as those in deciles 1, 2, 3, and 4. Mid-cap stocks are those in deciles 5, 6, and 7, while large stocks are those in deciles 8, 9, and 10.

[11] 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).

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 investor sentiment measure (detailed below) that captures the net trading of individuals probably reflects imbalances in the market as a whole.

We construct a daily measure of investor sentiment 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 in the calendar year. Specifically, we define Net Investor Sentiment (NIS) for stock *i* on day *t* as:

$$NIS_{i,t} = \frac{\text{Individual buy dollar volume}_{\text{i,t}} - \text{Individual sell dollar volume}_{\text{i,t}}}{\text{Average daily dollar volume in calendar year}_{\text{i,year}\in\{2000,2001,2002\,:\,\text{t}\in\text{year}\}}}$$

Since our goal is to identify periods in which individuals are accumulating or selling an unusual amount of shares, we use in our analysis the deviations of this measure from its mean. In other words, we define $ANIS_{i,t}$ (Abnormal $NIS_{i,t}$) as $NIS_{i,t}$ minus the average of NIS_i over the sample period.

3. Dynamic Relation between Investor Sentiment and Returns

In this section we examine the relation between individual investor sentiment and returns. In the first subsection we examine the extent to which our sentiment measure is related to past returns, as well as its ability to forecast future returns. In the second subsection we examine how our sentiment measure relates to volatility.

3.1 Investor Sentiment and Returns

We start by aggregating daily investor sentiment to create a weekly ANIS measure and identify those weeks where either positive or negative ANIS is the most pronounced. This is done by comparing each stock's ANIS value in a given week (the formation week) with the values of ANIS in the previous 9 weeks, and placing the stocks in decile portfolios. In other words, decile 1 contains stocks for which ANIS in the formation week is more negative than the stocks' own ANIS in the previous 9 weeks. We call this decile the "intense selling portfolio." Similarly, decile 10 contains stocks with the most positive ANIS relative to the previous 9 weeks, and we call this decile the "intense buying portfolio." 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.

Panel A of Table 2 presents the cumulative market-adjusted returns for these four individual investor sentiment portfolios.[12] 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 CR(*k*). The first line of Panel A 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 3.97%, and the mean excess return in the five days prior to that week is 1.92%. These returns are highly statistically significant. The last line of the panel 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 –2.54%, 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 Panel A of 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 Korea. As we mentioned in the introduction, one interpretation of these results is that individuals effectively provide liquidity to

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

institutions, selling shares when the buying pressure from institutions pushes prices up and buying shares when the selling pressure from institutions pushes prices down.

The table also provides evidence on returns following intense individual buying and selling activity. We observe positive excess returns following weeks in which individuals accumulate an unusual number of shares. The portfolio of stocks in decile 10 earns 0.33% market-adjusted returns in the week after the intense buying and 1.40% 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.[13]

These results should be contrasted with the common characterization of individual investors as "noise" traders who lose money on average. Indeed, the evidence is 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. While the effect should be symmetric in the sense that liquidity provision should 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 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. We use the percentage effective spread (the distance from the transaction price to the quote midpoint divided by

[13] We also looked at whether such dynamic relations (contrarian patterns and predictability) exist between the value-weighted market return and a value-weighted measure of the individual investor sentiment, but no statistically significant patterns were found.

the quote midpoint) as a proxy for the liquidity of a stock.[14] The larger the effective spread, the greater the price movement on trades and therefore the less liquid the stock. In Panel B of Table 2 we sort stocks each week according to the average percentage effective spread and put them into three groups: small, medium, and large.[15] 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: 0.9% in the 20 days following portfolio formation in the small spread group, 1.53% in the medium spread group, and 1.87% in the large spread group. These results are consistent with the hypothesis that individuals generate excess returns by accommodating the liquidity needs of institutions.

While these findings of return predictability can be interpreted as evidence of market inefficiency, they are at least qualitatively consistent with what we would expect if individuals provide liquidity to the market and profit from their service. Of course, one could argue that the magnitude of the profits associated with these trades suggest that the compensation for liquidity provision is too high. However, it should be noted that a strategy designed to exploit this presumed inefficiency would be quite transaction-intensive and entail taking on some risk.[16]

We performed additional analyses to examine the robustness of our findings to the methodology we employ. The reason we adopted the methodology of forming deciles by comparing a stock's ANIS in the formation week to its own past ANIS was because the

---

[14] The bid and ask prices are taken from the TAQ data base.

[15] 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. For example, a stock will be classified as low liquidity if its effective spread is larger due to worsened liquidity on that week even if normally it has a lower effective spread and on other weeks it is in the medium or small spread groups.

[16] We should also note that 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. In fact, Account Type information is not available even to the specialists who oversee trading on the NYSE floor, but rather is collected in the CAUD files for regulatory purposes.

impact of trading imbalances on future prices is likely to be related to the stock's ability to absorb order flow. Subrahmanyam (2003) makes a similar point stating that inventory control effects (that create the return reversals) predict a downward pressure on the price of a stock in the absolute sense, not a relative (cross-sectional) sense. Our methodology, 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.[17]

Nonetheless, to verify that these results are not driven by the manner in which we form the ANIS deciles, we also used a cross-sectional sorting procedure where each week all stocks are sorted on ANIS (relative to each other) and grouped into ten deciles. We then repeated the analysis for the four portfolios of individual buying and selling as in Table 2. The results were similar, and both contrarian tendencies and the return predictability on buying were statistically significant.

We also examined the robustness of our results to different definitions of excess returns. Specifically, we repeated the analysis with excess returns from a market model regression, with industry-adjusted returns and with raw returns.[18] The results were similar with all return definitions, and both the contrarian pattern and return predictability results were statistically significant.

3.2 Investor sentiment and volatility

Another potential explanation for the excess returns following individual buys is that individuals systematically buy stocks when they become riskier. Perhaps, institutions are

---

[17] The methodology also does not use any future information in the classification in order to make sure that the effects we document can be viewed in terms of return predictability. Note that despite the fact that ANIS is normalized by the average dollar volume in a calendar year, when we compare it to past weeks, the denominator is the same and the comparison is done only by considering the numerator that consists of the weekly volume imbalance.

[18] 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

either more risk averse, or perhaps more savvy, and thus choose to sell their shares when they have information that suggests that the risk of a stock is increasing. It has also been suggested that the activities of individual investors, i.e., the noise traders in the behavioral finance literature, make stocks more volatile or riskier.[19]

We cannot reject such interpretations off hand, as both the beta and the weekly standard deviation of returns of the individuals' buying portfolio are higher than those of their selling portfolio.[20] However, this observation may simply be an artifact of transitory changes in volatility around intense trading by individuals due to the price pressure and return reversals that are associated with liquidity provision; As such these would not represent fundamental differences in the risk attributes of stocks individuals buy as opposed to those they sell.

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 investor 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 3, which presents the cross-sectional mean of these abnormal volatility measures in each ANIS portfolio, tells us how volatility of returns evolves around the trading of individuals.

---

[19] See, for example, De Long, Shleifer, Summers, and Waldmann (1990a) and Shleifer and Summers (1990).

[20] The beta of the intense buying portfolio is 0.1027 higher than the beta of the intense selling portfolio and the standard deviation of the intense buying portfolio is 0.0033 higher than the standard deviation of the intense selling portfolio.

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 exactly at the level of the average volatility at $k = -20$, then increases to 0.0014 above average volatility at $k = -5$, reaches 0.0021 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 0.0021 and statistically significant, and the decrease of –0.0029 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 (0.0038) is observed from –20 to 0 before intense buying activity (decile 10), and most of it is subsequently reversed (–0.0026) 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.

4. Short-Horizon (Weekly) Predictability of Returns

In this section, we examine how our evidence relates to the Jegadeesh (1990) and Lehmann (1990) evidence 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. Alternatively, if the return reversals reflect what Lehmann (1990) characterizes as inefficiencies in the market for short-term liquidity, it is possible that the imbalances in individual buying and selling can explain the return reversals.[21] In this section's first subsection we examine how the ANIS effect interacts with both the return reversal effect and the turnover effect previously documented by Gervais, Kaniel, and Mingelgrin (2001). In the second subsection we examine the return reversal effect over previous

[21] Jegadeesh and Titman (1995) provides evidence that indicates that the daily return reversals exist mainly in relatively illiquid stocks.

time periods to provide some evidence on the extent to which the phenomenon we examine is stable over time.

4.1 The Relation between ANIS, Turnover, and Returns

To examine this issue, stocks are sorted each week into five quintiles of weekly returns, where quintile 1 (5) contains the stocks with the most negative (positive) return. Each week stocks are also put into five ANIS quintiles according to the value of ANIS that week relative to the values of ANIS of that same stock in the previous nine weeks (as in Section 3). Quintile 1 contains stocks with the most negative ANIS, or net selling, while quintile 5 contains stocks with the most positive ANIS, or net buying. Then, 25 portfolios are formed as the intersection of the five return quintiles and the five ANIS quintiles. For each of the portfolios we compute the market-adjusted return in the week following the formation week.[22]

Panel A of Table 4 focuses on the sentiment of individual investors and reports the time-series averages of the weekly market-adjusted returns for the 25 portfolios. Looking across the columns of the different return quintiles, no simple pattern of return reversal can be found. Take, for example, the row of ANIS quintile 1 (individual selling). The market-adjusted return in the subsequent week for return quintile 1 (most negative current week return) is –0.26%, for return quintile 3 it is –0.13%, and for return quintile 5 it is –0.25%. Similarly, in the row of ANIS quintile 5, both quintile 1 and quintile 5 of return are positive (i.e., a reversal after negative returns but a continuation after positive returns). 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

[22] We examined the robustness of our findings to different definitions of returns by repeating the analysis using raw returns, market-model-adjusted returns, and industry-adjusted returns (as in Section 3.1). Our conclusions from all these return definitions were the same.

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 ANIS quintiles.[23]

On the other hand, there is a pronounced pattern within each quintile of past returns going from past individual selling (ANIS quintile 1) to past individual buying (ANIS 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. For example, the market-adjusted return in the following week in the column of return quintile 1 is –0.26% for ANIS quintile 1, 0.03% for ANIS quintile 3, and 0.41% for ANIS quintile 5. 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 (ANIS quintile 5) and selling those stocks experiencing intense individual selling (ANIS quintile 1) in each return quintile. All these portfolios realize statistically significant positive payoffs, ranging from 0.30% to 0.67% per week.[24]

It is also possible that ANIS contains the same information contained in trading volume, which would suggest that our result is a restatement of the findings in Gervais, Kaniel, and Mingelgrin (2001) that high returns follow high volume. We therefore repeat the analysis sorting the stocks each week into five quintiles of weekly ANIS 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 ANIS quintiles (the turnover of a stock on a certain week is compared to the turnover of the same stock in the previous ten weeks). Based on these

---

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

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

5X5 sorts, 25 portfolios are formed as the intersection of the five turnover quintiles and five ANIS quintiles, and their returns are calculated.

Panel B of Table 4 reveals that the information in the investor sentiment measure is distinct from that in turnover, and both provide independent information about future returns. In particular, the strategy of buying the stocks in ANIS quintile 5 and selling the stocks in ANIS 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 ANIS row.

The finding that both the sentiment of individual investors and turnover have significant ability to 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 sentiment could potentially explain the high-volume return premium, leaving no role for turnover. Our findings, however, suggest that turnover and ANIS contain different information and neither of them subsumes the other.

To examine turnover, ANIS, and, past returns simultaneously we estimate a series of Fama and MacBeth (1973) regressions.[25] Table 5 presents the estimates of a regression of returns in week *t* on a set of dummy variables that represent week *t*-1 return

[25] 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).

quintiles, turnover quintiles, and ANIS quintiles. The results from these regressions are consistent with the findings in the previous table that consider the relation between ANIS, turnover, and past returns in two by two sorts. In particular, we find that ANIS and turnover are both strong predictors of future returns in these multiple regressions. After controlling for ANIS and turnover, however, past return quintiles provide no information about future returns.[26]

We also run a Fama-MacBeth regression specification where the current week's return is regressed on 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 specification is comparable with past literature (e.g., Jegadeesh, 1990; Subrahmanyam, 2003) that document a significant past return predictability effect. Table 6 presents the results of the regression analysis for the entire sample and for three size groups.

We first run a univariate regression of the subsequent week's return on each of the three predictive variables: return, ANIS, and turnover. We use a transformation of ANIS 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 ANIS value that week relative to the ANIS of that same stock in the previous nine weeks, where decile 1 (10) contains stocks with the most negative (positive) ANIS. We then use the decile rank of each stock on each week (the ANISDecile variable) as an independent variable in the

[26] 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 different from zero.

regressions.[27] Similarly, we use a transformation of turnover into decile ranks (as we do for ANIS) because Gervais, Kaniel, and Mingelgrin (2001) found such a transformation of volume useful in predicting returns.

In Panel A of Table 6 we use CRSP returns to be consistent with most of the papers in the return predictability literature. Looking at the results of the regressions using all stocks in the sample, the mean coefficient on past returns is negative and statistically significant, consistent with the weekly return reversal that was documented in the literature. The mean coefficient on ANISDecile is positive and highly statistically significant. Finally, the mean coefficient on TurnoverDecile is positive and highly statistically significant, consistent with the high-volume return premium phenomenon.

Even more interesting is the last model that uses all three explanatory variables together. The mean coefficient on past returns decreases slightly in magnitude, from −0.0243 in the univariate model to −0.0226 in the multivariate model, and remains statistically significant. Both the mean coefficients on ANISDecile and TurnoverDecile are very similar in the univariate and multivariate settings and are highly significant. The strong showing of both ANISDecile and TurnoverDecile in the multivariate regression supports our conclusion from Table 4 and Table 5 that these two variables contain different information. These findings suggest that we have identified an important new predictor variable with respect to weekly returns—the sentiment of individual investors.

The results of the regression analysis in Panel A of Table 6 give a somewhat different picture of return reversals from the analysis of 25 portfolios in Table 4 and the regressions using dummy variables for past return quintiles in Table 5. While it was difficult to discern a pattern of predictability along the return dimension in the analysis of

[27] For robustness, we also ran the regressions using ANIS, rather then the ANIS decile ranks, as the independent variable. This specification is similar in spirit to the cross-sectional robustness tests that we conducted in Section 3.1. The results were similar in that the mean coefficient on ANIS was positive and statistically significant in all the models (univariate and multivariate).

portfolios or return quintiles, here the mean coefficient on past return remains significant and does not change much even after ANIS is added as an explanatory variable.

We wanted to examine the robustness of these results to two issues: bid-ask bounce and nonsynchronous trading. 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. We therefore used the TAQ database to create a return series from end-of-day quote midpoints.[28] Since the prices used for constructing the return series are the midpoints between the bid and the ask, this series completely eliminates the bid-ask bounce problem. Also, 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 6 presents the results of the regressions with the midquote returns. In the regressions using all stocks, the mean coefficient on past returns in the univariate model is smaller in magnitude than the one in Panel A with the CRSP returns (-0.0185 versus -0.0243). In the multivariate model, in the presence of ANIS and turnover, the mean coefficient on past returns decreases in magnitude even further and is no longer statistically significant. The mean coefficient on ANIS, on the other hand, is the same irrespective on whether we are using returns from closing prices or midquotes, and is highly statistically significant. Therefore, it seems that the ability of past returns to

[28] 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.

predict future returns no longer exists in the Fama-MacBeth multivariate specification when we control for bid-ask bounce and nonsynchronous trading.

We also report in both panels of Table 6 the results of the regression analysis done separately for three groups of stocks: small stocks, mid-cap stocks, and large stocks. 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. Significant mean coefficients on ANISDecile and TurnoverDecile are found across the three size groups. However, the coefficient on past returns is reliably different than zero only for the small firms.

4.2 A Historical Perspective on Return Reversals

Since the three-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 three-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 7 indicate that the return reversal phenomenon has been changing. The second column of Table 7 shows a very clear trend in the estimated mean 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 1988–1990 period (−0.0940) to the 2000–2002 period (−0.0243). In fact, the magnitude has been at an all-time low since 1994. 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, 2000–2002. 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 three-year period.

Why are we observing such a trend? Perhaps the publication of this potential "inefficiency" might have prompted sophisticated traders to take advantage of it and eliminate the pattern. The past decade has seen an increase in the activity of proprietary trading desks utilizing quantitative strategies that are aimed at taking advantage of various price patterns. The growing ability to carry out computerized trading strategies together with a decline in trading costs could have brought an end to this inefficiency, if it were indeed an inefficiency.

This raises the question of what could have been (or still is in the case of small stocks) the reason for observing this phenomenon. Lehmann (1990) suggests that the weekly return reversal is due to inefficiency in the market for liquidity services. Mase (1999) concludes that weekly return predictability in the UK is due to irrational trading—overreaction and a subsequent correction—rather than frictions in the market for liquidity.

More recently, Subrahmanyam (2003) provides a model that captures both liquidity provision by risk-averse agents and irrationality in the sense of overconfidence. In his framework, reversals due to liquidity provision by risk-averse agents generate a relation between expected returns and past order flow while reversals due to overconfidence do not. Subrahmanyam tests his model using monthly returns and net trade imbalances signed using the Lee and Ready (1991) algorithm. He finds no significant relation between returns and his order flow measure, concluding that the liquidity provision hypothesis is not supported by the data.

Our results contrast with those presented by Subrahmanyam in that we find a very significant relation between returns and the past order flow imbalance of individual investors in a linear framework similar to the one he implements. The difference can be due to our ability to better measure the order flow imbalance of those agents who presumably provide liquidity, individuals. Subrahmanyam computes his measure using all trades in the market without the ability to focus on one class of investors or the other, and he needs to use an algorithm for signing trades that introduces errors into his measure while we know the true direction of orders. Another potential explanation for the difference is that we examine weekly predictability (due to the short sample period for which we obtained data), while he looks at monthly predictability.[29]

Our results suggest that liquidity provision may induce short-horizon return reversals. Both the effects of investor sentiment and turnover on returns are very prominent in recent data. On the other hand, the weekly return predictability using past return that was identified and investigated in the literature seems to be disappearing.

5. Is Individual Investor Sentiment Correlated Across Stocks?

In this section we examine whether individual investor sentiment is systematic in the sense that it affects all stocks at the same time. This is important because the behavioral finance literature suggests that, if indeed individual investors are "noise" traders, such systematic variation in their sentiment would affect expected returns. The argument describing our motivation for examining this issue 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

[29] Cooper (1999) finds significant weekly return predictability using past return and volume for large stocks, but his sample period ends in 1993. The evidence in Table 7 suggests that it is probably no longer possible nowadays to find predictability using past return information in mid-cap and large stocks.

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 investor sentiment measure and look at the percentage of variance of ANIS 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,[30] and look at the mean and standard deviation of the percentage of variance across the 1,000 random sub-samples. 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).[31]

Panel A of Table 8 shows the results of the principal component analysis of the individual investor sentiment measure and also of daily returns. The daily return analysis is shown to provide a sense of the magnitude of co-movement observed in the cross section of stocks. For example, 21.89% 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 5.8%, and therefore the difference between these two numbers, roughly 16.09%, is a better measure of the structure in the real data. The analysis of ANIS reveals very little evidence of correlated actions of

[30] 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.
[31] 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).

individual investors across stocks. Indeed, the first (and largest) principal component of ANIS explains only 1.33% of the variance (adjusted using the simulated data).

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 8 presents the results. Contrary to what one might have expected based on the above papers, the percentage of the ANIS variance explained by the first five principal components is lower for small stocks (3.20% for decile 1) than for large stocks (9.91% 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.[32] Our results indicate that it is difficult to find such correlated actions of individuals who trade on the NYSE. These results may suggest that finding a systematic influence of individual investor sentiment on expected returns may be difficult.

6. Conclusions

Our analysis of the trading of individual investors on the NYSE reveals that they behave as contrarians. The underlying reason for why individuals act in such a way is not well

---

[32] 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.

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. When large investors choose to accumulate shares, their buy orders push prices up inducing contrarian individuals to sell. Similarly, when large investors choose to sell shares, they push prices down and attract individual buyers.

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 potential traders that provide liquidity.

The evidence in this paper is consistent with the view that in the recent period that corresponds to an increase in institutional trading, a short-term liquidity provider could have generated profits by mimicking the trades of individual investors. There is also anecdotal evidence that suggests 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.[33] Indeed, the presence of these traders may be responsible for the reduction in the return reversals observed by Jegadeesh (1990) and Lehmann (1990).

Then why don't the strategies implemented by these short-term traders eliminate the excess returns associated with individual trades? 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 added risk associated with undertaking the liquidity-supplying trading strategies. Moreover, it may be the case that the mechanical strategies are unable to implement the strategies implicitly implemented by individual investors (the ANIS measure is not public information and institutions cannot simply use it to formulate their strategies). For example, in addition to buying past losers, our evidence suggests that individual investors tend to buy stocks that are temporarily more risky,

[33] For example, Automated Trading Desk (ATD) is one of the firms that pioneered the use of computerized expert systems applied to liquidity provision. While today 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.

which might be difficult to implement with a purely mechanical strategy. Perhaps institutions overreact to temporary increases in risk, and in doing so provide an opportunity for individual investors.

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 1,920 stocks and 752 trading days. For the analysis in Panel A of Table 8 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 sentiment 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 1.21% 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 8 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

Amihud, Y., and Mendelson, J., 1986. Asset pricing and the bid-ask spread. *Journal of Financial Economics* 17, 223-249,

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.

Brennan, M. J., and Subrahmanyam, A., 1996. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics* 41, 441-464.

Chan, L. K., and Lakonishok, J., 1995. The behavior of stock prices around institutional trades. *Journal of Finance* 50, 1147-1174.

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.

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.

Cooper, M., 1999. Filter rules based on price and volume in individual security overreaction. *Review of Financial Studies* 12, 901-935.

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., and Keloharju, M., 2001. What makes investors trade? *Journal of Finance* 56, 589-616.

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.

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

Pastor, L., and Stambaugh, R. F., 2003. Liquidity risk and expected stock returns. *Journal of Political Economy* 111, 642-685.

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

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

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., 2003. On distinguishing between rationales for short-horizon predictability of stock returns. Working paper, University of California at Los Angeles.

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

Wermers, R., 2000. Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *Journal of Finance* 55, 1655-1703.

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, 2002 with records in the CRSP database. We use ticker symbol and CUSIP to match the stocks to a special dataset containing aggregated buying and selling volume of individuals and institutions that was provided to us by the NYSE. There are 1,920 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 Dollar Volume, defined as the sum of executed buy and sell orders, and the Executed Order Size (in dollars and 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	5,698.9	61.62	2.37	0.0735	4,775.5	799.8
	Median	941.2	22.31	1.96	0.0622	1,148.0	675.5
Small stocks	Mean	289.3	13.05	2.15	0.0881	656.9	948.1
	Median	271.7	10.97	1.43	0.0745	356.8	773.6
Medium stocks	Mean	1,283.5	25.81	3.08	0.0711	2,119.2	738.4
	Median	1,180.5	23.93	2.41	0.0623	1,388.3	645.4
Large stocks	Mean	15,001.4	140.65	3.05	0.0634	12,404.0	691.4
	Median	5,256.9	38.36	2.48	0.0566	5,223.2	634.5

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 investor sentiment measure (ANIS). For each week in the sample period, we use the last 10 weeks (including the current week) to form ANIS deciles. Each stock is put into one of ten deciles according to the value of ANIS in the current week relative to its value in the previous ten weeks. Decile 1 contains the stocks with the most intense selling (negative ANIS) while decile 10 contains the stocks with the most intense buying (positive ANIS). In Panel A 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. In Panel B we partition the sample each week into three groups according to the average percentage effective spread of the stocks. We then present the cumulative return results for the Intense Individual Buying portfolio (decile 10) of each spread group. ** indicates significance at the 1% level and * indicates significance at the 5% level (both against a two-sided alternative).

Panel A: Returns around Individual Trading

Portfolio		**k=–20**	**k=–15**	**k=–10**	**k=–5**	**k=+5**	**k=+10**	**k=+15**	**k=+20**
Intense Selling	Mean	0.0397**	0.0374**	0.0302**	0.0192**	-0.0014	-0.0013	-0.0006	0.0009
(decile 1)	t-stat.	(13.42)	(15.19)	(15.80)	(14.82)	(-1.10)	(-0.66)	(-0.26)	(0.32)
Selling	Mean	0.0373**	0.0342**	0.0277**	0.0171**	-0.0016	-0.0016	-0.0011	0.006
(deciles 1&2)	t-stat.	(13.28)	(14.59)	(15.55)	(14.38)	(-1.37)	(-0.91)	(-0.48)	(0.25)
Buying	Mean	-0.0258**	-0.0249**	-0.0208**	-0.0138**	0.0032**	0.0066**	0.0096**	0.0134**
(deciles 9&10)	t-stat.	(-9.45)	(-10.86)	(-11.17)	(-10.84)	(2.76)	(3.75)	(4.34)	(5.10)
Intense Buying	Mean	-0.0254**	-0.0251**	-0.0218**	-0.0150**	0.0033**	0.0070**	0.0103**	0.0140**
(decile 10)	t-stat.	(-9.21)	(-10.67)	(-11.27)	(-11.33)	(2.77)	(3.89)	(4.66)	(5.37)

Panel B: Returns around Intense Individual Buying (decile 10) by Percentage Effective Spread Groups

Spread Group		**k=–20**	**k=–15**	**k=–10**	**k=–5**	**k=+5**	**k=+10**	**k=+15**	**k=+20**
Small %EffSprd	Mean	-0.0078**	-0.0097**	-0.0102**	-0.0080**	0.0012	0.0036**	0.0054**	0.0090**
	t-stat.	(-4.28)	(-5.86)	(-7.48)	(-7.66)	(1.23)	(2.61)	(3.70)	(5.28)
Medium %EffSprd	Mean	-0.0238**	-0.0240**	-0.0210**	-0.0146**	0.0036*	0.0085**	0.0118**	0.0153**
	t-stat.	(-8.37)	(-9.72)	(-10.18)	(-10.55)	(2.56)	(4.17)	(4.57)	(5.32)
Large %EffSprd	Mean	-0.0487**	-0.0448**	-0.0368**	-0.0240**	0.0053*	0.0097**	0.0152**	0.0187**
	t-stat.	(-8.77)	(-9.41)	(-9.55)	(-9.45)	(2.23)	(2.66)	(3.19)	(3.34)

Table 3
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 investor sentiment measure (ANIS). For each week in the sample period, we use the last 10 weeks (including the current week) to form ANIS deciles. Each stock is put into one of ten deciles according to the value of ANIS in the current week relative to its value in the previous ten weeks. Decile 1 contains the stocks with the most intense selling (negative ANIS) while decile 10 contains the stocks with the most intense buying (positive ANIS). 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)	0.0000 (-0.03)	0.0000 (-0.07)	0.0000 (0.04)	0.0014** (2.62)	0.0021** (4.14)	0.0003 (0.69)	-0.0004 (-0.80)	-0.0006 (-1.33)	-0.0008 (-1.52)	0.0021** (3.63)	-0.0029** (-4.41)	-0.0007 (-1.05)
Selling (deciles 1&2)	0.0001 (0.24)	0.0002 (0.30)	0.0003 (0.56)	0.0010 (1.92)	0.0010* (2.13)	-0.0001 (-0.20)	-0.0005 (-0.97)	-0.0007 (-1.45)	-0.0007 (-1.37)	0.0009 (1.56)	-0.0017** (-2.72)	-0.0008 (-1.13)
Buying (deciles 9&10)	-0.0009 (-1.90)	-0.0009 (-1.87)	-0.0005 (-0.99)	0.0006 (1.22)	0.0015** (2.80)	0.0009 (1.84)	0.0002 (0.46)	0.0002 (0.33)	0.0000 (0.04)	0.0024** (3.72)	-0.0015* (-2.37)	0.0009 (1.34)
Intense Buying (decile 10)	-0.0012* (-2.40)	-0.0011* (-2.36)	-0.0007 (-1.56)	0.0011* (2.08)	0.0027** (4.68)	0.0015** (2.80)	0.0004 (0.86)	0.0003 (0.55)	0.0001 (0.21)	0.0038** (5.72)	-0.0026** (-3.92)	0.0013 (1.81)

Table 4

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 investor sentiment measure (ANIS). For each week in the sample period, we use the last 10 weeks (including the current week) to form ANIS quintiles. Each stock is put into one of the five quintiles according to the value of ANIS in the current week relative to its value in the previous nine weeks (where quintile 1 has stocks with more negative ANIS, or more selling, and quintile 5 has stocks with more positive ANIS, 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 ANIS 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 the sentiment of individual investors. The last two rows of the panel give the payoff to the strategy of buying ANIS quintile 5 and selling ANIS 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 ANIS. 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 ANIS

		Return(t)						
		Q1 (<0)	Q2	Q3	Q4	Q5 (>0)	Q5–Q1	t-statistic
ANIS(t)	Q1 (<0)	-0.0026	-0.0022	-0.0013	-0.0003	-0.0025	0.0001	-0.05
	Q2	-0.0012	-0.0011	-0.0007	-0.0009	-0.0034	-0.0022	0.95
	Q3	0.0003	-0.0019	-0.0002	0.0008	-0.0012	-0.0015	0.70
	Q4	0.0023	-0.0002	0.0013	0.0014	0.0006	-0.0017	0.80
	Q5 (>0)	0.0041	0.0011	0.0023	0.0027	0.0025	-0.0016	0.80
	Q5–Q1	0.0067**	0.0032**	0.0036**	0.0030**	0.0050**		
	t-statistic	4.28	2.84	4.56	3.08	3.51		

Panel B: Weekly Return Predictability using Past Turnover and ANIS

		Turnover(t)						
		Q1 (low)	Q2	Q3	Q4	Q5 (high)	Q5–Q1	t-statistic
ANIS(t)	Q1 (<0)	-0.0050	-0.0029	-0.0022	-0.0006	0.0000	0.0052**	3.06
	Q2	-0.0040	-0.0039	-0.0021	0.0007	0.0008	0.0048**	3.49
	Q3	-0.0042	-0.0026	-0.0003	0.0007	0.0030	0.0072**	4.72
	Q4	-0.0029	-0.0006	0.0022	0.0039	0.0045	0.0074**	4.09
	Q5 (>0)	0.0011	0.0001	0.0005	0.0034	0.0058	0.0047**	2.65
	Q5–Q1	0.0063**	0.0030*	0.0027*	0.0040**	0.0058**		
	t-statistic	3.68	2.51	2.55	2.92	3.91		

Table 5
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 follow similar procedure for creating quintile dummy variables for the ANIS(t) and Turnover(t) variables. Construction of the individual investor sentiment measure (ANIS) 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	0.0070**	0.0063*	0.0077**	0.0071**
(t-statistic)	(2.81)	(2.23)	(3.06)	(2.80)
ANIS Q1	-0.0046**	-0.0055**	-0.0040**	-0.0041**
(t-statistic)	(-6.95)	(-5.02)	(-5.49)	(-4.65)
ANIS Q2	-0.0036**	-0.0050**	-0.0028**	-0.0026**
(t-statistic)	(-5.38)	(-4.75)	(-3.22)	(-3.38)
ANIS Q3	-0.0026**	-0.0036**	-0.0024**	-0.0014
(t-statistic)	(-4.39)	(-3.43)	(-3.04)	(-1.76)
ANIS Q4	-0.0011	-0.0012	-0.0017*	-0.0005
(t-statistic)	(-1.84)	(-1.11)	(-1.99)	(-0.58)
Turnover Q1	-0.0062**	-0.0075**	-0.0027*	-0.0028*
(t-statistic)	(-6.43)	(-5.39)	(-2.48)	(-2.15)
Turnover Q2	-0.0054**	-0.0074**	-0.0022*	-0.0031**
(t-statistic)	(-6.44)	(-5.89)	(-2.33)	(-2.93)
Turnover Q3	-0.0036**	-0.0046**	-0.0008	-0.0030**
(t-statistic)	(-4.76)	(-4.05)	(-1.06)	(-3.00)
Turnover Q4	-0.0018**	-0.0021	-0.0011	-0.0013
(t-statistic)	(-2.75)	(-1.83)	(-1.38)	(-1.40)
Return Q1	0.0017	0.0042**	-0.0018	-0.0002
(t-statistic)	(1.17)	(2.89)	(-1.02)	(-0.07)
Return Q2	0.0003	0.0006	-0.0004	-0.0018
(t-statistic)	(0.23)	(0.45)	(-0.31)	(-1.10)
Return Q3	0.0003	0.0014	-0.0007	-0.0018
(t-statistic)	(0.36)	(1.28)	(-0.52)	(-1.29)
Return Q4	0.0000	0.0016	-0.0018	-0.0011
(t-statistic)	(0.03)	(1.54)	(-1.63)	(-1.04)

Table 6
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), ANISDecile(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 investor sentiment (ANIS) measure is described in section 2, and the ANISDecile 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)	ANISDecile(t) (t-statistic)	Turnover Decile(t) (t-statistic)
All Stocks	0.0021	-0.0243**		
	(0.94)	(-2.80)		
	-0.0013		0.0006**	
	(-0.60)		(7.55)	
	-0.0029			0.0008**
	(-1.19)			(6.36)
	-0.0057*	-0.0226**	0.0005**	0.0009**
	(-2.44)	(-2.62)	(6.84)	(7.29)
Small Stocks	0.0006	-0.0337**		
	(0.24)	(-3.55)		
	-0.0038		0.0008**	
	(-1.52)		(6.63)	
	-0.0057*			0.0011**
	(-2.11)			(7.14)
	-0.0092**	-0.0335**	0.0007**	0.0011**
	(-3.45)	(-3.53)	(5.73)	(7.67)
Mid-Cap Stocks	0.0033	-0.0068		
	(1.47)	(-0.59)		
	0.0009		0.0005**	
	(0.39)		(5.00)	
	0.0016			0.0003**
	(0.66)			(2.68)
	-0.0010	-0.0038	0.0005**	0.0003**
	(-0.43)	(-0.33)	(4.91)	(2.84)
Large Stocks	0.0027	-0.0213		
	(1.24)	(-1.41)		
	-0.0002		0.0005**	
	(-0.08)		(4.44)	
	-0.0003			0.0004**
	(-0.13)			(3.21)
	-0.0027	-0.0159	0.0005**	0.0005**
	(-1.10)	(-1.04)	(4.51)	(3.85)

Panel B: Midquote Returns from the TAQ Database

Size Groups	Intercept (t-statistic)	Return(t) (t-statistic)	ANISDecile(t) (t-statistic)	Turnover Decile(t) (t-statistic)
All Stocks	0.0017	-0.0165*		
	(0.80)	(-2.02)		
	-0.0015		0.0006**	
	(-0.71)		(7.54)	
	-0.0030			0.0008**
	(-1.23)			(6.39)
	-0.0058*	-0.0148	0.0005**	0.0008**
	(-2.49)	-1.81	(7.17)	(7.21)
Small Stocks	0.0003	-0.0218**		
	(0.13)	(-2.67)		
	-0.0041		0.0007**	
	(-1.64)		(6.70)	
	-0.0057*			0.0010**
	(-2.12)			(7.71)
	-0.0093**	-0.0215**	0.0007**	0.0010**
	(-3.48)	(-2.64)	(5.95)	(8.20)
Mid-Cap Stocks	0.0030	-0.0036		
	(1.33)	(-0.32)		
	0.0007		0.0004**	
	(0.33)		(4.74)	
	0.0014			0.0003**
	(0.56)			(2.58)
	-0.0011	-0.0007	0.0004**	0.0003**
	(-0.46)	(-0.06)	(4.73)	(2.66)
Large Stocks	0.0023	-0.0185		
	(1.07)	(-1.22)		
	-0.0005		0.0005**	
	(-0.21)		(4.33)	
	-0.0006			0.0004**
	(-0.25)			(3.12)
	-0.0030	-0.0132	0.0004**	0.0005**
	(-1.23)	(-0.86)	(4.48)	(3.72)

Table 7
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 6) uses three years of data (2000-2002), we examine historical trends by running the regressions for non-overlapping three-year periods going back from 2002 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 – 1966	0.0026 (1.88)	-0.0777** (-9.06)	0.0034* (2.12)	-0.0916** (-9.66)	0.0023 (1.64)	-0.0778** (-7.51)	0.0019 (1.51)	-0.0528** (-5.52)
1967 – 1969	0.0029 (1.65)	-0.0823** (-12.40)	0.0044* (2.12)	-0.0918** (-11.25)	0.0026 (1.47)	-0.0771** (-9.84)	0.0015 (1.00)	-0.0781** (-9.86)
1970 – 1972	0.0017 (0.75)	-0.0884** (-10.67)	0.0014 (0.51)	-0.1162** (-13.08)	0.0019 (0.82)	-0.0685** (-6.42)	0.0023 (1.16)	-0.0653** (-7.50)
1973 – 1975	0.0001 (0.04)	-0.1060** (-12.87)	0.0006 (0.18)	-0.1366** (-15.92)	0.0002 (0.06)	-0.0909** (-9.43)	-0.0005 (-0.18)	-0.0681** (-6.49)
1976 – 1978	0.0044* (2.48)	-0.0831** (-10.45)	0.0061** (2.75)	-0.0988** (-11.93)	0.0044* (2.56)	-0.0797** (-8.95)	0.0018 (1.23)	-0.0716** (-8.78)
1979 – 1981	0.0039* (2.11)	-0.0690** (-11.23)	0.0046* (2.26)	-0.0800** (-13.10)	0.0039* (2.10)	-0.0715** (-8.77)	0.0030 (1.70)	-0.0554** (-5.58)
1982 – 1984	0.0046* (2.42)	-0.0655** (-12.00)	0.0053** (2.69)	-0.0680** (-10.30)	0.0046* (2.32)	-0.0724** (-9.34)	0.0042* (2.19)	-0.0691** (-7.67)
1985 – 1987	0.0031 (1.20)	-0.0735** (-9.50)	0.0021 (0.73)	-0.0802** (-9.05)	0.0035 (1.36)	-0.0750** (-8.34)	0.0043 (1.86)	-0.0739** (-6.82)
1988 – 1990	0.0016 (0.98)	-0.0940** (-6.21)	0.0006 (0.31)	-0.1161** (-5.65)	0.0017 (1.02)	-0.0309** (-3.37)	0.0025 (1.56)	-0.0427** (-4.16)
1991 – 1993	0.0056** (3.43)	-0.0791** (-11.68)	0.0068** (3.06)	-0.0948** (-10.49)	0.0051** (3.50)	-0.0517** (-5.47)	0.0043** (3.45)	-0.0511** (-5.98)
1994 – 1996	0.0030** (2.69)	-0.0555** (-9.40)	0.0031* (2.56)	-0.0750** (-8.66)	0.0028* (2.37)	-0.0153* (-2.05)	0.0034** (2.98)	-0.0431** (-5.31)
1997 – 1999	0.0023 (1.13)	-0.0401** (-4.78)	0.0015 (0.69)	-0.0469** (-5.78)	0.0024 (1.12)	-0.0182 (-1.16)	0.0028 (1.58)	-0.0256* (-2.43)
2000 – 2002	0.0016 (0.77)	-0.0243** (-2.95)	0.0018 (0.81)	-0.0393** (-4.35)	0.0022 (1.04)	0.0061 (0.57)	0.0012 (0.56)	-0.0120 (-0.75)

Table 8
Principal Component Analysis

This table presents a principal component analysis of returns and the investor sentiment measure of individuals (ANIS) 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 ANIS of individuals) 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 ANIS of individuals and institutions. 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	0.1330	0.0295	0.2189	0.2807
	Real Std.	0.0081	0.0030	0.0099	0.0101
	Sim. Mean	0.0121	0.0118	0.0580	0.1119
	Diff.	0.1209	0.0177	0.1609	0.1688
ANIS	Real Mean	0.0254	0.0217	0.0967	0.1642
	Real Std.	0.0020	0.0015	0.0037	0.0044
	Sim. Mean	0.0121	0.0118	0.0580	0.1119
	Diff.	0.0133	0.0099	0.0386	0.0523

Panel B: Percentage of Variance 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 1-5	Real	0.0880	0.0903	0.0838	0.0840	0.0889	0.0989	0.1026	0.1017	0.1144	0.1550
	Diff.	0.0320	0.0343	0.0278	0.0280	0.0330	0.0429	0.0465	0.0457	0.0583	0.0991
PC 1-10	Real	0.1572	0.1570	0.1483	0.1509	0.1550	0.1629	0.1677	0.1675	0.1808	0.2266
	Diff.	0.0491	0.0489	0.0402	0.0428	0.0472	0.0548	0.0596	0.0594	0.0727	0.1188