Can Individual Investors Beat the Market?

Joshua D. Coval, David Hirshleifer and Tyler Shumway*

December 2002

Abstract

We document strong persistence in the performance of trades of individual investors. Investors classified in the top 10 percent place other trades that on average earn excess returns of 15 basis points per day. A rolling-forward strategy of going long firms purchased by previously successful investors and shorting firms purchased by previously unsuccessful investors results in excess returns of 5 basis points per day. These returns are not confined to small stocks nor to stocks in which the investors are likely to have inside information. Our results suggest that skillful individual investors exploit market inefficiencies to earn abnormal profits, above and beyond any profits available from well-known strategies based upon size, value, or momentum.

*We thank Malcolm Baker, Randy Cohen, Laura Frieder, Mark Grinblatt, Mark Kamstra, Terry Odean, and seminar participants at the Atlanta Federal Reserve Bank, Harvard Business School, UCLA, the University of Florida, the University of Georgia, the University of Maryland, the University of Rochester, Virginia Tech, and the 2002 Western Finance Association Meetings for helpful comments and discussions, and Greg Allenby for his generous help with the data used here. Coval is at the Harvard Business School *(jcoval@hbs.edu)*, Hirshleifer is at the Fisher College of Business at Ohio State University *(hirshleifer_2@cob.osu.edu)*, and Shumway is at the University of Michigan Business School *(shumway@umich.edu)*. Shumway acknowledges financial support from the Bank One Corporation Assistant Professorship and the NTT fellowship of the Mitsui Life Center.

Financial economists have debated the efficient market hypothesis (EMH) for decades. In most formal tests, econometricians measure the profitability of specified strategies designed to exploit possible mispricing. Such tests are good at identifying whether a particular strategy represented a potential profit opportunity. However, such tests may fail to identify a market inefficiency if the econometrician has chosen a strategy that is unrelated to investor biases.

Furthermore, such tests do not identify whether the proposed strategies were actually exploited, or even recognized by sophisticated traders at the time of their investments. Often econometricians employ financial, econometric, and computational databases and technology that were not readily available to traders at the time.[1] Thus, some authors have suggested that the traditional efficient market hypothesis is too strong, and have proposed milder notions of efficiency that reflect reasonable constraints on the ability of intelligent investors to process information (see, e.g., the 'adaptive efficiency' argument of Daniel and Titman (1999)). In a similar spirit, we offer an approach for evaluating market efficiency that is based upon not just whether profit opportunities are in principle available, but whether some set of real investors have demonstrated abnormal skill in generating actual trading profits. Our evidence suggests that even this milder form of market efficiency is violated.

According to the EMH, investors' risk-adjusted performance should be random: unless they possess relevant private information, they should neither consistently beat the market nor should they, in the absence of transaction costs, consistently underperform the market. If those individual investors who have performed abnormally well in the past continue to perform abnormally well in the future by an amount that is not explained by mere chance, market efficiency may be violated. However, it is possible that investors whose abnormal performance persists are exploiting superior private information about fundamentals, rather than superior skill at identifying and exploiting market mispricing. Therefore, in addition to testing for performance persistence, we test whether abnormal performance is persistent among mid- to large-cap stocks, which presumably have less information asymmetry. In addition, we examine whether persistence of abnormal performance is confined to the few stocks in which investors transact frequently (and thus are more likely to possess private information) or whether it is present even in the companies in which our investors transact only once. Finally, we examine whether some investors persistently *underperform*; under an information story, such a finding would, implausibly, require investors to have superior information and then trade the wrong way.[2]

[1]For example, although historical price data is in principle public information that is costlessly available to all, in reality such data is costly, as evidenced by the fees researchers pay for access to electronic databases.

[2]Investors who systematically underperform might at first seem to present a paradox for the market inefficiency theory as well. On the one hand they display a superior skill at identifying inefficient prices,

An advantage of using individual trader performance to evaluate market efficiency is that this approach vastly expands the set of strategies indirectly being tested and the set of observable variables on which trades potentially may be conditioned. The burden is no longer on the econometrician to identify and measure the variables used in constructing trading strategies. Our approach obviates the need to run thousands of diverse tests and then speculate as to how to discount the statistical significance of these tests to adjust for datamining.

At first glance, it would seem that a search for evidence that individual traders outperform the market is not very promising. Individual traders are often regarded as at best uninformed, at worst fools. The noise trader approach to securities markets, for example, identifies individual investors as generating demands that are generally driven by liquidity or psychological considerations unrelated to the information about underlying security values (see, e.g., Black (1986), De Long et al. (1990) and Lee et al. (1991)). Several studies have documented the poor average performance of individual traders relative to the market and to institutional traders. For example, individual traders appear to trade too much, maintain underdiversified portfolios, and hold onto losing positions for too long.[3]

However, not all individual traders do poorly in their investments. Indeed, as Barber and Odean (2000) note, the top-performing quartile of the individual accounts in their dataset outperform the market on average by 0.5 percent per month. Their findings raise the fascinating question of whether some individual investors have superior skill, and are able to profit thereby. Of course, given the low diversification of many of the accounts, one expects substantial cross-sectional variation in account performance by chance even if there are no differences in skill. The central question we address in this paper, therefore, is whether individual investors who earn profits on their trades are merely lucky, or whether some of them are indeed skillful.

The issue of whether superior performance persistence exists has been examined most extensively for mutual funds. Most studies of mutual funds find that the abnormal performance of the average funds lags that of the overall market.[4] Similarly, only limited evidence exists suggesting that those funds that outperform can be expected

and on the other hand they trade in the wrong direction in association with the inefficiency. The puzzle is resolved by recognizing that if members of a group of investors are subject to common misperceptions, their total trading as a group will move prices in a direction adverse to their desired trades. Thus, an inefficient markets story will involve not just smart traders who make money by exploiting inefficiency, but foolish traders who lose money in the processing of generating the inefficiency.

[3]See, e.g., Blume and Friend (1975), Ferris, Haugen, and Makhija (1988), Odean (1997), Odean (1998), Barber and Odean (2000), Grinblatt and Keloharju (2001). Cohen et al. (2001) and Hirshleifer et al. (2001) also report evidence suggesting that individual investors trade unprofitably in response to cash flow news and earnings announcements.

[4]See Carhart (1995), Malkiel (1995), Chevalier and Ellison (1999), and Daniel et al. (1997). The exception is Wermers (2000) who, after controlling for cash drag, finds positive average excess returns.

to continue to do so in the future.[5]

While few would expect individual traders to be, on average, better informed than mutual fund managers, there are compelling reasons to believe that individual traders are better positioned to *exploit* a given informational advantage. First, individual traders almost always trade smaller positions than professional traders. As a result, the pressure that their trades impart on prices is likely to be much less. This makes them far better positioned to trade using strategies that exploit smaller or shorter-term deviations from fundamental values. Second, individual traders are less constrained than mutual funds to hold a diversified portfolio or to track the market or a given benchmark.

From the standpoint of the researcher seeking to detect performance persistence, the transaction-level datasets of individual accounts are far superior to mutual fund data, which are generally available only at a quarterly frequency. If the profit opportunities exploited by investors are transient, then tests that rely on transactions reported at a quarterly frequency are considerably disadvantaged.

Using the transaction record of over 110,000 accounts at a major discount brokerage, we conduct a variety of tests of individual performance persistence. First, to examine whether traders exhibit statistically significant persistence in their performance, we divide the sample in half and examine the correlation in the excess performance of an account's trades between the two halves. While we measure an investor's performance at a variety of horizons, the magnitude of excess returns earned on his trades drops off sharply after five to ten trading days. Thus, we typically focus on the average return he earns on his trades over the subsequent week as our measure of performance. Moreover, to ensure that our measures of trader profits are not driven by any price pressure created by the trade, in most of our tests we wait a day before beginning our measurement of returns. We also measure performance as the p-value of a one-sided test of whether the mean return earned by the account over the sample half is positive, and then compute the correlation of the probability across sample halves.

There is some debate as to the proper way to adjust for risk in evaluating investment performance. It is common to use factor benchmarks such as the 3-factor model of Fama and French (1992), or to control for characteristics such as book/market, size,

[5] Lehman and Modest (1987), Grinblatt and Titman (1992), Hendricks et al. (1993), Goetzmann and Ibbotsen (1994), Brown and Goetzmann (1995), Elton et al. (1996), and Wermers (1996) all document evidence of persistence in mutual fund performance. However, Carhart (1992, 1997) and Wermers (2000) find that most of the persistence can be explained by persistence in mutual fund expense ratios and momentum in stock returns. Baks et al. (2001) employ a Bayesian approach to detect managers with positive expected alphas.

or momentum (see, e.g., Daniel et al (1997)). However, each of these characteristics can be a proxy for market mispricing, as can the factor loadings in factors that are generated by portfolios based upon such characteristics. Thus, such benchmarks can absorb some of the abnormal performance that the test is trying to measure (for recent discussions, see e.g., Loughran and Ritter (2000) and Daniel, Hirshleifer and Teoh (2001)). We use as our benchmarks the Fama French (1992) 3-Factor Model and the Daniel et al. (1997) characteristics-based adjustment. Readers who regard these benchmarks as reflecting models of risk can regard this as a fair test for abnormal performance of investors. Readers who regard these benchmarks as capturing mispricing may regard our tests as showing the ability of sophisticated individual investors to earn profits above and beyond any gains or losses they earn based upon other well-known effects such as value, size and momentum.

We find that trader performance, whether measured as an average excess return or as a p-value, is consistently correlated across the two sample halves. The correlations are approximately 10 percent, and differ from zero with high significance. The positive correlation survives a variety of robustness checks, including comparing even and odd quarters, removing the smallest one-third of the sample of CRSP stocks, and removing an account's trade in a stock that the investor has traded previously (a possible indication that the individual has access to inside information). To see whether the persistence is due to stock selection ability or to trader ability (or negative ability) to time the market, we recompute performance replacing individual stock returns with overall market returns. In this way, we can test whether individuals consistently purchase stocks before the market rises. When individual stock returns are replaced with the overall market return in this way, we find only very limited evidence that accounts have persistence in their ability to time the market.

Next, to examine the economic significance of performance persistence, we classify each trade of each account according to the performance of *all other* trades placed by that account. In this way, we maximize our precision in classifying an account according to the investor's degree of skill. We call this a complementary image procedure, in analogy to the phenomenon in visual perception in which a figure is identified through its contrast with a complementary background. The complementary image procedure uses ex post information, and therefore does not provide a feasible trading strategy. However, the procedure does strictly quarantine the classification stage for a trade from the performance of that trade itself. This is essential, as otherwise we would induce a mechanical bias in which trades classified as coming from skillful traders are more likely to be profitable, even if there were no persistence. The classification for each trade is made by sorting according to the one-sided p-value in the test of the hypothesis that all other trades by that account have a positive rather than zero (possibly adjusted) mean return. Finally, we form decile portfolios and calculate the average of the returns earned on each trade during the subsequent week.

The difference between the returns of the top and bottom portfolios is striking. Trades in the top decile earn excess returns between 12 and 15 basis points per day during the following week. Trades in the bottom decile lose between 11 and 12 basis points per day. Over the entire holding period of the trade[6], the average trade in the top decile earns five basis points per day whereas the average trade in the bottom decile loses one basis point per day. The results are highly statistically significant and are invariant to using a factor- or characteristic-based risk/mispricing adjustment. The results are also robust to the removal of the smallest third of CRSP stocks and to the removal of trades in stocks traded more than once by the account. This suggests that the top traders earn economically large returns from their stock selection skills in a wide range of companies.

Finally, to investigate whether the information contained in account trading behavior offers profitable trading opportunities to those with access to this information, we construct and test the profitability of a trading strategy that uses only ex ante data at each point in time. On each date, we rank all traders who have traded at least 25 times up to that point. We rank them according to the one sided p-value for testing whether the mean return is positive. Next, for each quintile of traders, we construct a portfolio consisting of all stocks purchased by traders in that quintile during the previous week, weighted by the stocks' market values. We then construct a zero-cost trading strategy that is long the portfolio of the top quintile and short the portfolio of the bottom quintile. The returns of this strategy are then benchmark-adjusted using factor and characteristic-based adjustments.

Using the one-week holding period, this strategy earns excess returns of 5 basis points per day, or 13.7 percent per annum. Again, the results are robust to removing small stocks and trades in stocks an account has traded previously. The strategy is also assessed using a one-day and a one-month holding period. Using the one-day holding period, the return is 7 basis points per day, while the daily return of one-month holding period is indistinguishable from zero.

The paper proceeds as follows. In Section I we describe the data. Section II reports the results of the across-sample correlation tests and the estimation of returns of trades of accounts ranked according to the performance of their other trades. Section II also discusses the results of a trading strategy designed to exploit information contained in the trades of well- and poorly-performing accounts. Section III provides some interpretation of our evidence and Section IV concludes.

[6]The median holding period for active accounts is 199 days.

I. Data

This paper studies a dataset provided by a large discount brokerage firm on the trades placed by 115,856 accounts from January 1990 through November 1996.[7] Table 1 reports summary statistics. Since many of our tests focus on the 16,668 accounts that placed at least 25 trades during the sample period, we also report statistics for this subset. Overall, the average account placed 15 purchases of an average size of $8,599 in 9.2 different companies, though, not surprisingly, the median and standard deviation indicate substantial right-skewness in the distribution. The median account placed six purchase trades in four different companies at an average value of $4,369. For the subset of purchases that were later (at least partially) sold in our sample period, the holding period for the average (median) account was 378.11 (293) days.

The lower panel of Table 1 reports summary statistics for accounts that traded a minimum of 25 times during the sample period. The average account traded an average size of $10,301 in 66.4 trades in 36.8 different companies. Again the median figures are somewhat lower, with the median account trading 43 times in 26 different companies at an average value of $5675. Not surprisingly, the average holding period for the active accounts, 244.33 days, is considerably lower than the overall figure.

The key challenge we face in this inquiry is that we only have six years of data. This severely limits our power to assess the abilities of the individuals in our dataset. We therefore design our procedures to maximize power. To the extent that the profitable trading opportunities available to skillful individuals are short-lived, the variability of the unexpected component of a trade's return will account for an increasing fraction of total return variability as the holding period grows. Thus, for short-lived profit opportunities (such as event-related trading), any inference about abnormal expected returns is likely to be considerably easier when the focus is on shorter horizons, which the transactions data allow. With this in mind, we typically focus on the returns individuals obtain from their trades during the week that follows their trades.

An additional step we take to mitigate the inference problem is that, for most of the tests, we restrict our attention to accounts that have traded at least 25 times. While this removes more than 99,000 accounts from consideration, it ensures that for each account we study we have sufficient data to estimate trading profits with some accuracy.

For the factor-based risk adjustment, we estimate time-series regressions of the

[7] While the discount brokerage data include files with information about account holdings and trades, we only use the file that records account trades. There are 126,488 accounts in the trades file, but only 115,856 accounts had at least one purchase during the sample period.

return on each stock net of the Treasury-bill rate on several factors: the excess of the CRSP value-weighted market over the Treasury-bill rate, a size factor, a book-to-market factor, and one lag of each of these factors to adjust for the possibility of non-trading biases. For each month, we estimate these regressions using daily data for the calendar year finishing at the end of the month. We take the excess return of each stock during the month in question to be the sum of the intercept in these regressions plus the error term, or equivalently, the realized return minus the sum of the factor loadings times the realized value of each of the factors. Both the size and the book-to-market factors are calculated by taking the equal weighted average of the top three value-weighted size and book-to-market decile portfolio returns and subtracting the average of the bottom three decile portfolio returns.

For the characteristic-based risk adjustment, we follow a procedure similar to the approach used by Daniel et al. (1997) (DGTW). Specifically, we rank each stock into quintiles based on its market capitalization at the end of the previous month, its book-to-market ratio based on its most recently announced book equity value (lagged by at least 60 days to ensure public availability) and its momentum status. To determine the momentum status of each stock, we sort stocks each month into deciles based on their return over the previous three months. Any stock that has been in the highest decile during one of the past three months is considered a winner stock, while any stock that has been in the lowest decile is considered a loser stock. A stock that was in the both groups during the last three months is assigned to it most recent classification. Stocks that are neither losers nor winners are designated as neither, resulting in three possible momentum categories. Combining these three momentum categories with five size and five book-to-market categories results in seventy-five possible classifications for each stock. We calculate daily equal-weighted average returns for each of these seventy-five stock classifications, taking the characteristic-adjusted return of a particular stock to be its realized return minus the average return to a stock with its classification.

Our tests for stock selection ability focus solely on the performance of trades that initiate or expand existing positions in companies. We ignore all sales of shares. Our rationale for this is that we expect that sales are often not strongly driven by specific analysis of or private information about the sold stock. Liquidity needs, or the reversing of a position taken long ago in order to diversify may motivate many sales. Such sales may also be motivated by a desire to move into other firms expected to outperform the market. Since few accounts place short-sale trades, we ignore them as well. In contrast, we regard the purchase of a particular stock (as contrasted with the alternative of investing in a mutual fund) as a relatively clear indication that the investor expects that stock to outperform the market.

Using the excess return series, we then calculate the average daily returns earned during the days that immediately follow a given purchase. Specifically, if an account purchases shares in a company on a particular date, for our tests that use a weekly horizon, we calculate the average daily return in that company during the next five trading days. Returns calculated using a one-day horizon use only the subsequent trading day's return, whereas those using a one-month horizon use the subsequent 20 trading days' returns. We typically do not include the trading day's return in our calculations to ensure that any price pressure created by the purchase– particularly of small companies– does not distort our results in favor of the trader.

II. Results

A. Return Correlations

We begin with a simple correlation test for persistence in account performance in which average account returns are compared across the two sample halves. To be considered in our calculations, we require accounts to have traded at least 25 times during the first half of our sample. The six year sample is split at the end of the fourth year to ensure that roughly an equivalent number of accounts have traded at least 25 times in both sample halves. To make sure our results are not contaminated by, for example, individuals who trade more frequently if their performance is good, we place no minimum trade restriction on the second half of the sample. We then calculate the correlation in mean excess returns across the two sample halves.

We calculate correlations of raw and excess returns across the two sample halves. Risks are adjusted using the Fama and French (1992) 3-Factor Model and using DGTW characteristic portfolios. To account for the fact that average returns are calculated with varying precisions across accounts (and across sample halves), we calculate two additional return correlations. The first compares the ratio of the mean return to the return's standard deviation (a return/risk ratio) across the two sample halves. For the second, we compute the (one-sided) p-value associated with the t-statistic of the hypothesis that a given account's excess return is positive during the sample half. We then calculate the correlation in the account p-values across the two sample halves. Finally, as a robustness check, we also calculate the correlation in returns obtained in even and odd quarters. The results for the three return correlations are reported in the first three panels of Table 2.

The correlation in performance across the sample is consistently around 10 percent and highly statistically significant. The results are significant for both correlation calculations and are largely invariant to whether or how we adjust for risk. The correlations are also consistently positive and significant for each of our three performance measures: the simple average returns, the return-risk ratios, and the p-values. The correlations are also robust to splitting the sample into even and odd quarters instead of halves. Thus, the results of Table 2 provide the first evidence of persistence in the performance of individual traders.

Lastly, to see whether the persistence is due to trader ability (or negative ability) to time the market, we recompute our tests replacing individual stock returns with overall value-weighted market returns. This tests whether individuals consistently purchase stocks before the market rises. When individual stock returns are replaced with the overall market return, most of the evidence of performance persistence disappears. However, the p-value correlations retain most of its significance, suggesting perhaps some persistence in market timing. Overall, though, it appears that the persistence in the performance of individual traders' trades come primarily from stock selection and not market timing.

B. Performance Classification of Traders

While the above results indicate clear persistence in trader performance, they do not provide an economic measure of the level of this persistence. An interesting economic measure, for instance, would be the level of future returns that can be expected of traders identified as among the top 10 percent. Similarly, it would be of interest to know how poorly the bottom 10 percent of traders might be expected to perform in subsequent trades. To investigate the economic significance of persistence in trader performance, we classify traders according to the performance of their trades and then see how well this classification explains the returns of subsequent trades.

Since we only have six years of data, and since many traders have not accumulated a sufficient number of trades to be accurately classified until fairly late in the sample, we employ what we call a complementary image procedure to maximize our power to identify able traders. In this procedure, for each trade placed by a given trader, we use *all other* trades he has placed in our dataset to calculate his average return and the p-value for testing the hypothesis that the mean return of the other trades is positive. That is, to maximize the accuracy of our classification of the trader, we use trades placed in the future as well as those in placed in the past in assessing a trader's ability at a given point in time.

Clearly, this does not represent an implementable trading strategy for someone observing the individual trades as they were occurring. In order to test whether there is a profitable trading strategy based upon mimicking individual investor trades, we later consider a rolling forward procedure. The purpose of the complementary image procedure is not to design a strategy for making profits, but to address the scientific question of whether some traders exhibit superior skill. Although the complementary image procedure uses ex post data, it does not do so in a way that biases the measurement of traders' profits. In predicting the profit for a given week, the procedure omits the profit outcome for that week from the set of data used to identify the set of smart traders. Under the null hypothesis that excess returns are independent over time, this procedure should not generate excess returns.

So, as discussed above, corresponding to each trade of a given trader is an average return of all other trades placed by this trader, and a p-value that this average return is positive. We then sort all trades according to the corresponding average returns and p-values, form deciles, and calculate the average returns of the trades within each decile. We write the average return of trader j in all trades except trade k as

$$\hat{r}_j^k = \sum_{l, l \neq k}^{n_j} \frac{r_{j,l}}{n_j - 1}, \tag{1}$$

where n_j is the number of trades placed by trader j and $r_{j,l}$ is the return earned by trader j on trade l during the subsequent five trading days. Using this, the average return of the trades in decile i can be expressed as

$$r_i = \sum_j \frac{\sum_k r_{j,k} I\left(\hat{r}_j^k \in [\underline{r}_i, \overline{r}_i]\right)}{I\left(\hat{r}_j^k \in [\underline{r}_i, \overline{r}_i]\right)}, \tag{2}$$

where $r_{j,k}$ is the return earned by trader j on trade k and $I\left(\hat{r}_j^k \in [\underline{r}_i, \overline{r}_i]\right)$ is an indicator variable which is one if $\hat{r}_j^k$ is within the limits of decile i, $\underline{r}_i$ and $\overline{r}_i$.

Table 3 reports the average returns of the trades in each decile during the days following the trades' placement. Portfolios are formed according to the p-values that the *raw* returns of each trader's other trades during the five trading days after he places them are positive. We include only trades of traders that have placed at least 25 trades. In column 1, the average raw return in investors' other trades ('classifying trades'), $\hat{r}$, is reported for each decile. The classifying trade average returns range from negative 24 basis points to positive 108 basis points per day during the week after they are placed. Since the returns are in raw terms, most deciles have positive average returns in their other trades.

Column 2 contains the raw return same-day return, $r_{i,t}^*$, earned on the trades from the time the trade is placed the market close. We do not adjust these returns for risk since, without time-and-sales data, it is not apparent how to properly benchmark intraday returns. For most deciles, same-day returns are on average negative. Most of the accounts appear to concede between 25 and 35 basis points on the day their trades are placed. This is likely due to the bid-ask spread component of transaction costs that the traders incur in executing their trades, as discussed in Barber and Odean (2000). Interestingly, however, the final two deciles appear to concede far less in same-day returns costs. The top decile loses only 10 basis points on the day the trade is executed, and decile nine actually earns an average of 17 basis points by the end of trading. Relative to the bottom decile, both figures are highly significant. Thus, even when we focus on same-day returns, the accounts vary widely in terms of their ability to initiate positions at low cost.

Columns 3 through 8 report average daily excess returns for each of the portfolios during the days that follow the placement of their trades. Returns are benchmark-adjusted using the Fama-French 3-factor model. A wide difference in returns exists between the portfolios classified as having low performance and those classified with high performance. While the bottom portfolio loses between 5 and 14 basis points per day during the 5 subsequent trading days, the top portfolio averages gains of between 3 and 24 basis points. The difference between the two portfolios begins at 30 basis points and declines steadily to 8 basis points by the fifth day. Even after two weeks, a significant difference between the two portfolios remains, with the high performance portfolio outperforming the low portfolio by 6 basis points on the tenth trading day.

The final column of Table 3 reports the average daily holding period return, $r_{i,H}$, for trades within each decile. Remarkably, an economically and statistically significant spread of 6 basis points remains between the top and bottom deciles of trades. Most of the spread in holding period returns comes from the top decile. Indeed, the first four deciles all have average returns near zero and only the final two deciles have returns above three basis points per day. This suggests that although individuals may underperform the market over short horizons by trading excessively, over the longer term most performance persistence is concentrated among the individuals with positive ability.

While the results thus far provide a strong indication of persistent differences in the ability of different individuals to select stocks, a variety of potential concerns remain. First, because we have sorted accounts according to the raw returns of their other trades, we may be sorting on their willingness to assume some sort of risk not captured by market covariance, size, book/market, or momentum. This concern suggests that it may be desirable to sort according to a risk-adjusted performance of the accounts.[8]

Second, our risk-adjustment procedure may not be ideal. For instance, those that view momentum as proxying for some unknown risk-factor may be interested in seeing whether our results are robust to the removal of the component of returns that is attributable to a momentum strategy. More generally, it is useful to verify whether abnormal performance is robust to the employment of a characteristic-based adjustment that adjusts for momentum.

Third, our sorting of portfolios according to the p-value that their trades have earned positive profits may be suspect, particularly because we require that they have traded at least 25 times. For example, it could be the case that investors who have done well in the past trade more often because they believe they have ability.[9] Similarly, accounts that have done poorly in the past may be more inclined to trade aggressively to make up for past losses.[10] This may result in a post-selection bias for persistently lucky or unlucky accounts that consequently generate the requisite 25 trades. Thus, despite a likely reduction in our power to identify trader ability correctly, it is useful to rerun our tests using all accounts that have more than one trade, sorting only on returns.

Finally, to determine whether our results are driven by trading in stocks in which investors have private information, we rerun our tests removing all trades in companies in which the investor's account has transacted more than once. It seems unlikely that many investors obtain private information sporadically in a wide range of companies. This restriction therefore focuses the test on whether there is superior skill at identifying and exploiting market mispricing. Each of the above tests and robustness checks are described in Table 4.

The first column of Table 4 reports the average excess return to each decile. The portfolios differ markedly in their average excess returns. Trades placed by accounts whose other trades average returns that are among the bottom 10 percent of all trades lose 4.9 basis points per day during the next five trading days. In contrast, trades placed by accounts ranking in the top decile earn 19.4 basis points per day. When these returns are benchmark-adjusted the picture remains the same. Using the factor-based adjustment, the trades of accounts in the bottom decile lose 12 basis points per day, whereas those of accounts in the top decile earn 15 basis points per day. Both figures are significantly different from zero and their difference, 27.5 basis

[8] Of course, if we have errors in our benchmark, these will induce performance persistence when there is none. The fact that persistence exists when we sort according to raw returns helps alleviate this concern.

[9] See Barber and Odean (2000) for supportive evidence.

[10] Coval and Shumway (2001) document such behavior among market makers in the CBOT US Treasury Bond pit.

points, is highly significant. Because accounts are sorted according to the excess returns of their other trades, this average daily return differential is somewhat higher than the average over the first five trading days in Table 3, reflecting the improvement in accuracy in classifying the accounts. The characteristic-based adjustment results in a slightly lower spread of 22.8 basis points, but one that is still highly statistically and economically significant.

The persistence of the poor performance is notable, since it seems to indicate a special ability to underperform the market. The losses of these investors are far greater than the losses of the average individual investor documented by Odean (1999), and our decile 1 findings further indicate that this poor performance is present even after controlling for momentum. The systematic ability of some individuals to underperform indicates that access to inside information is not the primary source of abnormal performance in our sample.

This raises the question of what does cause these individuals to underperform. By now there are several models of investor psychology and prices in which imperfectly rational investors, in the process of producing market mispricing, on average lose money to more sophisticated 'arbitrageurs' (see, e.g., Hirshleifer (2001) for a review of recent models). Our findings are consistent with individual investors being a heterogeneous group that includes both foolish and sophisticated traders.

Although the portfolios are constructed using ex-post data, the return differentials nonetheless raise doubts about the efficient markets hypothesis prediction that excess returns on an account's trades will be independent draws. Either we have adjusted for risk incorrectly and thus we have induced some cross-sectional correlation in the returns through our measurement procedure, or the accounts have significant dispersion in their alphas.

One possibility, as mentioned earlier, is that our restriction that accounts must have traded at least 25 times over the full sample period is introducing a subtle post-selection bias. To control for this possibility, the fifth column sorts all trades of all accounts according to the average return earned on the account's other trades, adjusted for risk using the 3-factor model. We would expect his classification to be far less precise as a segregator of skillful and lucky investors. Furthermore, it results in extreme portfolios having a bias towards accounts that trade infrequently. Nevertheless, this classification produces an average return differential that is consistent with the previous findings. Accounts in the bottom decile place trades that lose 8.8 basis points per day, whereas accounts in the top decile earn 11 basis points per day during the week following their trades.

A second possibility is that frequent account trading in the same stock generates a correlation in trade returns. To control for this possibility, we reclassify accounts using only trades made in stocks they trade once during our sample. These results are reported in the final column of Table 4. Although the statistical significance declines somewhat (due to the fact that we have 40 percent fewer observations), the overall result is unchanged. Although traders in the bottom decile no longer perform so poorly in subsequent trades, traders in the top decile continue to place trades that earn nearly 10 basis points per day. The spread between the top and bottom deciles, 12.7 basis point per day, remains highly significant.

This finding casts further doubt upon the hypothesis that the abnormal performance we find is due to traders trading on inside information. While it is possible that a subset of the accounts have inside information about a company or two (i.e. in their employer or friend's firm), it seems doubtful that a large number of accounts have access to inside information in a broad set of companies. Finally, to see whether the results are concentrated in small, illiquid stocks, or stocks for which individual investors are likely to be insiders, we rerun our classification using only the largest two-thirds of all CRSP firms. Once again, the results (not reported here) remain essentially the same.

C. An Trading Strategy Based upon Observing Individual Investor Trades

The results thus far indicate that a subset of individual investors have ability of some sort. However, it is not yet clear whether these results offer a trading strategy for an observer to exploit the information contained in the accounts' trades. To investigate the real-time returns offered by individual trade information, we construct zero-cost portfolios that go long all the trades of accounts that have performed well up to the current date and go short all the trades of accounts that have performed poorly up to the current date. As with our earlier tests, to ensure that any price pressure created by trades does not influence our results, we wait until the day after the trade is executed to begin measuring returns.

Since we only have six years of data, and much of this is used to assess traders' performance, our power to detect abnormal returns is somewhat limited. To maximize our power to detect abnormal performance, there is a tradeoff. If we only include trades of accounts with mean returns significantly different from zero, we more reliably focus on the trades of more skillful versus less skillful traders. However, to the extent that, at times, only a limited number of accounts can be classified as unusually good (or bad), such a portfolio will be highly undiversified.[11] Since we only have one thousand days over which to measure our strategy's expected return, such lack of diversification can result in the unexpected component of returns becoming so variable that inference is impossible. On the other hand, if we are lax in our criteria for including accounts in the strategy, a larger fraction of the trades we mimic are from accounts lacking in special skill.

To strike a balance, we only consider accounts that have traded at least 25 times up to the current date, but we sort them into quintile portfolios to ensure that our portfolio is diversified. Furthermore, we only measure the returns to our strategy on days when there are at least 25 stocks in the top and bottom portfolios. Specifically, we rank all accounts that have traded at least 25 times up to the current date by the p-value that their excess return is positive. We then compute value-weighted returns of all the stocks purchased during the last five days by all accounts in each of the performance quintiles. Specifically, the return to portfolio i on date t is calculated as follows:

$$r_{i,t} = \sum_j \frac{MV_{j,t}}{\sum_k MV_{k,t} I_{i,k,t}} r_{j,t} I_{i,j,t}, \tag{3}$$

where MV_j is the market value of firm j on date t, $r_{j,t}$ is the return to firm j on date t, and $I_{i,j,t}$ is an indicator variable which is one if an account in portfolio i has purchased firm j within the holding period preceding date t and zero otherwise.

Using the strategy return defined in equation (3), we calculate the excess return to the strategy that goes long the top quintile and short the bottom quintile. We use a 4-factor model to adjust returns for risk, adding a momentum factor to the Fama French (1992) 3-factor model.[12] When we employ the 4-factor model, we calculate a *raw* return in equation (3) and then regress the difference between the top and bottom portfolio daily return on the four factors. When we benchmark-adjust using the characteristic-based adjustment, equation (3) is calculated using the individual firm characteristic-adjusted returns. The reported results focus on the 4-factor risk adjustment though highly similar results are obtained using the characteristic-based adjustment. To ensure that trading in small, illiquid firms does not drive the results, we remove the smallest (by capitalization) third of all CRSP firms from the sample. Finally, we examine the returns to the trading strategy using three portfolio formation horizons: daily, weekly, and monthly. The results are reported in Table 5.

Beginning with the one-week holding period, the strategy generates excess returns of 5.1 basis points per day. When only a market factor is used to risk-adjust, the

[11] Our reliance on value-weighted portfolios makes any lack of diversification even more pronounced.

[12] We construct our momentum factor by taking the difference of the equal-weighted average returns of the winner portfolio and the loser portfolio identified by our characteristic adjustment algorithm.

returns are 4.4 basis points per day. Both figures are significant at the 5 percent level but not the 1 percent level. If we measure returns at the daily horizon – that is, on the trading day following the trade placement – the results are slightly stronger in economic and statistical terms. The 4-factor excess returns are 6.8 basis points per day and those adjusted using the market factor are 5.6 basis points per day. When we move to the one-month horizon, however, the results essentially disappear, falling below a basis point per day and losing any statistical significance. The disappearance of significance at the monthly horizon suggests that the ability of traders to select stocks that earn abnormal returns may be confined to fairly short horizons. Alternatively, it may simply illustrate how difficult inference becomes as horizons lengthen relative to the horizon of the investors' profit opportunities, so that a trade's alpha declines relative to the unexpected component of returns.

As described above, estimating the returns to a feasible trading strategy based on our data involves a careful balance. If we base our strategy on fewer traders with more extreme past performance, the variability of our results increases. If we base our strategy on more traders in an effort to reduce the variability of our results, the average performance of the strategy declines.

In the tests reported in Table 5, we form top and bottom trader-mimicking portfolios based on the top and bottom quintiles of all ranked traders. We require all ranked traders to have at least 25 previous trades and we require both the top and bottom trader mimicking portfolios to consist of at least 25 stocks on any particular day. For the returns calculated over one week, our requirements mean that out of 1,205 possible trading days, we can only evaluate the returns to our strategy on 1,072 days. If we define the top and bottom trader mimicking portfolios by taking the top and bottom deciles of ranked traders, the strategy return can only be estimated on 945 days. Using the top and bottom deciles, the one week intercepts become 4.4 ($t = 1.74$) basis points for the 4-factor model and 5.1 basis points ($t = 2.14$) for the market model. If we define the top and bottom portfolios as the top and bottom thirty percent of traders, the number of valid return days becomes 1,101. The intercepts become 3.8 basis points ($t = 2.40$) for the 4-factor model and 3.1 basis points ($t = 2.05$) for the market model. If we estimate the regressions reported in Table 5 with a weighted least squares technique that assigns weights to each observation that are proportional to the square root of the number of stocks in the top and bottom portfolios, the intercepts become 4.7 b.p. ($t = 3.1$) for the 4-factor model and 4.9 b.p. ($t = 3.1$) for the four-factor model.

III. Discussion

The above results raise several interesting questions. To begin with, why are the results so much stronger, both economically and statistically, when portfolios are formed using all available trade data? Do the additional datapoints really add that much power? It turns out they do. Their contribution is twofold. First, the additional data gives us far more information with which to classify traders. With the real time trading strategy, a trader can be classified using only data up to the point in time of a given trade. Thus, the accuracy with which a trader is classified improves steadily across time. Conversely, when all data are used, each trade can be classified as if it is the last. This not only allows us to rank each trader far more accurately. It also allows us to use more trades of more traders. For instance, using our minimum of 25 trades, a trader who has placed 27 trades will only have two trades considered for our real time strategy's portfolio. On the other hand, when ex-post data are used in classifying trades, all 27 can count towards portfolio return calculations.

A second question raised by the results is whether they offer the opportunity for essentially riskless trading profits to anyone with access to the account trade data. With only six years of data, the answer is no. While there is a five basis point spread in the real time daily returns of the top and bottom quintiles, considering that the positions are turned over every five trading days, transaction costs will dominate any excess returns earned by the strategy. This is likely to be the case even though the five basis point spread is generated without trading in the smallest third of all CRSP firms.

On the other hand, if one has access to more than six years of data, account performance can be classified with high precision. So the excess return spread from year six onward is likely to resemble those calculated in Table 4. Unless round trip trading costs exceed one percent, it is likely that the excess returns will remain significantly positive.

A final question is whether these returns are generated by trading on inside information. While it would be surprising if small accounts at a discount brokerage had access to inside information, it is possible that a subset of the accounts have access to important information about their employer or that of a spouse or friend. However, the fact that the results are not significantly altered when we remove all trades placed by a given account multiple times in the same company suggests that this is not the source of our findings. Moreover, the abnormal performance of the better accounts is also essentially unchanged when we remove trades in the smallest third of the firms in the sample. It seems highly doubtful that a large number of accounts have access to inside information in a broad set of medium and large capitalization companies.

IV. Conclusion

Individual investors have been much maligned in the finance literature in recent years. This paper shows that not all individual investors are foolish. Traders that can be classified among the top 10 percent place trades that earn excess returns of between 12 and 15 basis points per day during the following week. These findings are robust to how one adjusts for risk, to the removal of small stocks from the sample, and to the removal of any companies in which the account has traded more than once. On the other hand, individual investors that consistently place underperforming trades also exist. Traders classified among the bottom 10 percent of all traders place trades that can expect to lose up to 12 basis points per day during the subsequent week.

Our finding that some individual investors have superior investment skills, and that others systematically underperform, suggests a new perspective on the issue of whether on average individual traders foolishly trade too much. As discussed earlier, previous studies have shown that individual investors on average lose money in their trades. However, if traders vary widely in terms of their ability to select investments, and if they learn about and develop this ability through trading, it may in fact be rational for some investors to trade frequently and at a loss, in the hope of future gains.[13] It is still discouraging that, net of costs, individual investors as a group lose money on average. On the other hand, if traders who learn that they have unusual ability move their accounts to lower-cost or higher-leveraged trading venues (e.g. options markets), evidence drawn solely from stock trades may focus on those investors who are still in the process of learning– either how to trade, or about whether they are good traders. (Any such tendency to change venues would mitigate the returns obtainable by mimicking the trades of smart traders in our stock-trading sample, which suggests that true skill differences may be even greater than our estimates.)

Finally, this evidence does not support the efficient market hypothesis. The ability of individual traders at a discount brokerage to select outperforming companies is not confined to small firms or only a few firms in which the traders transact frequently; and some investors persistently trade so as to underperform. These findings suggest that investors' persistent abnormal performance is not derived primarily from trading on

[13] As mentioned earlier, those investors who have superior ability at *losing money* relative to the market may be individuals who are, in equilibrium, contributing to the *creation* of market inefficiencies. In principle such an investor has a clear opportunity to learn how to make abnormal profits by reversing his trading strategy.

inside information. The broad ability of some individual investors to achieve abnormal performance implies a violation of semi-strong form market efficiency. An interesting further question is whether the large brokerage companies are aware of the value of the information contained in their customers' trades.

References

Baks, Klaas P., Andrew Metrick, and Jessica Wachter, 2001, "Should investors avoid all actively managed mutual funds? A study in Bayesian performance evaluation," *Journal of Finance*, 56, 45-85.

Barber, Brad, and Terrance Odean, 2000, "Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors," *Journal of Finance*, 55, 773-806.

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

Blume, Marshall and Friend, Irwin, "The Asset Structure of Individual Portfolios with Some Implications for Utility Functions," *Journal of Finance*, 30, 585-604.

Brown, Stephen J., William N. Goetzmann, Roger Ibbotson, and Stephen Ross, 1992, "Survivorship bias in performance studies," *Review of Financial Studies* 5, 553-580.

Carhart, Mark M., 1995, "Survivorship Bias and Mutual Fund Performance," Ph.D. thesis. Graduate School of Business, University of Chicago.

Carhart, Mark M., 1997, "On persistence in mutual fund performance," *Journal of Finance*, 52, 57-82.

Chevalier, Judith, and Glenn Ellison, 1999, "Are some mutual fund managers better than others? Cross-sectional patterns in behavior and performance," *Journal of Finance* 54, 875-900.

Choe, Hyuk and Kho, Bong-Chan and Stulz, René M., 2000, "Do domestic investors have more valuable information about individual stocks than foreign investors?" Ohio State University, Fisher College of Business Dice Center WP 2000-21

Cohen, Randolph B., Paul A. Gompers, and Tuomo Vuolteenaho, 2001, "Who Underreacts to Cash-Flow News? Evidence from Trading between Individuals and Institutions," Working Paper, Harvard University.

Coval, Joshua D. and Tyler Shumway, 2001, "Do behavioural biases affect prices?", Working Paper, University of Michigan.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, "Measuring mutual fund performance with characteristic based benchmarks," *Journal of Finance* 52, 1035-1058.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, "A theory of overconfidence, self-attribution, and security market under- and overreactions," *Journal of Finance*, 53, 1839-1886.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 2001, "Overconfidence, Arbitrage, and Equilibrium Asset Pricing," *Journal of Finance*, 56, 921-965.

Daniel, Kent D. and Sheridan Titman, 1999 "Market Efficiency in an Irrational World," *Financial Analysts' Journal*, 55(6), November/December, (1999):28-40

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, "Noise trader risk in financial markets," *Journal of Political Economy*, 98, 703-738.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 1996, "The persistence of risk-adjusted mutual fund performance," Journal of Business 69, 133-157.

Fama, Eugene F. and Kenneth R. French, 1992, "The Cross Section of Expected Returns," *Journal of Finance*, 47, 427-465.

Ferris, S. P., R. A. Haugen and A. K. Makhija, "Predicting Contemporary Volume with Historic Volume at Differential Price Levels: Evidence Supporting the Disposition Effect," *Journal of Finance*, 43(3), (1988):677-697

Grinblatt, Mark and Matti Keloharju, 2001, "How Distance, Language and Culture Influence Stockholdings and Trades," *Journal of Finance*, 56, 3, (2001):1053-1073

Goetzmann, William N., and Roger G. Ibbotson, 1994, "Do winners repeat? Patterns in mutual fund performance," *Journal of Portfolio Management* 20, 9-18.

Grinblatt, Mark, and Sheridan Titman, 1992, "The persistence of mutual fund performance," *Journal of Finance* 47, 1977-1984.

Hendricks, Darryl, Jayendu Patel, and Richard Zeckhauser, 1993, "Hot hands in mutual funds: The persistence of performance 1974-1988," *Journal of Finance* 48, 93-130.

Hirshleifer, David, 2001, "Investor Psychology and Asset Pricing," *Journal of Finance*, 64, 4, 1533-1597.

Hirshleifer, David, James Myers, Linda Myers and Siew Hong Teoh, 2002, "Do Individual Investors Drive Post-Earnings Announcement Drift?" Fisher College of Business.

Lee, Charles, M.C., Andrei Shleifer, and Richard H. Thaler, 1991, "Investor sentiment and the closed-end mutual funds," *Journal of Finance*, 46, 75-109.

Lehman, Bruce N., and David Modest, 1987, "Mutual fund performance evaluation:

A comparison of benchmarks and a benchmark of comparisons," *Journal of Finance* 42, 233-265.

Loughran, Tim and Jay R. Ritter, 2000, "Uniformly Least Powerful Tests of Market Efficiency," *Journal of Financial Economics*, 55, 361-389.

Malkiel, Burton, 1995, "Returns from investing in mutual funds 1971 to 1991," *Journal of Finance* 50, 549-572.

Odean, Terrance, 1998, "Are Investors Reluctant to Realize Their Losses?" *Journal of Finance*, 53, 1775-1798.

Odean, Terrance, 1999, "Do Investors Trade Too Much?" *American Economic Review*, 89, 1279-1298.

Wermers, Russ, 1997, "Momentum investment strategies of mutual funds, performance persistence, and survivorship bias," Working paper, University of Colorado.

Wermers, Russ, 2000, "Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses," *Journal of Finance* 55, 1655-1695.

Table 1: Summary Statistics

Table 1 reports summary statistics for the entire set of accounts and for the subset of accounts that have traded at least 25 times during the sample period. Average holding period is the average holding period for an account of all purchases that were sold later in the dataset. For consecutive buys or sells in a given company, we calculate the time between the last purchase and the first sale.

Variable (per account)	Mean	Median	St. Dev.	Min.	Max.
Full Sample ($n = 115,856$)					
Number of Purchases	15.06	6	38.43	1	3,167
Average Dollar Value	8,599	4,369	28,031	0.1	6,011,360
Number of Different Firms Purchased	9.20	4	19.68	1	1,523
Number of Purchases Sold Later	9.96	4	25.03	0	2,209
Average Holding Period (days)	378.11	293	321.34	0	2,100
Accounts with at least 25 trades ($n = 16,668$)					
Number of Purchases	66.40	43	83.61	25	3167
Average Dollar Value	10,301	5,675	18,071	50	692,524
Number of Different Firms Purchased	36.83	26	41.16	1	1,523
Number of Purchases Sold Later	32.92	22	48.60	0	2,209
Average Holding Period (days)	244.33	199	192.37	0	1,870

Table 2: Correlation Tests of Performance Persistence

Table 2 reports correlations across the two sample halves between in the average returns earned on firms purchased by a given account during the five trading days that follow the purchase. The correlations of sample halves split the sample in half at the end of the fourth year and calculate the correlation in performance across the two sample halves. Only accounts with at least 25 trades during the first four years are included in the calculations. The even/odd quarter correlations divide all trades into those that occur during the first and third quarter of the year and those that occur during the second and fourth. Only accounts with at least 25 trades in odd quarters are considered in the calculations. The p-values are calculated using a t-distribution and a t-score that average returns are positive. The 3-factor risk-adjusted return correlations regress returns on daily realizations of the SMB, HML and RMRF factors. The DGTW characteristic-adjusted returns subtract from a given firm's daily return the daily return to the equivalent size, book-to-market, and momentum portfolio. The market-timing returns replace the daily risk-adjusted return of a given firm with the corresponding daily return of the value-weighted market portfolio. P-values are in parentheses.

	Sample Halves		Even/Odd Quarters	
Variable	Pearson	Rank Order	Pearson	Rank Order
Raw Returns				
Mean Return	0.055	0.091	0.060	0.079
	(0.005)	(0.000)	(0.004)	(0.000)
Mean Ret. / StDev.	0.097	0.109	0.093	0.094
	(0.000)	(0.000)	(0.000)	(0.000)
p-value	0.114	0.120	0.098	0.107
	(0.000)	(0.000)	(0.000)	(0.000)
3-Factor Risk-Adjusted Returns				
Mean Return	0.069	0.092	0.122	0.108
	(0.000)	(0.000)	(0.000)	(0.000)
Mean Ret. / StDev.	0.090	0.094	0.119	0.102
	(0.000)	(0.000)	(0.000)	(0.000)
p-value	0.101	0.103	0.099	0.101
	(0.000)	(0.000)	(0.000)	(0.000)
DGTW Characteristic-Adjusted Returns				
Mean Return	0.089	0.099	0.021	0.033
	(0.001)	(0.000)	(0.475)	(0.265)
Mean Ret. / StDev.	0.091	0.109	0.049	0.052
	(0.001)	(0.000)	(0.093)	(0.079)
p-value	0.112	0.111	0.070	0.071
	(0.000)	(0.000)	(0.018)	(0.016)
Market-Timing Returns				
Mean Return	0.009	0.012	0.006	0.009
	(0.572)	(0.437)	(0.723)	(0.576)
Mean Ret. / StDev.	-0.016	0.007	0.057	0.021
	(0.345)	(0.679)	(0.000)	(0.208)
p-value	0.029	0.050	0.039	0.073
	(0.082)	(0.003)	(0.022)	(0.000)

Table 3: One-Day Portfolio Returns: Complementary Image Procedure

Table 3 reports the average daily return of trades that have been sorted into deciles according to the assessed ability of the trader. Portfolios are formed according to the raw returns of each trader's other trades during the five trading days after he places them. These portfolios only include trades of traders that have placed at least 25 trades. Column 1 reports the average raw return in traders' other trades for each decile, $\hat{r}$. Future one-day returns are calculated for each trading day after the trade is executed. Column 2 reports the average raw return earned from the time the trade is executed to the same-day close, $r^*_{i,t}$. Columns 3 through 8 report the risk-adjusted returns of the trade on the trading day one, two, three, four, five, and ten trading days after the trade is placed (i.e. $r_{i,t+n}$ where i is the portfolio number and n is the number of trading days ahead). Column 9 reports the average daily return calculated over the holding period of the given trade, $r_{i,H}$. Returns are risk-adjusted using a 3-factor Fama/French model. T-statistics are in parentheses.

Portfolio	$\hat{r}_i$	$r^*_{i,t}$	$r_{i,t+1}$	$r_{i,t+2}$	$r_{i,t+3}$	$r_{i,t+4}$	$r_{i,t+5}$	$r_{i,t+10}$	$r_{i,H}$
1 (low)	-0.0024	-0.0029	-0.0006	-0.0013	-0.0009	-0.0014	-0.0005	-0.0002	-0.0001
2	0.0012	-0.0035	0.0000	-0.0003	-0.0007	-0.0008	-0.0005	-0.0002	0.0000
3	0.0026	-0.0026	0.0003	-0.0002	-0.0004	-0.0004	-0.0003	-0.0001	0.0001
4	0.0030	-0.0033	0.0002	0.0000	0.0000	-0.0006	-0.0002	0.0001	0.0001
5	0.0045	-0.0031	0.0008	0.0005	0.0003	-0.0003	-0.0002	0.0002	0.0003
6	0.0046	-0.0023	0.0006	0.0000	0.0002	-0.0002	0.0003	0.0003	0.0002
7	0.0057	-0.0027	0.0012	0.0005	0.0002	-0.0003	0.0001	0.0003	0.0002
8	0.0075	-0.0026	0.0013	0.0010	0.0005	0.0001	0.0003	0.0002	0.0003
9	0.0073	-0.0017	0.0012	0.0007	0.0004	0.0002	0.0001	0.0001	0.0004
10 (high)	0.0108	-0.0010	0.0024	0.0012	0.0010	0.0006	0.0003	0.0004	0.0005
10-1		0.0019	0.0030	0.0025	0.0019	0.0020	0.0008	0.0006	0.0006
		(13.7)	(20.3)	(18.3)	(14.1)	(15.3)	(6.2)	(5.0)	(8.2)

Table 4: Five-Day Portfolio Returns: Complementary Image Procedure

Table 4 reports the average daily return of trades that have been sorted into deciles according to the assessed ability of the trader. Returns are calculated by averaging the returns of the firm over the five days after it was purchased. The first five columns of numbers are for portfolios that have been formed according to the p-value that the trader's other trades have a positive average return. These portfolios only include trades of traders that have placed at least 25 trades. The first column reports the average daily return in excess of the risk-free rate for each portfolio. The next column reports the standard deviation of this excess return. The next two columns report 3-factor risk-adjusted and DGTW characteristic-adjusted returns. The final two columns report 3-factor risk-adjusted returns for portfolios formed using two alternative sorting procedures. In the second-last column, stocks are sorted according to the raw returns earned by that trader in his other trades, regardless of how few trades the trader has placed. In the final column, only stocks that have been traded once by a given account are included in the calculation of probabilities and returns. T-statistics are in parentheses.

	Excess Returns				Sorted on returns No min. trade number	Stocks traded only once
Portfolio	Mean	Std. Dev.	3-Factor Alpha	DGTW Alpha	3-Factor Alpha	3-Factor Alpha
1 (low)	-0.00049	0.01556	-0.00123	-0.00109	-0.00088	-0.00030
2	0.00020	0.01588	-0.00034	0.00002	-0.00049	-0.00040
3	0.00036	0.01593	-0.00019	-0.00043	-0.00023	-0.00031
4	0.00063	0.01516	-0.00011	0.00028	-0.00014	0.00011
5	0.00088	0.01529	0.00041	-0.00025	0.00031	0.00028
6	0.00082	0.01528	0.00011	0.00029	0.00006	0.00046
7	0.00109	0.01539	0.00030	0.00027	0.00053	-0.00025
8	0.00116	0.01523	0.00100	0.00067	0.00058	0.00054
9	0.00162	0.01482	0.00053	0.00039	0.00045	0.00004
10 (high)	0.00194	0.01511	0.00152	0.00119	0.00110	0.00096
10-1			0.00275	0.00228	0.00198	0.00127
			(21.7)	(18.0)	(17.5)	(7.0)

Table 5: Trading Strategy Returns

Table 5 reports the results of a regression of a trading strategy's return on the daily realizations (and lagged realizations) of four factors: the market return minus the risk-free rate (RMRF), the return of high minus low book-to-market stocks (HML), the return of small minus large stocks (SMB), and the return of a momentum portfolio that is long past winners and short past losers (MOM). Portfolios are constructed by sorting on each date accounts that have traded at least 25 times up to that date based on the p-values of their past trades. Only the largest two-thirds of all CRSP stocks are included in portfolios. For the three holding periods, returns are measured using the first trading day (One Day), first five trading days (One Week), and first twenty trading days (One Month) after the trade is placed. The returns are then value-weighted within each portfolio. The strategy's returns are constructed by going long the top quintile and short the bottom quintile on days when at least 25 stocks are in each quintile. T-statistics are in parentheses.

Factor-Adjusted Returns: High Minus Low Portfolio (Daily Returns)

	Holding Period					
Variable	One Day		One Week		One Month	
Intercept	0.00056	0.00068	0.00044	0.00051	0.00001	0.00003
	(2.5)	(2.9)	(2.3)	(2.5)	(0.1)	(0.44)
$RMRF_t$	-0.08337	-0.14988	-0.06742	-0.05512	-0.01661	-0.01423
	(-2.1)	(-2.9)	(-1.9)	(-1.2)	(-1.29)	(-0.85)
HML_t		-0.20600		-0.03326		-0.00531
		(-3.6)		(-0.7)		(-0.31)
SMB_t		-0.02491		0.01311		0.03023
		(-0.4)		(0.2)		(1.4)
MOM_t		-0.00311		-0.03714		0.03605
		(-0.1)		(-1.1)		(2.82)
$RMRF_{t-1}$		0.12728		0.05398		0.00152
		(2.3)		(1.2)		(0.09)
HML_{t-1}		-0.06553		0.000318		-0.01444
		(-1.0)		(0.0)		(-0.67)
SMB_{t-1}		-0.00197		-0.0291		-0.00264
		(0.0)		(-0.6)		(-0.16)
MOM_{t-1}		0.06540		0.09720		-0.01418
		(1.6)		(2.8)		(-1.12)
n	911	911	1072	1072	1191	1191