View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

A RESEARCH REPORT FROM STOCKHOLM INSTITUTE FOR FINANCIAL RESEARCH

All Guts, No Glory: Trading and Diversification among Online Investors

ANDERS E.S. ANDERSON

NO 25 — JUNE 2004

Stockholm Institute
for Financial Research

Stockholm Institute for Financial Research (SIFR) is a private and independent non-profit organization established at the initiative of members of the financial industry and actors from the academic arena. SIFR was launched in January 2001 and is situated in the center of Stockholm. Magnus Dahlquist serves as director of the Institute. The mission of SIFR is to:

- Conduct and stimulate high quality research on issues in financial economics, where there are promising prospects for practical applications,
- Disseminate research results through publications, seminars, conferences, and other meetings, and
- Establish a natural channel of communication about research issues in finance between the academic world and the financial sector.

The activities of SIFR are supported by a foundation based on donations from Swedish financial institutions. Major contributions have been made by: AFA, Alecta, Alfred Berg, AMF Pension, Brummer & Partners, Carnegie, Handelsbanken, Kapitalmarknadsgruppen, Länsförsäkringar, Nordea, Svenska Fondhandlareföreningen, and Östgöta Enskilda Bank.

Sveriges Riksbank funds a position as visiting professor at SIFR.

SIFR also gratefully acknowledges research grants received from Bankforskningsinstitutet, Föreningsbankens Forskningsstiftelse, Jan Wallanders och Tom Hedelius Stiftelse, Riksbankens Jubileumsfond, and Torsten och Ragnar Söderbergs stiftelser.

Stockholm Institute for Financial Research, Saltmätargatan 19A 11, SE-113 59 Stockholm, Sweden
Phone: +46-8-728 51 20, Fax: +46-8-728 51 30, E-mail: info@sifr.org, Web: www.sifr.org

All Guts, No Glory: Trading and Diversification among Online Investors

Anders E. S. Anderson

Stockholm Institute
for Financial Research

All Guts, No Glory: Trading and Diversification among Online Investors

Anders E. S. Anderson*

April 19, 2004

Abstract

I explore cross-sectional portfolio performance in a sample containing 324,736 transactions conducted by 16,831 investors at an Internet discount brokerage firm during the period May 1999 to March 2002. On average, investors hold undiversified portfolios, show a strong preference for risk, and trade aggressively. I measure performance using a panel data model, and explain the cross-sectional variation using investors' turnover, portfolio size and degree of diversification. I find that turnover is harmful to performance due to fees, and is therefore more predominant among investors with small portfolios. It is argued that the degree of diversification is a proxy for investor skill, and it has a separate and distinct positive effect on performance. These findings are helpful in explaining the overall result that investors underperform the market by around 8.5% per year on average.

Keywords: Investor behavior; performance evaluation; panel data models.
JEL codes: G11, D14, C33.

*Financial support from *Jan Wallanders och Tom Hedelius Stiftelse* is gratefully acknowledged. I thank my advisors, Magnus Dahlquist and Paul Söderlind, for their comments and ongoing support; Peter Englund; Mattias Nilsson; seminar participants at Stockholm Institute for Financial Research (SIFR), and the 2003 Arne Ryde Workshop in Finance at Lund University for helpful comments. This work was completed during my time at SIFR and I would like to express my gratitude for the hospitality I received there. Finally, I am grateful to the online broker who wishes to remain anonymous, and *Finansinspektionen* who provided me with the data. This research would not have been possible without their generosity. Send correspondence to: Anders Anderson, Stockholm Institute for Financial Research, Saltmätargatan 19A, SE-113 59 Stockholm, Sweden, Phone: +46-8-7285133, Fax: +46-8-7285130, E-mail: anders.anderson@sifr.org .

1 Introduction

The stock market boom of the late 1990s is, by most standards, unprecedented in stock market history. Even if many companies in traditional industries were valued at historical highs, the market was given an extra boost by sky rocketing prices stemming from the newly emerging information technology sector. Interest in the stock market surged. A new category of financial intermediaries, namely online brokers, provided low-cost stock market access that was mainly aimed at small investors. The Stockholm Stock Exchange reports that, in 1997, these companies accounted for 1% of the value and 3% of the transactions on the exchange. By 2000, this had risen to 4% of the value and 18% of the transactions. These aggregate figures suggest that online brokers have attracted a clientele of small investors that trade actively. In March 2000 the market peaked, and then entered into a bear market that was to become one of the worst ever. These turbulent times provide the data for this study on the investor performance of a group of online traders.

This paper aims to quantify and measure the relative effects on performance of investment behavior. Since online investing is fairly new, there exists little previous research on the performance of online traders, even though they are predicted to grow in number (see Barber and Odean, 2001b for a survey). Online investors are well suited for studying individual investor behavior, since intermediation between the broker and the investor is kept at a minimum. Individual investors are also more likely to suffer from behavioral biases than investment professionals; overconfident investors are likely to trade more vigorously and hold undiversified portfolios.

The data were made available by an online broker and cover all transactions since the start in May 1999 up to and including March 2002. The 324,736 transactions in common stocks are distributed over 16,831 investors who enter sequentially. The investors are, on average, relatively young, predominantly male, and aggressive traders.

The average turnover rate implies that investors buy and sell their portfolio more than twice a year. The 20% that trade the most turn their portfolios around about seven times a year. In addition, investors are not well diversified. The median number of stocks in the portfolio is two, and 18% of the investors hold only one stock.

The investors in this sample show a preference for risk in general, and technology stocks in particular. The average beta is above 1.4, and among the investors who only hold one stock, 80% choose a stock belonging to the technology industry. In contrast, among the investors who hold four or more stocks, the average technology sector weight is only 55%. Diversification is therefore not only related to idiosyncratic risk, but to industry selection as well. Furthermore, I find evidence that investors who are more highly diversified systematically hold stocks that perform better *within* industries on average.

This effect remains when accounting for differences in portfolio risk and size, and suggests that investment skill is related to the degree of diversification.

I propose a method for retrieving individual, monthly portfolio returns directly from transaction data that is new to the literature of individual investor performance. Portfolio returns are measured relative to *passive returns*, which are the returns of the portfolio investors held at the beginning of the month, and therefore exclude monthly trading. It is found that most of what is lost due to trading can be related to fees, or 32 basis points per month compared with a total of 37. Investors in the top trading quintile lose around 95 basis points per month compared with those who do not trade. This result, however, is primarily driven by investors with small portfolios who are more sensitive to fixed fees. The average investor spends around 3.8% per year of their portfolio wealth on fees; this is more than twice the charge of a standard mutual fund.

Figure 1A shows that the equally-weighted mean of the market-adjusted return is -2.07% per month. In this paper I show how to decompose the return into the three parts discussed: the component due to the choice of industry, intra-industry selection—"stock-picking," and trading.

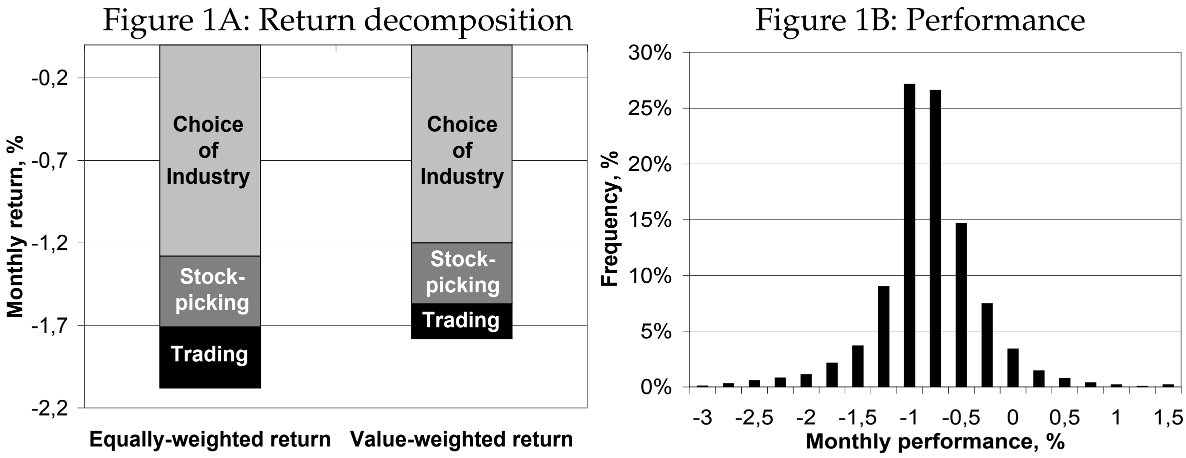
The choice of industry is most important in explaining the return difference to the market, reflecting the heavy tilt toward technology stocks and bad market timing. This may simply reflect investors' preferences for high risk. Risk is likely to be less of a problem when measuring stock-picking ability. The investors lose 43 basis points per month from choosing stocks that underperform any given industry on average. Trading costs are roughly equally important; 37 basis points per month are lost compared with the passive portfolio held at the beginning of the month. The value-weighted return suggests that fees, in small transactions and for small portfolio sizes, drive this result.

In addition to the uni-dimensional effects of trading and stock selection documented above, the main contribution of this paper lies in quantifying the relative importance of these characteristics for investor performance taken together.

Figure 1B shows the frequency distribution of investors' average performance, generated from estimates in a panel regression. The average performance here is a function of investors' turnover, size of portfolio and number of stocks held. It is possible to generate quite substantial cross-sectional variation in abnormal performance from the data by using these characteristics. The average underperformance is 74 basis points (or around 8.5% in annualized terms). The poor performance is likely to be due to the fact that these investors' portfolios are *too small*, that they trade *too much* and are *less* experienced *on average* compared with other stock market participants. An additional percentage increase in turnover hurts investor performance by 1.7 basis points per month. Investors whose portfolio values are twice as high as the sample average gain an additional 11 basis points. Similarly, investors who hold one stock more than average, i.e., four rather than three,

Figure 1: Summary of key results

Figure 1A depicts the equally and value-weighted mean of a decomposition of the monthly market-adjusted return. The bar labelled "Choice of industry" refers to the return difference between the market and the chosen industry. "Stock-picking" measures the deviation from the individual stocks selected and the industry benchmark. "Trading" measures the return difference between the portfolio that includes trading and the portfolio held at the beginning of the month. Figure 1B displays a histogram of the 16,831 investors' monthly abnormal performance generated by the coefficient estimates of panel regression Model V in Table 8.



perform 7 basis points better. The characteristics are also related to risk; investors that are older, women, trade more, and are more diversified all take less systematic risk.

The remainder of the paper is organized as follows. Section 2 discusses the theoretical foundations of trading and stock selection, as well as the previous empirical evidence, within the framework of individual investor behavior. Section 3 presents the transaction data. Section 4 begins by explaining how portfolio returns are retrieved from transactions and then presents the results. Section 5 concludes.

2 Trading and stock selection

This paper links individual investor performance to both trading behavior and portfolio strategies. Explaining the findings by rational behavior is not unproblematic, given the overall poor performance of investors' portfolios.

The first question that arises is: Why do these individuals trade in such vast quantities? The no-trade theorem states that prices fully reflect information and when new information arrives, it is immediately incorporated into prices. If this were the case, there would be no trading at all.

But there may be informational asymmetries that drive trading. Grossman and Stiglitz (1980) derive an equilibrium from when the marginal benefits and costs of trading equate. Varian (1989) shows that trading can occur if investors have different priors of a risky assets mean. While this may explain why trading occurs, it offers little explanation as to what drives the priors. If differences in information drive trading, we would expect to see such investors compensated for the cost. The available evidence from individual investors in fact suggests the opposite: trading erodes returns. Heaton and Lucas (1996) propose that individuals trade in financial assets to buffer idiosyncratic income shocks in order to smooth consumption over time. Even if this provides another fully rational explanation for trading, it is difficult to see why this insurance should be valued at such high transaction costs. Investors could instead trade in mutual funds at a much lower cost.

The trading behavior of individual investors has often been attributed to overconfidence, as proposed by De Long, Shleifer, Summers, and Waldmann (1990), Kyle and Wang (1997), Daniel, Hirshleifer, and Subrahmanyam (1998), among others. In the psychology literature, overconfidence serves as a label—at least from a theoretical viewpoint—of two broad classes of cognitive biases.

The first, and most common, definition of overconfidence is the tendency for individuals to understate the uncertainty regarding their own estimates. When experimental subjects are asked to form confidence bounds around their point estimates, the outcome typically falls outside of the bound much more often than expected if people were well calibrated. This phenomenon is found to be task dependent, meaning that the evidence is strongest in tasks that subjects find difficult.[1]

The second manifestation of overconfidence is that people are unrealistically optimistic about their own ability. In a classic survey among students, Svensson (1981) finds that 82% rank themselves to be among the 30% of drivers with the highest driving safety. Such a belief can be linked to the concept of priors mentioned above, because it implies that individuals may overstate the significance of the information they may acquire. Furthermore, Langer and Roth (1975) find that individuals tend to ascribe success to their own ability and failure to bad luck. Such an *illusion of control* is therefore closely related to overconfidence.

In financial models, overconfident investors are those who hold unrealistic beliefs of how high their returns will be and how precisely these can be estimated. It is reasonable to believe that overconfidence may be more prevalent among individual investors, since money management is regarded as a difficult task for most people. In addition, feedback in terms of relative performance is very noisy, and therefore the ability to learn from behavior is low. In the previous literature, overconfidence has primarily been associated with excessive trading, but in principle, it could also lead to a lack of diversification. Investors overestimating the significance of the information they obtain regarding a particular stock may feel that investing in this stock is more attractive than investing in a more diversified portfolio.

In the previous literature on individual investor performance, Schlarbaum, Lewellen, and Lease (1978) match purchases to sales and find that a round-trip transaction costs

[1]For a review of these results, see McClelland and Bolger (1994).

around 3.5% in commissions. Investors in their sample more than compensate for this cost in their trading. These results have been contested by Barber and Odean (2000), who point out that if investors are more likely to realize gains than losses, this methodology is likely to produce overly favorable estimates of investor returns.[2]

Barber and Odean (2000), in contrast, measure returns from position statements, implicitly assuming that all trades are conducted at the end of the month, and estimate trading costs separately. They find that the average round-trip trade costs approximately 3% in commissions and 1% on the bid-ask spread for a round-trip transaction. An aggregated portfolio consisting of the top quintile of active traders loses as much as 6.5% annually compared to the market due to these costs.

A related result by Barber and Odean (2002) is that online investors significantly underperform a size-matched sample of investors who did not go online. They find that young men with high portfolio turnover are more likely to go online—and once they do—trade even more. They attribute this finding to three factors. First, it is argued that men are more overconfident than women, and will therefore be more likely to switch to online trading. Second, there is an illusion of control; in other words, investors who go online falsely perceive risk to be lower when they are able to monitor their portfolio instantly. Third, they propose that another psychological concept—*cognitive dissonance*—can reinforce trading activity. Cognitive dissonance occurs when individuals rationalize a behavior on the basis of prior beliefs. If the belief is that high performance is associated with intense trading activity and constant monitoring of the portfolio, it is precisely this behavior that such individuals will show.

Glaser and Weber (2003) conduct a survey among investors at an online broker, and are therefore able to test directly how different measures of overconfidence relate to trading volume. They find evidence that trading volume is related to the second manifestation of overconfidence, rather than the first: investors who believe they are above average trade more.

Overconfidence can explain trading behavior and lack of diversification, but not *which* stocks investors choose to buy. To gain a better understanding of investment strategies, we borrow a different concept from the psychology literature, namely the availability heuristic. Individuals have a clear tendency to underestimate risks when the context is familiar or available. Slovic, Fischhoff, and Lichtenstein (1982) find that individuals underestimate by far the risk of dying of common diseases, but overestimate the risk of rare and dramatic accidents. Even if accidents are rare, they attract much more attention when they occur. It is thus easy to attribute too much weight to such casual observations.

Odean (1999) reports that investors, on average, sell stocks that outperform those they

[2] This argument relies on the findings of Shefrin and Statman (1985), Weber and Camerer (1998), and Odean (1998) that investors hold on to the losers and sell the winners in their portfolios.

buy at a cost of around 3% per year. This cannot easily be explained by overconfidence, but is attributed to investors following naive investment strategies. Barber and Odean (2003) argue that people's buying behavior of stocks is subject to an availability bias, as stocks bought are more likely to have had an extreme return performance (positive *and* negative) or have had high media coverage. Benartzi and Thaler (2001) suggest another form of availability bias when choosing among mutual funds: the *1/n-heuristic*. They find that the number of funds available for selection determines the allocation, with investors naively splitting their investments in equal proportions across the funds. These examples provide evidence of systematic effects on buying behavior and should therefore be added to the previous evidence on the reluctance to sell losers.

It is still not clear in what ways naive strategies erode performance, unless they are negatively correlated with return patterns. Grinblatt and Keloharju (2000) find that the degree of sophistication matters for performance. They argue that households are likely to be less sophisticated investors and find that they trade to the opposite of investment professionals, such as institutions. In their sample, households act contrarian, and on average, lose from having such a strategy.

From the results cited here, it is tempting to generalize about private investors, who as Coval, Hirshleifer, and Shumway (2003) put it; "are often regarded as at best uninformed, at worst fools." However, they do find persistence in the performance of the top ten percent of most successful traders. This serves as an important reminder that not all private investors perform poorly, even if many do.

3 Data

An important advantage of studying an online broker is that the orders are placed directly by the investors. Although it is possible to place orders over the telephone as well, this constitutes a very small part of the transactions. There is therefore little direct interaction between the investor and the broker, which could otherwise be a source of concern when making inferences about performance across various investor groups. A drawback is that such investors cannot be regarded as being representative of any other group than online investors in general. As low fees are the main form of competition for online brokers when attracting customers, active traders are likely to be self-selected.

There are no tax exemptions for any accounts, as there are for the Keogh or 401(k) scheme in the U.S., where taxes are deferred. Furthermore, Swedish tax rules do not distinguish between the holding period of stocks, as is common in many countries. The tax rate is a flat 30% rate for the net of all realized gains and losses for private investors. It is therefore possible to aggregate all portfolio holdings across individuals, even if they in some cases possess more than one account.

The transaction file includes all trades in common Swedish stock for all customers at an Internet brokerage firm from the time it was established. The data stretches from mid-April 1999 until the end of March 2002, or 35 calendar months.[3] This file contains transaction prices, volumes and fees for each traded stock, as well as an individual identification tag that shows account number, age and gender for each trade. In addition, data are collected on closing prices for 521 distinct stock ticker names corresponding to the transaction file.[4]

From the original sample of 340,612 transactions distributed over 20,799 investors, I make the following exclusions: Accounts owned by minors, those under the age of 18 in the first year of trading, are excluded as it is unclear if they are independently managed. Portfolios worth less than or equal to SEK 1,000 in the first month are excluded, since apart from their being small, there is also very little trading in these portfolios. These small portfolios are not likely to be important for the investor, and the very fact that they are not traded may indicate that they are also judged by the investor as being too small. Finally, I exclude investors that trade but never owned a portfolio at month-end in the sample for selectivity reasons. An investor enters the sample by either buying or depositing stocks. When categorizing investors by trading activity, the first month is excluded if there were no deposits, because it may not be representative of how active the trader is.[5] By excluding these observations, we obtain a sample which is hereafter referred to as non-entering observations.

The fact that investors enter sequentially is displayed by Figure 2, along with the price level of a value-weighted Swedish stock index. There were 900 investors active in the sample at the end of 1999. By the end of 2000, the number of investors had grown to 11,261 and by 2001 they were 12,569. At most, which was in the last month of the sample, there were 13,917 investors active at the same time. Even if the pace at which investors entered is interesting in itself, it is not possible to know if they were new in the stock market or if they were experienced traders that switch between brokers.

There are 2,914 investors leaving before the sample period ends, but the attrition rate is relatively stable around the mean of 1.4% per month.[6] This is roughly four times as high as that found by Brown, Goetzmann, Ibbotson, and Ross (1992) in a sample of U.S. mutual funds. Odean (1999) analyzes active accounts at the beginning of his sample period and

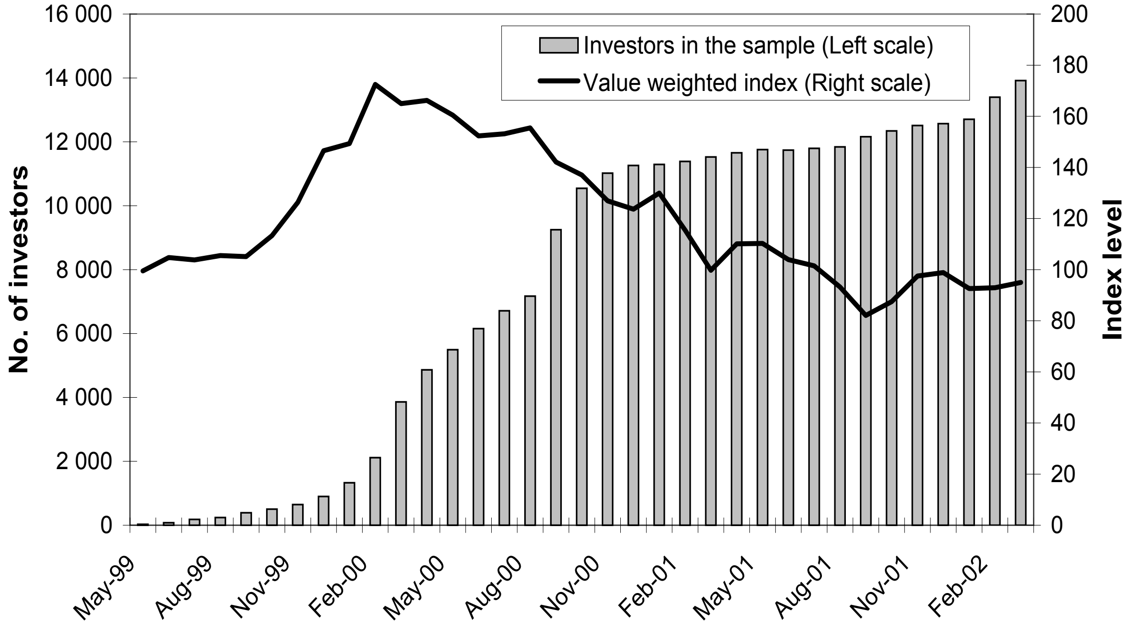
[3] In effect, I exclude all April 1999 transactions from the sample in order to get full calendar months of data. However, I calculate the portfolios held at the end of April 1999, and thus any positions from this period are included in the data.

[4] I include stocks from all official listings in Sweden. The prices were collected from *OM Stockholmsbörsen*, *Nordic Growth Market*, *Aktietorget* and *Nya Marknaden*.

[5] We do not wish to distinguish between an investor who begins her career by depositing stocks—and therefore records a zero turnover—and an investor buying the same portfolio, who will record a turnover of 50%.

[6] The stable attrition rate supports the hypothesis that most investors leave for exogenous reasons and there are only 41 instances in which investors go bankrupt, i.e., record a return of -100%.

Figure 2: Investors in the sample and the price level of stocks
The price path of the Swedish value-weighted index is plotted with a solid line (right scale) where the price is normalized to 100 on the last of April 1999. The bars denote the number of investors active in the sample each month (left scale).



find that 55% of the accounts fall out of the sample during the seven-year period in his study, which suggests a mean attrition rate of around 0.65% per month. When investors leave the sample, but are not replaced, there could potentially be a survivorship bias in favor of more successful investors that continue trading. The sample under consideration here contains all investors, and we could therefore measure the effect survivorship has on performance. However, it is clear that the data set across investors, for the most part, covers the bear market that followed the peak of the stock market boom in March 2000.

3.1 Sample summary

A description of the 324,736 transactions and 16,831 investors in the sample is presented in Table 1. There are 287,723 buy and sell transactions and 37,013 deposits and redemptions in all.

The average purchase is lower than the average sale, but as the number of purchases exceeds sales, the total value purchased is larger than the value sold. It is likely that part of this difference is attributable to new investors coming into the sample, thereby net investing in the market.

There is a great deal of skewness in the transactions, as evidenced by the mean being higher than the 75th percentile in all cases. This has implications for fees, as they are fixed within certain value brackets.[7] Therefore, small trades will be costly if measured

[7]The standard fee charge in Swedish kronor, is SEK 89 (approx. USD 10) for each transaction. For each

Table 1: Data Description: Transactions and Portfolios

Descriptive statistics of the transaction data are displayed in Panel A. The purchases and sales fees are averaged over the number of trades. Portfolio size in Panel B is determined by the first observation of total capital (as defined in the main text) for each investor. The mean turnover, number of observations, trades, stocks and technology weight are first averaged for each investor over the months they appear in the sample. USD 1 corresponds to about SEK 9 during the sample period.

Panel A: Transactions							
	No. of Obs.	Mean	Standard deviation	25th Percentile	Median	75th Percentile	Total value (MSEK)
Purchases, SEK	169,471	51,128	187,622	4,500	12,400	39,000	8,664.71
Purchases, fee in %		1.69	3.71	0.22	0.61	1.56	17.43
Sales, SEK	118,252	69,922	224,441	7,052	19,750	59,400	8,268.47
Sales, fee in %		1.94	12.48	0.16	0.48	1.12	13.34
Deposits	30,543	44,071	266,988	2,755	9,550	25,800	1,346.07
Redemptions	6,470	40,360	259,436	1,577	4,560	22,300	261.13
Panel B: Monthly portfolios							
	No. of Obs.	Mean	Standard deviation	25th Percentile	Median	75th Percentile	Total value (MSEK)
Portfolio obs.	265,342	15.77	8.44	8	18	23	n/a
Portfolio size	16,831	92,347	418,549	6,200	17,700	53,750	1,554.30
Turnover, SEK	16,831	74,392	846,646	616	2,532	11,281	16,963.95
Turnover, %	16,831	17.93	35.72	2.94	6,90	17.68	n/a
Number of trades	16,831	1.19	3.51	0.16	0.44	1.00	n/a
Number of stocks	16,831	3.30	2.95	1.38	2.36	4.00	n/a
Technology weight, %	16,831	66.52	35.09	38.28	78.41	100.00	n/a
Panel C: Investor demographics							
	No. of Obs.	Mean	Standard deviation	25th Percentile	Median	75th Percentile	Proportion, %
Age, All	16,831	38.95	12.35	29	37	48	100.00
Age, Men	13,768	38.43	12.20	29	36	47	81.80
Age, Women	3,063	41.24	12.74	31	39	51	18.20

as an average per transaction as in Table 1. This is illustrated by the fact that the mean purchase and sale fees are 1.69% and 1.94% measured on an average trade basis, whereas the value-weighted fees, obtained by dividing total trade value by the sum of fees, are as low as 0.20% and 0.16%, respectively. The median trade implies that a round-trip transaction costs around 1%. The sharp differences in value and trade weighted fees alone suggest that there may be considerable differences in performance depending on the size of the trades, which ultimately is related to portfolio size.

There are 16,831 investors in the sample from which 265,342 portfolios are recon-

200,000 interval of trade value above 20,000, there is an additional charge of SEK 119. However, for the most active investors, with more than 75 trades per quarter, the charge is only 5 basis points of the value of the trade, or a minimum of SEK 79 per transaction.

structed. To obtain a measure of portfolio size that is unrelated to investor returns, the first monthly observation of portfolio capital is used.[8] Portfolio size varies substantially between investors: the mean is SEK 92,347, and the median SEK 17,700.

The median for portfolio size in this sample is close to the figures from *Statistics Sweden* for the overall population. The median Swede owned Swedish stocks worth between SEK 20,000 and SEK 15,000 at the end of 1999 and 2001 respectively, but the corresponding average is much higher at SEK 319,000 and SEK 183,000.[9] The relative difference between the means and medians between time periods indicates that new investors enter the market. Between these dates, the share of the population that owned individual stocks rose from 16% to almost 22%. This is an unobserved variable in the sample, but it does suggest that a fair share of the investors studied here are new to the stock market.

The median investor in the sample holds an average of 2.33 stocks. The higher mean suggests that there are a minority of investors with a much higher degree of diversification across holdings. That the mean and median investor holds few stocks may not be so surprising given the relatively small value of the portfolio. In fact, roughly 18%, or 3,030 investors, hold only one stock. This feature of the data implies that the idiosyncratic component of individual portfolio returns is high.

The sample consists of 82% men, making it similar to the sample of online traders in Barber and Odean (2002). However, the median age is considerably lower. The median age of all investors is 37, with no significant difference in the age distribution between men and women. Therefore, the composition of investors broadly supports the hypothesis of Barber and Odean (2001a) that overconfidence is related to gender. If overconfident investors tend to self-select in becoming clients at online brokers, we may then expect to find them to be young and predominantly male.

Turnover is measured by dividing the total value of monthly trades by two times the value of the portfolio holdings each month. The average monthly turnover for each investor is almost 18%. The annualized turnover would therefore be 216%, implying that these investors flip their portfolios more than twice a year. By comparison, the Swedish stock market average turnover between 1999 and 2002 is around 62%.[10] This implies that turnover among the investors considered here is more than four times as high as the market in general. We also find considerable cross-sectional variation in trading, as the median investor only turns around 7.5% of the portfolio. Even if the median investor's turnover is much lower at an average yearly turnover rate of 90%, it is still well above the

[8]Portfolio capital includes the value of the portfolio at the beginning of the month as well as the value of any net purchases during the month. Portfolio capital is formally defined in Appendix A.

[9]The sharp difference between medians and means is even more extreme because the nationwide statistics include entrepreneurs who own very large stakes in their companies.

[10]This measure is constructed by dividing the value of all trades at the Stockholm Stock Exchange by two times the value of outstanding stock at year-end.

overall market mean.

Portfolio composition is analyzed with respect to industry classification.[11] The companies are categorized into nine industries: telecommunications, information technology, finance, health care, industrials, consumer goods, media, raw materials and services. Portfolio holdings are largely concentrated in two industry sectors: telecommunications and information technology, combined and hereafter referred to as technology. The median investor holds an average of 71% technology stocks, which represents a clear overweight of the sector. The technology sector has a predominant role as it represented between 34% and 48% of the value-weighted Swedish market index. On a relative basis, the mean investor in the sample allocates twice the weight to technology compared to the technology index weight.[12]

The strong tilt towards these stocks among investors can have several explanations. First, technology stocks are riskier, and investors may prefer to take higher risk. But rational investors diversify their portfolios to avoid idiosyncratic risk; they typically do not choose only one stock. Second, it is reasonable to assume that companies in the technology industry on average are smaller, and investors prefer small stocks. But the evidence for this is not very clear. The median company for the consumer goods, media and services industries are equally small or even smaller. Third, during this period, the technology sector offered a wider set of companies that investors could choose from. There is slightly more support for this, since only one other sector contains close to an equal number of stocks—industrials. This feature is relevant if investors follow the 1/n-heuristic as suggested by Benartzi and Thaler (2001). Fourth, it is important to keep in mind that during the Internet frenzy, new companies entered the stock market at an unprecedented pace. It is possible, or even likely, that the news flow was biased towards the technology sector. Barber and Odean (2003) also propose that naive investors select stocks that have experienced extreme price movements. This could also explain why risky technology stocks are overrepresented in the sample.

Naive investors may therefore react to signals that are unrelated to information for several reasons. But the rational principle of diversification could be contrasted with naive strategies. Sophisticated investors, who are less overconfident and prone to react to noise, are more likely to be better diversified than those following naive strategies. A preliminary investigation of such systematic effects of investor behavior can be studied in the correlations reported in Table 2.

Quite naturally, the number of stocks held and the value of the portfolio are highly

[11]The industry classification is made by *Affärsvärlden*, who also produce the value-weighted index used here.

[12]This measure is obtained by dividing the investors weight by the overall technology sector index weight, each month. It is not reported separately, because the results are similar to that of the absolute weight, which in turn are easier to interpret.

Table 2: Correlations: Individual characteristics

The table reports non-parametric Spearman correlations for 16,831 individual investor characteristics given in Table 1, excluding non-entering observations. All values are significantly different from zero at the 1% level, except the correlation between age and turnover which has a *p*-value of 9%.

	Portfolio size	Turnover	Number of trades	Number of stocks	Technology weight	Age
Portfolio size	1.00	-	-	-	-	-
Turnover	0.09	1.00	-	-	-	-
Number of trades	0.17	0.82	1.00	-	-	-
Number of stocks	0.24	0.10	0.31	1.00	-	-
Technology weight	-0.04	0.03	-0.02	-0.22	1.00	-
Age	0.08	0.01	0.04	0.15	-0.11	1.00

positively correlated, as are turnover and the number of trades. The fact that age and portfolio value is positively related can be an indication that portfolio value in turn is correlated with (unobservable) overall wealth. Age and technology weight are negatively related, suggesting that stocks within this industry are more popular among younger investors.

Two correlations are more interesting than others. The first is portfolio value, which is positively correlated with both the number of trades and turnover. This is in contradiction to the common apprehension that trading is most frequent among small investors. The second finding is a substantial negative correlation between the number of stocks in the portfolio and the technology weight. There is, of course, a binary choice of industry when few stocks are held, such that diversification must by necessity be related to stock holdings. What is not so obvious, as the negative correlation suggests, is that investors *on average* choose a lower technology sector exposure when holding more stocks. This indicates that investors pursue different strategies depending on portfolio composition.

The positive correlation between portfolio turnover and size, and the negative correlation between technology weight and diversification, suggest a general pattern. To analyze these two features of the data in more depth, the investors are sorted into quintiles formed on the basis of these variables.

3.2 Turnover and portfolio size

Given the major differences in median and mean fees, the performance of small investors is likely to suffer due to their small-sized trades. Hence, it may be important to control for portfolio size when looking at performance. I apply a two-pass sorting procedure. In the first pass, the investors are sorted by turnover into five groups that contain approx-

Table 3: Quintiles sorted by turnover and portfolio size

Investors are first sorted into quintiles based on their average turnover, excluding entering observations. A second sorting is conducted on portfolio value, thereby partitioning each turnover quintile into five sub-quintiles based on portfolio size. Panel A and B report the means of turnover in percent per month and portfolio size in SEK. USD 1 corresponds to about SEK 9 during the sample period.

	Turnover, quintiles					
	(Low) 1	2	3	4	(High) 5	All
Panel A: Mean turnover, monthly %						
Portfolio size 1 (Small)	0.00	1.24	4.27	10.26	41.56	11.47
Portfolio size 2	<0.01	1.37	4.12	10.43	48.05	12.79
Portfolio size 3	<0.01	1.36	4.18	10.37	53.22	13.83
Portfolio size 4	0.00	1.30	4.20	10.43	62.89	15.76
Portfolio size 5 (Large)	<0.01	1.20	4.19	10.81	87.43	20.74
All	<0.01	1.29	4.19	10.46	58.63	14.92
Panel B: Mean portfolio size, SEK						
Portfolio size 1 (Small)	2,882	3,592	3,444	3,555	4,223	3,539
Portfolio size 2	5,142	8,346	8,268	9,449	12,432	8,727
Portfolio size 3	9,801	18,304	18,284	20,799	30,341	19,506
Portfolio size 4	22,049	44,489	41,540	45,988	76,944	46,201
Portfolio size 5 (Large)	174,725	542,453	341,642	311,246	548,729	383,808
All	42,907	123,401	82,612	78,185	134,618	92,347

imately 3,366 investors each. In the second pass, each turnover quintile is sorted into five equally sized sub-quintiles.[13] There are then about 673 investors in 25 groups in the turnover/ portfolio size dimension. Table 3 displays the means of turnover and size for each group. The main difference in trading activity between investors is that those in the lowest turnover quintile hardly ever trade, while those in the highest quintile trade extensively. Those who trade the most have a turnover of almost 59% per month, which on a yearly basis means that they buy and sell their portfolio almost seven times. In fact, the investors in the top turnover quintile account for more than 60% of the trades.

Among the investors in turnover quintile 5, those with the largest sized portfolios trade significantly more than all other groups. Among these investors, turnover is almost 90% per month. This, in turn, drives the overall result that the quintile with the largest portfolio size has the highest turnover rate. Comparing the overall means of turnover in Table 1 and Table 3, it falls from 18% to 15% when only the non-entering observations are included.

Portfolio size is also unevenly distributed across individuals. The mean size of the smallest quintile is around 100 times smaller than the largest quintile. The investors with

[13]This sorting procedure is therefore similar to that used by, e.g., Fama and French (1992) when exploring the book-to-market and size effect.

Table 4: Quintiles sorted by technology weight and number of stocks
Investors are first sorted into quintiles based on the average number of stocks in their portfolio. A second sorting is done on the investors' average technology weight, thereby partitioning each diversification quintile into five sub-quintiles based on the average technology weight. Panel A and B report the means of the number of stocks held and the technology weight in percent.

	Diversification quintiles: Number of stocks held					
	(Few) 1	2	3	4	(Many) 5	All
Panel A: Mean number of stocks						
Technology weight 1 (Low)	1.00	1.65	2.43	3.71	8.65	3.49
Technology weight 2	1.02	1.72	2.48	3.71	8.35	3.45
Technology weight 3	1.00	1.54	2.48	3.70	7.78	3.30
Technology weight 4	1.00	1.74	2.50	3.71	7.43	3.28
Technology weight 5 (High)	1.00	1.60	2.19	3.49	6.65	2.99
All	1.01	1.65	2.42	3.66	7.77	3.30
Panel B: Mean technology weight, %						
Technology weight 1 (Low)	0.06	7.44	13.08	14.97	13.63	9.83
Technology weight 2	91.15	53.81	49.14	45.70	36.08	55.17
Technology weight 3	100.00	89.96	79.51	68.43	54.06	78.39
Technology weight 4	100.00	100.00	97.02	87.38	71.83	91.24
Technology weight 5 (High)	100.00	100.00	100.00	99.30	90.93	98.04
All	78.22	70.22	67.73	63.14	53.30	66.52

the lowest trading activity clearly have smaller portfolios, around half the value of the overall sample. The data also suggest that those who trade the most have larger portfolios compared with the other size matched turnover quintiles.

3.3 Diversification and technology weight

An identical approach is used to investigate how stock diversification is related to investors technology weight. Investors are first sorted by the number of stocks held and then by the technology weight. Table 4 reveals that the mean number of stocks held among the 20% of investors that are least diversified is close to one. In the top quintile, they hold around eight stocks. Those who have a lower technology weight also have slightly more stocks than those who have the highest weight—3.49% compared with 2.99%.

The most striking result is found in Panel B, where the mean technology weight decreases monotonically with the number of stocks held from 78% to 53%. This means that four out of five investors who hold only one stock choose one in the technology sector. Consequently, there is strong evidence of a systematic effect of diversification that goes beyond that of simply holding stocks of different companies. Investors with more stocks in their portfolio choose to be less exposed to the technology sector. This suggests that investment strategies differ between groups of investors, and implies that there are underlying differences in behavior that could be related to investment skill.

4 Results

The previous data analysis reveals considerable cross-sectional variation in trading and diversification that may be helpful in explaining performance. At the outset, we may expect that excessive trading erodes performance, but the prior of how diversification should affect performance is not very clear. A random strategy—where investors hold a few stocks selected at random—should be related to idiosyncratic risk only, and be unrelated to mean returns. But this is true only if the strategy and associated expected returns are independent. Overconfident investors who follow naive investment strategies will underestimate risk, and their forecasts for expected returns will be overly favorable. Such investors are not only likely to take more systematic risk, but may also be less skilled in choosing which stocks to select. In this case, performance may vary with diversification. Overconfident investors who hold undiversified portfolios could be less skilled in choosing which stocks to select.

To facilitate such comparisons, the results are presented in two parts. The first part begins by reviewing how portfolio returns are constructed from transaction data and proposes a return decomposition. The market adjusted return can be split into components that are designed to identify the returns that can be associated with investor turnover and industry selection. The return differences are evaluated separately over the quintiles in these dimensions, as in the previous section. The second part presents a panel regression model. The individual characteristics are incorporated at the same time, such that we obtain marginal effects of those found in the first part.

4.1 Investor returns

The data are available in transaction form, from which portfolios are reconstructed. The key issue, when defining returns, is to identify the payoff and the corresponding capital that can be associated with it.[14] Only a brief summary of the method is presented here, without going into any details of the definitions. A more exhaustive explanation of how portfolio excess returns are calculated from transaction data is given in Appendix A. Three returns are used in the analysis: excess passive return, $R^P_{i,t}$; total excess return, $R_{i,t}$; and industry excess return, $R^{Ind}_{i,t}$. Each of them is explained below.

Passive excess return refers to the return of the portfolio held by investor i at date $t-1$, i.e., the beginning of the month. The payoff is calculated for each stock as the price change

[14] A related approach has been applied by Linnainmaa (2003), who investigates daytrades. However, the method considered here defines payoffs and required capital quite differently.

during month t times the number of stocks held at date $t-1$. The payoffs are then summed over all stocks in the portfolio, and normalized into a return by the value of the total position. This value, which is the required capital to finance the portfolio, is referred to as *position capital*. We denote the passive excess return adjusted with the 30-day T-bill rate $R^P_{i,t}$.

When investors trade, the payoff is calculated as follows for each stock. Suppose a transaction in a certain stock for an individual takes place at date d, which is at some point during month t. If the trade is a purchase and the stock is held throughout month t, the net proceeds are calculated from date d to t, and conversely, if it is a sale of a stock owned at date $t-1$, from $t-1$ to d. As is shown in Appendix A, intra-month transactions for each stock can be aggregated and averaged, so the net effect applies to what has already been stated. The sum of the payoffs over each stock in the portfolio is the value change of the portfolio during month t.

The key now is to identify the capital components associated with trading. The minimum capital requirement for each investor is assumed to be the position capital measured at date $t-1$. If purchases exceed sales, in cases requiring additional funds, these funds are added to the capital base and labelled *trading capital*.[15] Total capital is thus the sum of position capital and trading capital.

The portfolio return is in excess of the interest rate, which is adjusted for in the following way. Reducing the total payoff with the cost of position capital is straightforward, because this is the minimum cost for financing this portfolio. However, when there is trading, investors can be net sellers or net buyers. Interest is added to the payoff if they are net sellers, and deducted if they are net buyers. The interest associated with trading involves calculating the cash balance for each investor at each point in time during the month. This "fictitious cash account" therefore assures that net buyers or net sellers are charged or compensated for cash-flows at the going 30-day interest rate.[16] The resulting *total excess return* is denoted $R_{i,t}$, and includes all trades between $t-1$ and t. Therefore—if there is no trading—$R^P_{i,t}$ coincides with $R_{i,t}$.

The return measure does not include dividends. This exclusion will bias the returns measured here downwards. However, this bias is expected to be small. The overall market paid little in dividends during the period, and especially the growth firms held by the investors in the sample. For this reason, the market return used as benchmark does not include dividends.

The third and last return needed for the analysis is the *industry excess return*, $R^{Ind}_{i,t}$,

[15] Note that the timing of sales and purchases matters for the definition of the capital base. Consider a sale and a purchase of the same value. If the purchase precede the sale, capital is required whereas there is no effect if it were the other way around.

[16] The effect on returns stemming from the interest rate, however is small due to the high volatility of stock returns.

constructed as follows. The industry weights for the portfolio the investor holds at date $t-1$ are calculated. The industry return is the weighted average of the excess returns on the nine industry indexes, and therefore tracks the index composition of each investor's portfolio.

4.1.1 A simple return decomposition

The decomposition aims to clarify the return difference between a passive strategy (excluding trading) and an active strategy (including trading) as well as how a passive strategy relates to various benchmarks. By using the three returns, we can offer the following definitions. *Market-adjusted return* is defined as

$$\Delta IM_{i,t} \equiv R_{i,t} - R_{M,t},$$

where $R_{M,t}$ is the excess return of the value-weighted market benchmark. *Trade-adjusted return* is defined as

$$\Delta IP_{i,t} \equiv R_{i,t} - R^{P}_{i,t},$$

and serves as an approximation of the contribution of active trading. Passive return can be thought of in this setting as an own-benchmark return in the same spirit as proposed by Grinblatt and Titman (1993), who investigate the performance of mutual funds. They argue that any asset pricing model is sensitive to its particular assumptions, but the own-benchmark can serve as an intuitive and appealing means of comparison. Grinblatt and Titman use yearly and quarterly fixed portfolios when defining the benchmark portfolio. Here, passive return is defined on a monthly basis. Investors in this sample have a much higher turnover than mutual fund managers, and there is enough variation in a month to enable interesting comparisons. Passive return serves as a natural benchmark when investigating if rebalancing is profitable for investors. It should be noted that when there is no portfolio observation at the beginning of the month, we cannot observe a passive return. To make investor returns comparable with or without trading, only non-entering observations will be considered. More importantly, measured trading costs can be difficult to interpret if the first purchased portfolio is included. A buy-and-hold investor needs to buy the portfolio at some stage, but transaction costs are averaged over a very long time.

Market-adjusted industry return is written

$$\Delta INDM_{i,t} \equiv R^{Ind}_{i,t} - R_{M,t},$$

which is a measure of the return contribution stemming from the choice of industry compared to the market benchmark. It follows by construction that if an investor holds the market value weights, the difference is zero.

Industry-adjusted passive return is defined as

$$\Delta PIND_{i,t} \equiv R^P_{i,t} - R^{Ind}_{i,t},$$

and measures the difference between the actual portfolio held at the beginning of the month and a portfolio that tracks the return of the chosen industries. This can be interpreted as a measure of how well investors can select stocks *within* industries. Consider an investor who owns two stocks, but in different industries. Even if the industries underperform the market, the selected stocks might still outperform the chosen industries. This is exactly what is captured by the industry-adjusted passive return.

By using the definitions above, we can express the market-adjusted return as the sum of three components: the trade-adjusted return, the market-adjusted industry return, and the industry-adjusted passive return

$$\Delta IM_{i,t} = \Delta IP_{i,t} + \Delta INDM_{i,t} + \Delta PIND_{i,t}.$$

Furthermore, it follows that we are also able to define the *market-adjusted passive return* as

$$\Delta PM_{i,t} \equiv \Delta IM_{i,t} - \Delta IP_{i,t} = \Delta INDM_{i,t} + \Delta PIND_{i,t} = R^P_{i,t} - R_{M,t},$$

which then completes the link between the five definitions and three returns. Table 5 displays the results of this decomposition in four ways: returns with or without fees, and by weighing returns equally or with total capital.

The average investor in the sample had a monthly return of -3.38%, which implies an annualized excess return of a substantial -34%. The strong negative return indicates that investors, on average, have experienced very high losses. This can partly be explained by the fact that investors enter the market sequentially, as illustrated by Figure 2. The large number of investors who entered the sample late inevitably faced a weaker stock market.

The market-adjusted return makes a crude adjustment for such effects. Still, investors lose between 1.8% to 2.1% per month compared to the market, including fees. The equally-weighted means of the trade-adjusted return reveal that 32 basis points can be explained by fees alone. In annualized terms, 32 basis points per month means that the average investor paid around 3.8% per year of her portfolio value in fees. This is more than twice the annual fee charged by most mutual funds. Further, the effect of value-weighting investors on fees is clear, implying that large investors pay less in fees expressed as a percentage of the portfolio return.

There is a small but still negative difference in the trade-adjusted return even when fees are excluded. This is evidence that investors on average do not beat their own-benchmark defined by the portfolio held at the beginning of the month. The trade-adjusted mean when investors are value-weighted actually implies that large-sized in-

Table 5: Investor mean returns: A simple decomposition

The market-adjusted return is decomposed into three parts. The trade-adjusted return measures the effect of rebalancing. The difference between the market-adjusted return and the trade-adjusted return is labelled the market-adjusted passive return, which in turn has two components: The market-adjusted industry return measures the contribution stemming from industry selection with respect to the market, and the industry-adjusted passive return measures stock-picking ability within industries. There are 251,879 non-entering observations in the sample from which averages of 16,831 investor mean returns are constructed. The means for investor portfolio returns are weighted equally or with total capital as defined in the text. The effect of including or excluding fees is presented separately.

Returns, monthly %			With fees		Without fees	
Operation	Definition	Comment	Value-weighted	Equally-weighted	Value-weighted	Equally-weighted
None	Portfolio excess return, $R_{i,t}$	Investor total excess return.	-3.05 (0.05)	-3.38 (0.05)	-2.97 (0.05)	-3.07 (0.05)
None	Market-adjusted return, $\Delta IM = R_{i,t} - R_{M,t}$	Total return including monthly rebalancing in excess of market.	-1.78 (0.05)	-2.08 (0.05)	-1.70 (0.05)	-1.76 (0.05)
-	Trade-adjusted return, $\Delta IP = R_{i,t} - R^{P}_{i,t}$	Trading contribution from rebalancing.	-0.21 (0.02)	-0.37 (0.02)	-0.13 (0.02)	-0.05 (0.02)
=	Market-adjusted passive return, $\Delta PM = R^{P}_{i,t} - R_{M,t}$	Buy-and-hold return in excess of market.	→	→	-1.57 (0.05)	-1.71 (0.05)
-	Market-adjusted industry return, $\Delta INDM = R^{Ind}_{i,t} - R_{M,t}$	Industry contribution to buy-and-hold return.	→	→	-1.20 (0.02)	-1.28 (0.02)
=	Industry-adjusted passive return, $\Delta PIND = R^{P}_{i,t} - R^{Ind}_{i,t}$	Stock selection contribution to buy-and-hold return.	→	→	-0.37 (0.04)	-0.43 (0.04)

Mean standard errors in parentheses. All variables are significantly different from zero at the 1% level or lower, except equally-weighted ΔIP without fees, which is significant at the 5% level.

vestors lose more than median investors when fees are excluded. The fee itself only explains some 8 basis points of the total 21 points.

The difference between the market-adjusted return and the trade-adjusted return can be further analyzed and decomposed into two parts. Both of these are defined for passive portfolios, such that they are free from trading. Hence there is no need for a separate analysis with respect to fees.

The market-adjusted industry return shows the difference between the market return and the particular choice of industries. Most of what can explain the deviations from the market return is embedded here. Investors have chosen to invest in industries that have underperformed relative to the market, which is most likely a direct consequence of the strong tilt towards technology stocks. As this simple decomposition does not include any risk-adjustment, this effect might very well be a result of investors choosing higher systematic risk.

Risk is likely to be less problematic when evaluating the industry-adjusted passive return, as there is considerably less variation in risk within industries than between. The industry-adjusted passive return reveals that the investors on average lose around 43 basis points from choosing stocks that underperform any chosen industry. This is interesting, as it suggests that individuals may systematically choose stocks that underperform. Furthermore, the somewhat higher value-weighted return indicates that this pattern is more predominant among investors with small portfolios.

If we assume that all systematic risk is captured by industries, the decomposition suggests that investors underperform the market by around 80 basis points. In this case, trading and stock selection are roughly equally important. To examine these two features of the data in more depth, the following sections report the returns associated with the corresponding quintiles of Table 3 and Table 4.

4.1.2 Returns: Trading and portfolio size

We will now look more closely at how the trade-adjusted return is related across investor groups. The mean return is calculated for each of the 25 groups of investors defined in Section 3.2. This is also done for all investors in each quintile in the two dimensions, and finally for the whole sample. Table 6 reports the mean returns corresponding to the sample partition in Table 3. Associated standard errors are given in parentheses.

As there is virtually no trading in the lowest turnover quintile, there can be no deviation from the own-benchmark, and so the passive return equals the total return. This means that the difference in the first column of Table 6 is zero. It is clear that when trading activity increases, performance declines. The most active traders underperform their benchmark portfolio by 95 basis points compared with only 52 for traders in the fourth quintile. This general effect seems to be valid for all portfolio sizes, but the smallest investors contribute most to this general pattern.

Therefore, we find a size effect as well as a trading effect: the mean underperformance for the investors with the smallest portfolio size is 79 basis points, but only 15 for the largest. When fees are excluded from the analysis, there is still a weak size and turnover effect, but only 5 basis points are lost on average when fees are excluded. One reason for this is that investors, on average, pursue strategies that are unprofitable. Barber and Odean (2000) attribute a similar, but daily, effect to costs associated with the bid-ask spread.

It is somewhat puzzling that the largest investors who trade the most lose up to 42 basis points, excluding fees. When comparing Panel A and B, we see that fees only explain 17 basis points of the total trade-adjusted return. On the other hand, this group was also found to be trading more than twice as much as the smallest investors in Table 3. A net cost of 42 basis points may not be so conspicuous considering that almost 90% of the portfolio is traded in one month. Yet, somewhat surprisingly, the investors with the smallest portfolios that trade the most gain 31 basis points by trading, excluding fees. However, the performance in this group is so dispersed that it is insignificant.

Table 6: Mean trade-adjusted returns

The trade-adjusted return measures the difference between the total portfolio return and the passive return, which is the portfolio held at the beginning of the month. The return is reported with fees in Panel A, and without fees in Panel B. There are approximately 673 investors in each sub-quintile corresponding to the partition in the turnover and portfolio size dimensions in Table 3.

	Turnover quintiles					
	(Low) 1	2	3	4	(High) 5	All
Panel A: ΔIP, Trade-adjusted returns including fees, monthly %						
Size 1 (Small)	0.00	-0.20***	-0.53***	-1.22***	-2.01***	-0.79***
	(0.00)	(0.02)	(0.02)	(0.13)	(0.35)	(0.08)
Size 2	<0.01	-0.13***	-0.37***	-0.59***	-0.87***	-0.39***
	(<0.01)	(0.03)	(0.03)	(0.07)	(0.26)	(0.06)
Size 3	<0.01	-0.09***	-0.20***	-0.40***	-0.78***	-0.29***
	(<0.01)	(0.02)	(0.03)	(0.09)	(0.28)	(0.06)
Size 4	0.00	-0.06***	-0.15***	-0.28***	-0.51***	-0.20***
	(0.00)	(0.01)	(0.03)	(0.06)	(0.17)	(0.04)
Size 5 (Large)	<0.01	-0.01	-0.06***	-0.11***	-0.59***	-0.15***
	(<0.01)	(0.01)	(0.03)	(0.05)	(0.12)	(0.03)
All	<0.01	-0.10***	-0.26***	-0.52***	-0.95***	-0.37***
	(<0.01)	(<0.01)	(0.01)	(0.04)	(0.11)	(0.02)
Panel B: ΔIP, Trade-adjusted returns excluding fees, monthly %						
Size 1 (Small)	0.00	-0.05**	-0.07*	-0.21*	0.31	-0.01
	(0.00)	(0.02)	(0.04)	(0.12)	(0.33)	(0.07)
Size 2	<0.01	-0.04	-0.13***	-0.12*	0.11	-0.04
	(<0.01)	(0.03)	(0.03)	(0.07)	(0.26)	(0.05)
Size 3	<0.01	-0.03	-0.06**	-0.08	-0.17	-0.07
	(<0.01)	(0.02)	(0.03)	(0.09)	(0.28)	(0.06)
Size 4	0.00	-0.02*	-0.06**	-0.11*	-0.17	-0.07*
	(0.00)	(0.01)	(0.03)	(0.06)	(0.17)	(0.04)
Size 5 (Large)	<0.01	0.01	-0.01	-0.01	-0.42***	-0.09***
	(<0.01)	(0.01)	(0.02)	(0.05)	(0.12)	(0.03)
All	<0.01	-0.02**	-0.06***	-0.10***	-0.07	-0.05**
	(<0.01)	(<0.01)	(0.01)	(0.04)	(0.11)	(0.02)

Mean standard errors in parentheses. Significance levels for a *t*-test of the mean to be different from zero at the 10%, 5%, and 1% level are marked (*), (**), and (***).

4.1.3 Returns: Diversification and technology weights

The natural candidates to analyze the effect of diversification across stock holdings are the market-adjusted industry return and the industry-adjusted passive return. As an extension to Table 4, these returns are investigated across diversification and the technology weight, which here serves as a crude measure of risk.

The market-adjusted industry return measures how the choice of industry has affected investors portfolio return relative to the market. The column on the far right of Panel A in Table 7 reveals that the group of investors who underweighted the technology sector outperformed the value-weighted index. But since over 75% of the investors in this sample did the opposite, the means become negative moving down the column. There is a similar effect across the quintiles sorted by the degree of diversification. The market-adjusted industry return is more negative for investors with few stocks, which is most likely due to the technology weight that was found to be higher among these investors. Therefore, these results simply confirm that investors chose to carry a lot of risk, but faced unfavorable market conditions.

Panel B provides more interesting results in this respect. The industry-adjusted passive return controls for the industry choice for each investor at each point in time. Any relative deviation from this benchmark stems from the investor's choice of individual stocks *within* each industry. The risk among firms within industries is likely to be more similar. The overall result, which shows that 43 basis points are lost due to stock selection within the industries, is substantial.

There is little systematic variation across technology quintiles (moving vertically down the rightmost column of Panel B in Table 7). If any, those with lower technology weights appear to underperform their industry benchmark more than those with higher weights. Therefore, there is no evidence that the overall negative return stemming from which stocks to buy in a given industry is related to a preference for technology stocks.

There is a much clearer pattern found horizontally in Panel B of Table 7. Investors with few stocks underperform more relative to those with many stocks in their portfolios. One must bear in mind that the industry-adjusted passive return measures the relative performance of individual stocks and any mix of industries. A random strategy will "average out" investor returns over diversification quintiles if choices were independent. Choosing several stocks within a given industry should reduce the *variance* of such a portfolio, but not change the *mean*. Here, we find that virtually all investor groups with few stocks have inferior mean returns than those who have many. The ability to target individual stocks that perform better increases with the number of stocks held.

The diversification measure could be sensitive to investors going bankrupt. It is more likely that those investors who left the sample due to bankruptcy are to be found in the

Table 7: Mean industry-adjusted returns

The market-adjusted industry return in Panel A measures the return difference between the market and the chosen industry portfolio for each investor. The industry-adjusted passive return in Panel B measures the difference between the chosen industry portfolio and the actual chosen stocks of the portfolio held at the beginning of the month. There are approximately 673 investors in each sub-quintile corresponding to the partition in the technology weight and diversification dimensions in Table 4.

	Diversification quintiles: Number of stocks held					
	(Few) 1	2	3	4	(Many) 5	All
Panel A: ΔINDM, Market-adjusted industry returns, monthly %						
Technology weight 1 (Low)	2.05***	1.65***	1.33***	1.11***	0.93***	1.41***
	(0.11)	(0.08)	(0.07)	(0.07)	(0.06)	(0.04)
Technology weight 2	-2.38***	-0.60***	-0.49***	-0.27***	0.01	-0.75***
	(0.10)	(0.08)	(0.06)	(0.05)	(0.04)	(0.04)
Technology weight 3	-2.89***	-2.43***	-1.94***	-1.48***	-0.80***	-1.91***
	(0.12)	(0.06)	(0.07)	(0.07)	(0.05)	(0.04)
Technology weight 4	-2.88***	-2.56***	-2.81***	-2.41***	-1.60***	-2.45***
	(0.10)	(0.12)	(0.07)	(0.05)	(0.05)	(0.08)
Technology weight 5 (High)	-2.83***	-2.97***	-2.72***	-2.77***	-2.13***	-2.68***
	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.04)
All	-1.78***	-1.38***	-1.32***	-1.17***	-0.72***	-1.28***
	(0.06)	(0.05)	(0.04)	(0.04)	(0.03)	(0.02)
Panel B: ΔPIND, Industry-adjusted passive returns, monthly %						
Technology weight 1 (Low)	-1.27***	-0.16	-0.62***	-0.45***	-0.17**	-0.53***
	(0.41)	(0.28)	(0.17)	(0.14)	(0.09)	(0.11)
Technology weight 2	-0.40	-0.49**	-0.89***	-0.29*	-0.33***	-0.48***
	(0.27)	(0.24)	(0.20)	(0.15)	(0.08)	(0.09)
Technology weight 3	-0.90***	-0.11	-0.40***	-0.27*	-0.36***	-0.41***
	(0.31)	(0.18)	(0.15)	(0.15)	(0.12)	(0.09)
Technology weight 4	-0.22	-0.60***	-0.23*	-0.11	-0.38***	-0.31***
	(0.35)	(0.21)	(0.14)	(0.11)	(0.10)	(0.09)
Technology weight 5 (High)	-0.63**	-0.03	-0.61***	-0.57***	-0.26*	-0.42***
	(0.29)	(0.24)	(0.23)	(0.20)	(0.16)	(0.10)
All	-0.68***	-0.28***	-0.55***	-0.34***	-0.30***	-0.43***
	(0.15)	(0.10)	(0.08)	(0.07)	(0.05)	(0.04)
All, excl. bankrupt investors	-0.55***	-0.22***	-0.54***	-0.34***	-0.30***	-0.43***
	(0.14)	(0.10)	(0.08)	(0.07)	(0.05)	(0.04)

Mean standard errors in parentheses. Significance levels for a *t*-test of the mean to be different from zero at the 10%, 5%, and 1% level are marked (*), (**), and (***).

group holding only one stock. The bottom row in Panel B of Table 7 reports the means when these investors are excluded from the sample. Even if the performance rise for one-stock investors, they still underperform by almost twice the amount compared to those best diversified.

The systematic effect of diversification on performance suggests that this variable could be related to experience or skill, but the relatively high average underperformance could also be an indication that investors choose stocks that are riskier than their respective industry benchmarks. If this is the case, such risks should also be correlated across investors in the diversification dimension.

4.1.4 Summary of results from the return decomposition

In all, three results are obtained from the return decomposition. First, investors that trade more, lose more. This is found to be almost entirely related to fees, which in turn are related to the size of the portfolio. Large portfolios are less affected by fees due to the fee structure that involves minimum costs. Second, the high gear towards technology stocks in combination with bad market timing means that most of what is lost above the market-adjusted return is related to industry choice. Third and last, the number of stocks held is found to be related to how investors perform when adjusting for industry choice.

These preliminary findings are interesting from a descriptive viewpoint and to understand the data. On the other hand, to be able to make any firm statements about performance, there is a need to make riskadjustments and to control for interdependence among the measured effects.

4.2 Panel estimation

The natural starting point when building a model for portfolio evaluation is the Capital Asset Pricing Model (CAPM). When using the market return as a benchmark to assess the risk of a portfolio, it ignores common variation caused by time-varying expectations. Traditional, unconditional models can ascribe abnormal performance to an investment strategy that only relies on public information as shown, for instance, by Breen, Glosten, and Jagannathan (1989).

Further, given the size of this sample, modelling a separate beta for each investor is not a realistic option. On the other hand, it would be desirable to allow for heterogeneous preferences and investment strategies. The goal, therefore, is to allow beta to vary between investors in some predetermined and structured manner, while allowing for time-variation.

The asset pricing model suggested here is an extension of the conditional CAPM proposed by Ferson and Schadt (1996). Let us assume that investor returns can be described by

$$R_{i,t} = B_0 R_{M,t} + \sum_{k=1}^{K} B_k \left[y_{k,i} R_{M,t} \right] + \sum_{l=1}^{L} B_{l+K} \left[z_{l,t-1} R_{M,t} \right] + \varepsilon_{i,t}, \qquad \begin{array}{l} i = 1, \ldots, N, \\ t = 1, \ldots, T. \end{array} \tag{1}$$

There are i investors grouped into K investor risk characteristics. The investor characteristics are specific to each individual and hence fixed over time. In addition, there are L information or state variables $z_{l,t-1}$ which describe the investors' opportunity set and is the same for all individuals. The state variables represent information that is common and known to the investors in t. Lower case letters for the characteristics and information variables are deviations from unconditional means, $y_{k,i} = Y_{k,i} - \overline{Y}_{k,\cdot}$ and $z_{l,t-1} = Z_{l,t-1} - \overline{Z}_{l,\cdot}$. The excess return of the market benchmark is denoted by $R_{M,t}$ and $\varepsilon_{i,t}$ is an investor and time-specific disturbance term. The coefficient B_0 can be thought of as the average beta with $B_1, \ldots, B_{K+J}$ as linear response coefficients to investor characteristics and state variables.[17]

In this way, we obtain rich variation in the cross-section, but the individual characteristics are kept fixed over time as to keep the interpretation of the results clear. Similarly, the proxy for the information set across investors is kept constant, but varies over time.

A typical implementation of the model specified by equation (1) is to add intercept terms for each investor and then test the null hypothesis that they are jointly or individually zero. However, the interest here is to relate performance to investor characteristics. We already have reasons to believe that the strong prediction of market efficiency may not be applicable to online investors. Online investors are not well diversified. In addition, they face higher transaction and search costs than, for instance, mutual funds. They are also more likely to be subject to behavioral biases such as overconfidence, and follow naive strategies that may affect their performance.

In essence, it is of interest to model the intercept by controlling for the investor characteristics in various ways. Therefore, the regressions performed is of the form

$$R_{i,t} = A_0 + \sum_{j=1}^{J} A_j c_{j,i} + B_0 R_{M,t} + \sum_{k=1}^{K} B_k \left[y_{k,i} R_{M,t} \right] + \sum_{l=1}^{L} B_{l+K} \left[z_{l,t-1} R_{M,t} \right] + \varepsilon_{i,t}, \qquad \begin{array}{l} i = 1, \ldots, N, \\ t = 1, \ldots, T, \end{array} \tag{2}$$

[17]The average beta is the unconditional mean of the conditional beta with respect to the instruments. The linear response coefficients can be thought of as an approximation of a Taylor expansion around their means, ignoring the higher order terms if the responses are in fact non-linear.

where there are J controls for investor types $c_{j,i}$, which then vary between individuals. The controls are also demeaned to preserve the interpretation of A_0 as the measured average abnormal return.

We can identify the parameters in (2) with the following moment conditions

$$\begin{aligned} E\left(\varepsilon_{i,t}\right) &= 0, \quad \forall\, i, \\ E\left(\varepsilon_{i,t} c_{j,i}\right) &= 0, \quad \forall\, i, j, \\ E\left(\varepsilon_{i,t} R_{M,t}\right) &= 0, \quad \forall\, i, \\ E(\varepsilon_{i,t} R_{M,t} y_{k,i}) &= 0, \quad \forall\, i, k, \\ E(\varepsilon_{i,t} R_{M,t} z_{l,t-1}) &= 0, \quad \forall\, i, l, \end{aligned} \tag{3}$$

such that the error term is uncorrelated with the explanatory variables. The moment conditions in (3) are estimated with GMM, but the point estimates coincide with those obtained by OLS.[18] The main difference is that the variance-covariance matrix allows for both heteroscedasticity and autocorrelation. This is a desirable feature, because standard methods run a clear risk of overstating the precision of the estimates.[19]

The sample moment conditions corresponding to (3) are explicitly considered in Appendix B along with other details regarding the estimation procedure.

4.2.1 Selection of variables

The performance analysis is conducted directly in the panel. The previous preliminary analysis found that portfolio size, turnover, and the number of stocks in the portfolio can be important determinants of cross-sectional abnormal performance. These variables are therefore chosen to parameterize the intercept. The same variables are used to control for beta risk across investors. In addition, age and gender are included as controls for heterogeneous risk between investors. The paragraphs below explain these choices.

One of the reasons for the difference in average performance between equally and value-weighted performance may be that risk is related to portfolio size. This can be linked to a relative risk aversion argument: an individual could be prepared to gamble small amounts compared to the level of wealth. Such an investor is likely to take high risk compared to the investor who has more at stake. If this is an important feature of the data, it will be controlled for. Since portfolio size is extremely skewed, the logarithm of the individual size measure is used.

Turnover can be important in two ways. Technically, high turnover could mean that

[18] This follows directly from the OLS assumption $E(uX) = 0$, by substituting for u and solving for the parameters of the model.

[19] The standard OLS assumption referred to here is that errors are independently distributed. This is clearly too strict an assumption for the data set under consideration.

cash is held in the portfolio. This may affect the return measure, since trading capital increases, and ultimately lower our beta estimates.[20] Alternatively, high turnover investors might in fact choose less risky investment strategies. Further, as turnover is included among the intercept terms, it is also a desirable control variable for risk. This is also the case for diversification, defined by the number of stocks. In addition, the degree of diversification could also be correlated with risk, since the allocation to the technology sector varies with the number of stocks held.

Age will matter when old investors have less human capital as a resource for future income; they may prefer to take lower stock market risk. The reason, as shown by Bodie, Merton, and Samuelson (1991), is that such investors have less flexibility than younger investors to adjust their labor supply and consumption if savings were to deteriorate. In this sample, older people may simply find high beta technology stocks less attractive than young people do.

Barber and Odean (2001a) argue that men are more overconfident than women; their study confirms that men trade more, and therefore do not perform as well as women. If men are more overconfident, they may also load up on more systematic risk. Also, Levin, Snyder, and Chapman (1997) find, in an experimental setting involving gambles, that women tend to be more risk-averse than men. A gender dummy for women investors is therefore included in the riskadjustment.

When specifying the information or state variables, it is difficult to know what information is relevant. The work of Keim and Stambaugh (1986), and Campbell (1987) shows that lagged stock and money market variables can have significant predictive power for the market risk premium. With the obvious risk of data snooping, the stock index return, the level of the 30-day Treasury bill and the yield spread between a 10-year and 1-year government bond are included in the regressions.

4.2.2 Regression results

The first column of Table 8 marked Model I shows that investor monthly performance is around -1.29% in an unconditional single-index specification. The beta is around 1.4, reflecting that investors in this sample take on considerable market risk.

The second and third regression condition the beta on time variation and heterogeneity in the cross-section. The average performance increases to -0.74%, the average beta increases and its standard deviation falls. This shows that the conditional model indeed controls for important variation in the betas and that the unconditional specification is misleading. In fact, the average underperformance is no longer significant, even though it is still highly negative.

[20]This is only true for investors who liquidate or acquire a *total* net position. Rebalancing a portfolio does not imply a change in trading capital itself.

The objective of the panel model is more about seeing how performance varies with investor characteristics than making inferences about risk. The results for the conditional betas are therefore only discussed briefly. The betas vary significantly with the characteristics in the cross-section, but the effect is generally small. For instance, a positive, one-standard deviation change in age above the mean produces about the same effect as when the investor is female: a decrease in beta of around 0.03. The negative effects of turnover and diversification are larger but still small: about two or three times larger than for age. The beta coefficient for portfolio size was never significant, so it is excluded from all regressions. The relatively low variability in the betas raises some concern as to whether some cross-sectional variation in risk is not controlled for. However, a robustness check at the end of this section confirms all of the main results that follow from the inferences made in Table 8.

The adjusted R^2 is reported in Table 8, even though it is difficult to interpret when we have observations in two dimensions. Nevertheless it gives an indication of how substantial the idiosyncratic component of the returns is for the overall sample. About 30 percent of the total variation is explained by the models under consideration.

The key results regarding investor performance and characteristics are reiterated by the panel regressions. Model IV reports the results of including turnover and portfolio size alone, and Model V when the diversification variable is also included.

The separate effect of turnover is a loss in performance of 1.8 basis points for each percent increase in turnover. This relation is found to be somewhat convex; that is, the marginal effect of turnover diminishes for the majority of investors.[21] These results are quite devastating for most investors. The coefficient estimates imply that around 85 basis points are lost in monthly performance when controlling for portfolio size for the group of most active traders. One should keep in mind that the median portfolio in the sample is small, which is an important explanation for the turnover effect. The results of Model IV imply that investors whose portfolio is twice the size of the mean investor, or around SEK 35,000, gain almost 15 basis points compared to the sample average.

In the final model considered in Table 8—labelled Model V—the diversification variable is included in the regression. The coefficients for turnover are slightly reduced, but the marginal effect of portfolio size is much smaller. This is because the effect of portfolio size is somewhat crowded out by diversification, as the number of stocks and size are positively correlated variables. Performance is unlikely to be a linear function of the number of stocks held over a wider range of stock holdings, but tests for non-linearities did not produce any significant results. Investors who hold one more stock than the av-

[21] The break-point where the effect of turnover increases marginally is around 250%, and is overall positive at 500% per month. There are 16 investors that trade more than 500%. Their average monthly total excess return is -0.13% which is clearly above the overall sample average. Such “extreme traders” could be therefore be outperforming in the sample.

Table 8: Panel regression estimates: Main results

There are five regressions measuring performance in the panel. Model I is a simple unconditional, single-index model, and Model II conditions beta risk on the state variables as described in the text. Model III adjusts for risk in the cross-section as well, whereas IV and V characterize the intercept on trading, portfolio size, and the degree of diversification, measured as the number of stocks in the portfolio. The dependent variable is the investors' total excess return, and there are 251,879 observations and 16,831 investors in all cases.

Group of coeffi-cients	Model name	I Unconditional	II Conditional	III Conditional	IV Conditional	V Final model
	Dependent variable	Investor excess return	Investor excess return	Investor excess return	Investor excess return	Investor excess return
		Intercept parameters, %				
A_0	Average intercept	1.292* (0.679)	-0.747 (0.614)	-0.740 (0.607)	-0.740 (0.612)	-0.741 (0.613)
A_j	Turnover_i	-	-	-	-1.771*** (0.642)	-1.731*** (0.653)
	Squared turnover_i	-	-	-	0.380** (0.162)	0.352** (0.157)
	Log port-folio size_i	-	-	-	0.215** (0.105)	0.157* (0.087)
	Diversifi-cation_i	-	-	-	-	0.071* (0.041)
		Beta parameters				
B_0	Average Beta	1.415*** (0.105)	1.436*** (0.062)	1.444*** (0.061)	1.444*** (0.062)	1.443*** (0.062)
B_k	Log Age_i	-	-	-0.103*** (0.032)	-0.101*** (0.031)	-0.101*** (0.031)
	Female_i	-	-	-0.027* (0.015)	-0.028* (0.015)	-0.028* (0.015)
	Turnover_i	-	-	-0.154*** (0.040)	-0.182*** (0.037)	-0.185*** (0.037)
	Diversifi-cation_i	-	-	-0.036*** (0.006)	-0.035*** (0.006)	-0.033*** (0.007)
B_l	Lagged index return_{t-1}	-	2.115*** (0.690)	2.203*** (0.680)	2.206*** (0.678)	2.201*** (0.677)
	Long m. short bond_{t-1}	-	0.247** (0.110)	0.248** (0.111)	0.248** (0.112)	0.248** (0.112)
	Short rate_{t-1}	-	5.377*** (1.870)	5.117*** (1.904)	5.118*** (1.907)	5.132*** (1.902)
	Adjusted R^2	0.296	0.307	0.309	0.310	0.310

Newey and West (1987) standard errors, robust with respect to autocorrelation and heteroscedasticity, are given in parentheses. Significant parameter estimates at the 10%, 5%, and 1% level are marked (*), (**), and (***). The null hypothesis for the average beta is B_0=1.

erage investor gain an additional 7 basis points in performance. This means that those in the top quintile of diversified holdings gain some 30 basis points over the average.

The negative intercept of 74 basis points implies an annualized underperformance of about 8.5%. Barber and Odean (2000) find that a portfolio consisting of the top quintile of the most active investors loses around 7% annually compared to those who do not trade. The two results are related in that Barber and Odean's most active investors trade about as much as the *average* investor considered here.

To put the model to additional tests, the regressions of Table 9 use the specified final model with alternative assumptions. Model VI and VII substitute for the dependent variable, and instead use the investor return excluding fees and the passive return. The first return includes trading, but at zero cost; the second measures the return on the fixed portfolio held at the beginning of the month. When fees are excluded, the mean performance increases by roughly 21 basis points. When trading is disregarded altogether, it improves by 26 basis points. None of the intercepts is significant, but the means reiterate the evidence reported earlier that investors would have been better off not trading even if it was costless. The coefficients for turnover and portfolio size diminish, and are now insignificant. However, the coefficient for diversification is virtually unchanged. This is the case for both Model VI and VII, which consolidates the evidence that the parameter for diversification picks up performance that is unrelated to trading and portfolio size. There is no support for a more general negative effect of portfolio size that was found in Panel B of Table 3 when fees are excluded. The size-coefficient is positive but insignificant in both specifications that exclude fees.

As discussed previously, investors who on average trade more may hold a larger proportion of cash in their portfolio, which in turn may affect the measured risk. The coefficient for turnover in Model VII does not support this hypothesis. It is smaller, but still significantly negative, which indicates that high turnover investors also hold passive portfolios that are less risky on average.

The survivorship bias in the sample is likely to be large due to the high attrition rate. Only investors active at sample-end are included in Model VIII reported in Table 9, and the intercept shows that the average performance increases by 11 basis points. On a yearly basis, this means that the survivorship bias in the sample is in the vicinity of 1.3%, which is about double the size usually found for mutual funds. Further, the effect of diversification is virtually unchanged, indicating that this effect is not driven by investors leaving the sample.

The sample does not enable us to distinguish between individuals who are new to the stock market and those who have owned stocks before. A very crude way of defining new or inexperienced investors is to remove those investors who deposited their first portfolio. These individuals could not have been new to the stock market when they

Table 9: Panel regression estimates: Additional results

There are five regressions measuring performance in the panel all based on Model V in Table 8. Model VI and VII substitute for the dependent variable with the total return excluding fees and the passive return. Model VIII is specified for the subset of investors active at sample-end. Model IX excludes the investors who deposited stocks as they entered the sample, and Model X are the investors who entered the sample in February 2000 or earlier.

Group of coefficients	Model name	VI Costless	VII Passive	VIII Survivors	IX New	X Early
	Dependent variable	Excluding fees	Passive return	Investor excess return	Investor excess return	Investor excess return
	No. of investors	16,831	16,831	13,917	11,416	2,218
	No. of obs.	251,879	251,879	228,702	175,240	49,583
		Intercept parameters, %				
A_0	Average intercept	-0.533 (0.613)	-0.481 (0.617)	-0.729 (0.605)	-0.799* (0.445)	-0.929 (0.588)
A_j	$Turnover_i$	-0.761 (0.580)	-0.174 (0.180)	-2.181*** (0.782)	-1.902 (1.280)	-1.261** (0.635)
	Squared $turnover_i$	0.149 (0.142)	0.058 (0.126)	0.511** (0.247)	0.544 (0.417)	0.264** (0.118)
	Log portfolio $size_i$	0.049 (0.086)	0.030 (0.086)	0.159* (0.088)	0.037 (0.147)	-0.005 (0.096)
	$Diversification_i$	0.085** (0.041)	0.082** (0.041)	0.083* (0.044)	0.066** (0.032)	0.080*** (0.021)
		Beta parameters				
B_0	Average Beta	1.444*** (0.063)	1.449*** (0.062)	1.444*** (0.060)	1.287*** (0.043)	1.329*** (0.064)
B_k	Log Age_i	-0.099** (0.031)	-0.100*** (0.035)	-0.101*** (0.025)	-0.165*** (0.051)	-0.245** (0.096)
	$Female_i$	-0.028* (0.015)	-0.029* (0.015)	-0.027 (0.017)	-0.058** (0.023)	-0.035 (0.031)
	$Turnover_i$	-0.179*** (0.039)	-0.093*** (0.034)	-0.178*** (0.054)	0.068 (0.106)	-0.170*** (0.053)
	$Diversification_i$	-0.033*** (0.007)	-0.038*** (0.007)	-0.035*** (0.007)	-0.011*** (0.004)	-0.014*** (0.005)
B_l	Lagged index $return_{t-1}$	2.205*** (0.682)	2.299*** (0.676)	2.190*** (0.626)	1.643*** (0.589)	1.875** (0.911)
	Long m. short $bond_{t-1}$	0.247** (0.111)	0.241** (0.112)	0.254** (0.112)	0.090 (0.111)	0.093 (0.149)
	Short $rate_{t-1}$	5.160*** (1.889)	4.839** (1.900)	5.169*** (1.881)	1.794 (1.666)	3.439 (2.222)
	Adjusted R^2	0.310	0.311	0.327	0.303	0.275

Newey and West (1987) standard errors, robust with respect to autocorrelation and heteroscedasticity, are given in parentheses. Significant parameter estimates at the 10%, 5%, and 1% level are marked (*), (**), and (***). The null hypothesis for the average beta is B_0=1.

became investors at this brokerage firm. Model IX in Table 9 marked "New" reports the performance for those 11,416 investors who bought their first portfolio. The mean performance for this group is about 80 basis points, which is 6 basis points lower than the sample average. This is a small difference, which is also insignificant when modelled as a fixed effect in the total sample.

It would be interesting to discover whether those who entered the market early performed better than those who came in late. It is difficult to partition the sample into a "bull" and a "bear" market, because there are too few observations during the first part in order to enable any reasonable estimates. In addition, it may not be of much interest to find that some investors experienced high gains during the sharp upturn. There is considerable idiosyncratic noise, making it difficult to conclude if investors were market timers or simply lucky. But if these investors were clever enough to time the market in the upturn, one might also claim that they should have been able to perform better in the downturn. Model X takes the 2,218 "Early" investors who were active in the sample before March 2000 and measure the performance of this group alone. The mean performance of this group is actually much lower than for the whole sample. Since these investors lose 93 basis points on average, compared to 74 for the whole sample, there is no evidence that early investors perform better on average.

4.2.3 Robustness

As an additional test, the same regression model is applied to the 3,367 investors sorted into the highest and lowest quintiles by portfolio size, turnover and diversification. This specification is more demanding, as the regression coefficients now describe the variation within groups rather than across quintiles as in the full sample case. A crude measure of the effect between quintiles is now found in the overall means of the regressions. The results are reported in Table 10.

The average intercepts and betas all confirm the effects that were measured in the overall sample. The investors in the largest portfolio size quintile outperform the smallest by 26 basis points. Similarly, those with the lowest turnover gain 30 basis points more than those who trade the most on average, and those with many stocks in their portfolio gain 36 basis points more than the least diversified investors. The average beta for the investors with the largest portfolios is lower than for for those with the smallest. This suggests that there is a difference in average beta risk between these groups, even though it was not significant for the whole sample.

The parameter estimates for the intercept terms broadly confirm that the previous conclusions hold for the larger portfolios. This is important, because it confirms that the previously reported results are not driven by the many small-sized accounts in data, but are also a valid characterization of those with large portfolios.

Table 10: Panel regression estimates: Investors sorted into top and bottom quintiles

The six regressions are conducted on the first and fifth quintile for portfolio size, turnover and diversification. The dependent variable is total return, and each quintile contains around 3,366 investors. The parameters for turnover and number of stocks are excluded for quintile 1 in the regressions when they are sorted by turnover and number of stocks due to near collinearity.

Group of coefficients	Model name	Portf. size, Quintile 1	Portf. size, Quintile 5	Turnover, Quintile 1	Turnover, Quintile 5	Diversif., Quintile 1	Diversif., Quintile 5
	No. of obs.	43,160	54,372	36,373	37,041	41,523	54,787
		Intercept parameters, %					
A_0	Average intercept	-0.830 (0.800)	-0.570 (0.515)	-0.706 (0.731)	-1.074^{*} (0.641)	-0.890 (0.847)	-0.528 (0.460)
A_j	Turnover_i	-1.934 (1.441)	-1.113^{**} (0.497)	n/a	-0.757^{**} (0.343)	2.588 (1.895)	-1.876^{***} (0.575)
	Squared turnover_i	1.049 (1.196)	0.205^{*} (0.110)	n/a	0.192^{**} (0.097)	-0.691 (0.621)	0.349^{***} (0.124)
	Log portfolio size_i	0.101 (0.417)	-0.072 (0.124)	0.180 (0.207)	0.195^{*} (0.102)	0.135 (0.206)	0.111^{**} (0.055)
	Diversification_i	0.098 (0.192)	0.084^{***} (0.030)	0.109 (0.091)	0.015 (0.036)	n/a	0.048^{**} (0.020)
		Beta parameters					
B_0	Average Beta	1.584^{***} (0.084)	1.360^{***} (0.057)	1.452^{***} (0.057)	1.353^{***} (0.071)	1.645^{***} (0.083)	1.277^{***} (0.049)
B_k	Log Age_i	0.008 (0.046)	-0.172^{***} (0.061)	-0.061 (0.038)	-0.130 (0.087)	-0.140^{***} (0.029)	-0.104^{***} (0.028)
	Female_i	0.008 (0.018)	-0.077^{***} (0.015)	-0.037 (0.025)	-0.012 (0.025)	-0.012 (0.018)	-0.047^{***} (0.016)
	Turnover_i	-0.391^{***} (0.080)	-0.156^{***} (0.031)	n/a	-0.251^{***} (0.021)	-0.106 (0.078)	-0.130^{***} (0.034)
	Diversification_i	-0.083^{***} (0.026)	-0.024^{***} (0.005)	-0.061^{***} (0.018)	-0.014^{**} (0.006)	n/a	-0.015^{***} (0.003)
B_l	Lagged index return_{t-1}	3.600^{***} (0.846)	1.564^{**} (0.624)	2.133^{***} (0.543)	2.163^{**} (0.879)	2.815^{***} (0.787)	1.892^{***} (0.539)
	Long m. short bond_{t-1}	0.438^{***} (0.117)	0.150 (0.110)	0.220^{*} (0.133)	0.293^{**} (0.132)	0.366^{**} (0.149)	0.136 (0.091)
	Short rate_{t-1}	6.848^{*} (3.604)	4.472^{***} (1.557)	3.546^{*} (2.067)	7.781^{***} (2.208)	5.707^{**} (2.513)	3.882^{***} (1.481)
	Adjusted R^2	0.272	0.378	0.267	0.266	0.237	0.472

Newey and West (1987) standard errors, robust with respect to autocorrelation and heteroscedasticity, are given in parentheses. Significant parameter estimates at the 10%, 5%, and 1% level are marked (*), (**), and (***). The null hypothesis for the average beta is $B_0=1$.

The regression for the small portfolios is much noisier, and therefore many of the parameters are insignificant. This is also true for turnover and diversification, as investors are, on average, small in these quintiles as well. In these two last cases, parameters are excluded due to the problem of collinearity. There is little or no variation in turnover and diversification for the lowest quintiles, making these variables impossible to distinguish from their average intercept and beta coefficients.

Turnover has generally a negative effect on performance, except for investors who are least diversified. The coefficients here switch signs, and indicate a strong positive effect. This finding supports a learning behavior where some investors choose to diversify as they become aware of their unprofitable strategy. It could also be that some of these investors simply benefited from selling their stock. Due to the weak significance of this result, the only conclusive evidence is that trading does not harm the least diversified investors to the same extent as the other investor groups.

In conclusion, the general results broadly hold when partitioning the sample into investor groups, and the regression means reveal important differences between them. The parameters in the top quintiles for each group also indicate significant variation within the studied investor groups.

5 Conclusion

Investor performance can be attributed to several, partly interacting, investor characteristics. The discovered systematic pattern of investor performance deepens our knowledge of the trading behavior of online investors in general, and the relation between performance and characteristics in particular.

Online investors trade aggressively in small portfolios, which means that the commissions they pay are high even if fees are low in absolute terms. The average investor flips her portfolio twice a year, and the 20% who trade the most do so on average seven times a year. The marginal effects of turnover reveal that investors who do not trade gain around 25 basis points more per month than the average investor. But trading is not equally as harmful for those with larger portfolios. Portfolios that are twice the size of the sample average gain 15 basis points per month in performance. The combined effects of turnover and portfolio size are mainly related to fees, as they are insignificant when trading is costless.

The novel finding in this study is that undiversified investors systematically choose underperforming stocks in any given industry, and thus the degree of diversification is also important in explaining cross-sectional differences in performance. The quintile of investors who are best diversified earn 36 basis points per month more than those who are least diversified. The panel regressions confirm that diversification has a separate and distinct effect that is unrelated to portfolio size. Benartzi and Thaler (2001) suggest that mutual fund investors diversify naively over available assets. I find that it is the overall lack of diversification among equity investors that can be linked to performance. Undiversified investors are overconfident in their own stock-picking ability, because they are shown to take higher risks and underperform more. The choice of stocks could be explained by a naive strategy based on availability in the way proposed by Barber and Odean (2003). Undiversified investors show a clear preference for attention-grabbing technology stocks.

I propose that the explanation for the positive effect of diversification on performance lies in the degree of investor sophistication. Unsophisticated investors are more inclined to follow heuristics than common advice. It is tempting to conclude that individuals investing in one stock rather than a mutual fund are widely unaware of the most basic textbook advice on portfolio diversification. But we need to interpret with care, because they might have other holdings of financial assets than those observed in the sample. Therefore, this explanation is only speculative. On the other hand—if there are other stock holdings—the observed portfolio in this sample must contribute to the investor's overall utility in some way. I argue that this is possible, but unlikely.

First, the observed portfolio could provide necessary negative correlation to some other assets held so as to offset overall risk in the aggregate portfolio. I find this unlikely, considering that the stocks held are primarily high-beta, technology stocks. Second, investors may be constrained by being unable to borrow the funds needed to obtain the desired level of risk. This, I believe, is also unlikely. There are well-diversified mutual funds that track most industries, and that would serve as a low-cost alternative to these individual stocks. Third, investors might simply enjoy gambling, betting on single stocks for the sheer fun of it. Such motives are hard to reject, but they do not explain *why* these investors are less successful than others in selecting stocks.

An interesting question for future research is to understand how stock holdings relate to other investor characteristics, such as total wealth, occupation and education. Such variables are also likely to be useful proxies for investor sophistication, and in turn, the profitability of investment strategies.

In summary, most online investors behave contrary to conventional wisdom: They put all their eggs in one basket and count their chickens before they are hatched. Online investors showed guts in taking risks, but few gloried in it.

Appendix A: Measurement

Let $x_{n,i,t}$ be the number of shares of a stock n held by the individual i at the end of month t. A transaction d during month t is denoted by $x_{n,i,d}$, and super-indices B and S indicate whether it is a buy or a sell transaction. Similarly, associated actual purchasing and sales prices net of fees are denoted $p^B_{n,i,d}$ and $p^S_{n,i,d}$ for each of these transactions. In what follows, we also need the closing price for stock n on the last day of month t, which is labelled $p^C_{n,t}$. The stock position for individual i at the end of month t is

$$x_{n,i,t} = x_{n,i,t-1} + \sum_{d \in t} \left(x^B_{n,i,d} - x^S_{n,i,d} \right), \tag{A1}$$

which is the position at the beginning of month t plus the sum of buys and sells during the month, hereafter net purchases for short. In what follows, we will impose the restriction that $x_{n,i,t} \geq 0$, meaning that investors are not allowed to have outstanding negative positions at month-end.

A.1 Payoffs

Trading, position and total payoffs for each stock and individual are as follows. The position payoff is defined as

$$\Pi^P_{n,i,t} = x_{n,i,t-1} \cdot (p^C_{n,t} - p^C_{n,t-1}), \tag{A2}$$

which is simply the position at the beginning of the month times the change in price. The trading payoff in stock n for individual i during month t is given by

$$\Pi^T_{n,i,t} = \sum_{d \in t} p^S_{n,i,d} \cdot x^S_{n,i,d} - \sum_{d \in t} p^B_{n,i,d} \cdot x^B_{n,i,d} + \sum_{d \in t} \left(x^B_{n,i,d} - x^S_{n,i,d} \right) p^C_{n,t}. \tag{A3}$$

The first and second component of (A3) states the net sales revenue of stock n during month t, which is the value of sells minus buys at actual transacted prices. This value needs to be adjusted if the number of stocks sold exceeds sales, or vice versa. The third component of (A3) adjusts payoffs by the value of net purchases. There are two cases. If net purchases is positive, the payoff is adjusted by multiplying the net increase in the number of shares by the price at the end of the month. If sales exceed buys, there will be stocks included in the trading payoff by (A3) that are already accounted for by (A2). Therefore, the value of these shares at t is deducted from the trading payoff.[22]

Deposits of stocks are assumed to be transacted at the beginning of the month and

[22] The method applied is therefore related to that of Linnainmaa (2003), who investigate the profitability of daytrades. The main difference here is that payoffs for positions are invariant to which stocks are actually sold. Furthermore, the capital components associated with the payoffs that follow are quite differently defined.

redemptions at the end. Therefore, $x_{n,i,t-1}$ also includes all deposits of stocks made during the month. This is the most convenient way to include deposits since they cannot be regarded as traded stocks. It would be a mistake not to include redemptions and deposits as there would be at least some individuals who deposit their portfolio, but do not trade.

Investors are allowed to short-sell their stock with these definitions because the summation is invariant to the ordering of purchases and sales. The restriction only means that there must be a positive holding of each stock at the end of the month.[23]

Total payoff for each investor i in stock n is just the sum of trading and position payoff

$$\Pi_{n,i,t} = \Pi^T_{n,i,t} + \Pi^P_{n,i,t}. \tag{A4}$$

To find the payoff for the whole portfolio, we sum over n to obtain total portfolio payoff for individual i in month t

$$\Pi_{i,t} = \sum_n \Pi_{n,i,t} = \sum_n \Pi^T_{n,i,t} + \sum_n \Pi^P_{n,i,t}. \tag{A5}$$

A.2 Capital components

The task is to measure investors' ability to create value in their portfolios over a fixed time frame while accounting for trading. The key issue is to identify the capital tied to the payoff components at the portfolio level. The definition of trading capital is complicated by the fact that investors who trade extensively may turn around their portfolio many times per month. For instance, an investor may sell her complete holdings of one stock and invest in another during the month. The capital required for the initial holding and the trading capital needed for the purchase is one and the same.

To facilitate comparisons, we assume here that investors hold unleveraged portfolios. They are unconstrained in that they can borrow cash freely to cover the cost of any net purchases at the portfolio level. In this case, the capital required for trading is the minimum amount of money needed to finance the portfolio.

Similar to payoffs, we distinguish between position and trading capital as follows. Position capital is defined as

$$C^P_{i,t} = \sum_n \left(x_{n,i,t-1} \cdot p^C_{n,t-1} \right), \tag{A6}$$

which is simply the value of all stocks in the portfolio at the beginning of the month.

The amount of capital engaged in trading is determined in two steps. I begin by matching purchases and sales. For each investor, the trades are sorted on a stock by

[23] In the sample, this proved to be a minor problem as there were only 34 instances where it was needed to cover open short positions at month's end. This was done by dating the corresponding buy transaction at the beginning of the following month, $t+1$, as belonging to t.

stock basis in calendar time. Buy transactions are assumed to precede sales. This is to ensure that the investor does not borrow any stocks in the portfolio.

In step two, we begin by defining the traded value of any sale or buy as

$$TV_{i,d} = \left\{ \begin{array}{c} p_{i,d} \cdot x_{i,d}^J \text{ if } J = S \\ -p_{i,d} \cdot x_{i,d}^J \text{ if } J = B \end{array} \right\},$$

such that it represents the revenue of any sales and cost of any purchase independent of the stock that is traded. We then seek the lowest cost that is needed to finance the trading activity during the month. The trade values are ordered during the month from beginning to end for each investor regardless of which stock is traded, and the cash balance is calculated at each point in time. The lowest cumulative cash balance in month t is the minimum amount needed to finance the portfolio without leverage, and is written

$$C_{i,t}^T = -\min_d \left[\sum_{d \in t} TV_{i,d}, 0 \right], \tag{A7}$$

and is expressed as a positive number since we pick out the largest negative cash balance.

Total capital is the sum of position capital and trading capital,

$$C_{i,t} = C_{i,t}^P + C_{i,t}^T. \tag{A8}$$

Therefore, the capital base is only increased if trading incurs additional funding. But this is exactly what we want, because the investor who reallocates her investment without using additional funds will have the same capital base.

A.3 Simple returns

The simple portfolio return for investor i in month t is

$$r_{i,t} = \frac{\Pi_{i,t}}{C_{i,t}}. \tag{A9}$$

If no trading occurred in month t, it is easy to verify that this expression corresponds to

$$r_{i,t} = \left[\sum_n w_{n,i,t-1} \cdot \frac{p_{n,t}^C - p_{n,t-1}^C}{p_{n,t-1}^C} \right],$$

which is the weighted return of the portfolio held in $t-1$, and where $w_{n,i,t-1}$ is the weight of stock n held by individual i in $t-1$.

A.4 Excess returns

The obvious problem when constructing returns from the definitions above is that no account is taken of any alternative return on funds that is not invested in the market. For example, consider an investor who buys stocks at the end of the month. This portfolio will have a capital base that reflects the value of the additional purchases at the beginning of the month, but a stock return measured over a much shorter horizon.

This effect is mitigated by measuring excess returns, created as follows. It is assumed that the investor borrows at the available 30-day T-bill rate, r_{t-1}^F, in order to finance the portfolio. The interest that is attributable to the position component, $I_{i,t}^P$, is calculated as the cost of borrowing the value of the portfolio at the beginning of the month, i.e the first part of equation (A8).

If trading occurs, we seek the net interest paid for trading capital during the month. Interest is calculated for each transaction and summed over the month creating the fictitious revenue $I_{i,t}^T$ that corresponds to the interest that is attributable to the actual timing of purchases and sales.[24]

The excess return is therefore

$$R_{i,t} = \frac{\Pi_{i,t} + I_{i,t}^P + I_{i,t}^T}{C_{i,t} - \min\left[I_{i,t}^T, 0\right]}, \tag{A10}$$

where $I_{i,t}^P$ is always 0 or negative and $I_{i,t}^T$ is negative if there is a net cost of financing the monthly transactions. When trading capital is 0 but the investor is net selling, $I_{i,t}^T$ represents the interest earned on investments that is sold out of the portfolio. In this way, timing of the sale is properly accounted for since positive interest is added to the return measure. Both trading capital and $I_{i,t}^T$ can be positive if the investor only draws cash for a short time and for a small amount in comparison to sales revenues in a month.

The interest on trading capital is only added to the capital base if it is negative. This is because it is assumed that interest earned is paid out at the end of the month, but any costs must be covered by capital at the beginning of the month. It therefore ensures that returns are bounded at -1.

In the case of no trading, $I_{i,t}^T = 0$, we obtain the familiar definition of excess returns, which is

$$R_{i,t} = \frac{\Pi_{i,t} + I_{i,t}^P}{C_{i,t}^P} = r_{i,t} - r_{t-1}^F.$$

[24]Two assumptions apply: the borrowing and lending rates are the same and the effect of compounding is ignored.

A.5 Passive returns

The idea of comparing investor performance with their own-benchmark was originally proposed by Grinblatt and Titman (1993). Here, the passive return captures the same intuition and is defined as

$$R_{i,t}^{P} = \frac{\Pi_{i,t}^{P} - I_{i,t}^{P}}{C_{i,t}}, \tag{A11}$$

which is the position payoff divided by total capital corrected for interest. As the own-benchmark return measures the return of the portfolio since there was no trading during the month, we see that (A10) and (A11) are exactly the same, because then we have that

$$\Pi_{i,t} = \Pi_{i,t}^{P}.$$

The passive return measure uses total capital as a base. It is therefore assumed that whatever funds are used for net investments during the month are invested at the risk-free rate. The investors can only deviate from the benchmark by trading. During the month, investors can move in and out of the market as a whole or switch allocation between stocks. If these tactical changes in risk and reallocations are profitable, the investors earn a higher excess return on the traded portfolio than on the static own-benchmark.

A.6 Industry returns

The industry weight in industry v for individual i at time $t-1$, $w_{v,i,t-1}$, is obtained by summing the weights of any stocks n the individual holds in each industry v,

$$w_{v,i,t-1} = \sum_{n \in v} w_{n,i,t-1}. \tag{A12}$$

Thus, the investor's industry tracking return is the industry weight multiplied by the respective excess index return,

$$R_{i,t}^{Ind} = \sum_{v} R_{v,t}^{Ind} \cdot w_{v,i,t-1}. \tag{A13}$$

A.7 Turnover

As a measure of turnover, the total traded value is divided by two times total capital,

$$\text{Turnover}_{i,t} = \frac{\sum_{d \in t} |TV_{i,d}|}{2 \cdot C_{i,t}}. \tag{A14}$$

Turnover therefore includes both position and trading capital in the denominator. Total capital is doubled so we can interpret the measure as how often a portfolio is bought and purchased in a month. A turnover measure of one thus implies that the whole portfolio is sold and purchased.

A.8 Concluding example

The calculation of the various returns defined above is illustrated in Table 11 by two simplified examples of two investors holding or trading two different stocks.

Paul initially holds 50 A stocks and 100 B stocks at prices 90 and 50, respectively. He makes one trade during the month, which is an additional purchase of 100 stock A at price 92. At the end of the month, the stock prices are 100 for A and 45 for B. The total payoff for A is 1,300, of which 500 is attributable to the position and 800 to trading. Since there was no trading in B, trading payoff is 0, but there is a position payoff of -500. Total payoff is therefore 800.

The position capital needed to finance this portfolio is 4,500 for A and 5,000 for B, i.e., 9,500 in total. Furthermore, the additional stocks A bought cost 9,200. Since there are no more trades, this is the lowest cash balance during the month. Therefore, trading capital and total capital sums up to 18,700. All in all, this yields a total portfolio return of 4.28%. The passive return, given by the return on the portfolio held in $t-1$, is 0%. Turnover, which is the value of the purchases divided by two times total capital, is almost 25%, indicating that this month Paul bought and sold a quarter of his portfolio.

The other investor—Magnus—starts out with 100 A stocks and makes three trades. The ordering of the trades are marked by super-indices. He begins by buying 50 B stocks to price 45. Later, he sells 140 A stocks (such that in effect he short-sells 40 A stocks) at price 85. Finally, he decides to buy back 100 A stocks at price 95.

Magnus generates a trading payoff of 250 in B as the stocks that were bought at 40 are each worth 45 at month-end. The trading payoff for A is calculated as follows. The value of sales minus purchases is 2,680 and is adjusted by 40 stocks valued at 100, such that the trading payoff is -1,320 in all. A position payoff of 1,000 is recorded for the stocks owned at the beginning of the month and held to the end, such that the total payoff for A is -320.

We can convince ourselves that this is indeed correct by noting that the monthly mean purchasing price of A stocks is 92.50. Magnus owned 100 A shares that were worth 90 at the beginning of the month, and 100 shares was bought at 95. Magnus incurred a loss of 770 when 140 shares were sold at 87, but gained on the remaining 60 stocks that were kept to month-end. The 7.50 profit on each of these 60 shares amounts to 450. Losses and profits come to -320.

Position capital is defined by the value of the holdings, which in this case is 100 A stocks to the value of 9,000. We retrieve the trading capital by considering the order of the trades. The lowest value we obtain summing over the trades is 2,000, which is needed

Table 11: Two examples of return measure

The return measure is illustrated by two examples reflecting a one-month investment history of two investors. The monthly $t-1$ and t closing prices are 90 and 100 for stock A and 50 and 45 for stock B. The order for which the trades occur are marked by super-indices. Paul holds 50 units of A and 100 of B. He then buys 100 more A at price 92. Magnus holds 100 A stocks. Then 50 units of B are purchased at price 40, followed by 140 A sold at 87. Finally 100 A stocks are bought at price 95. For simplicity, the returns here are simple rather than excess returns used in the actual calculations.

	Paul*		Magnus*	
Assumptions	Stock A	Stock B	Stock A	Stock B
Closing price in *t-1*, $p^{C}_{n,t-1}$	90	50	90	50
Closing price in *t*, $p^{C}_{n,t}$	100	45	100	45
Initial position, $x_{n,i,d}$	50	100	100	0
Amount bought, $x^{B}_{n,i,d}$	100[1]	0	100[3]	50[1]
Price bought, $p^{B}_{n,i,d}$	92	-	95	40
Amount sold, $x^{S}_{n,i,d}$	0	0	140[2]	0
Price sold, $p^{S}_{n,i,d}$	-	-	87	-
Payoffs				
Trading payoff, $\Pi^{T}_{n,i}$	800	0	-1,320	250
Position payoff, $\Pi^{P}_{n,i}$	500	-500	1,000	0
Total payoff, $\Pi_{n,i}$	1,300	-500	-320	250
Capital requirements				
Trading capital, $C^{T}_{i,t}$		9,200		2,000
Position capital, $C^{P}_{i,t}$		9,500		9,000
Total capital, $C_{i,t}$		18,700		11,000
Returns				
Total return, $R_{i,t}$		4.28%		-0.64%
Passive return, $R^{P}_{i,t}$		0.00%		9.09%
Turnover		24.6%		107.6%

*) Any resemblance to actual persons or events are unintentional and purely coincidental.

to finance the first transaction.[25] Trading capital is therefore 2,000. Total capital is 11,000 which corresponds to the initial holding of 100 A stocks at price 90 plus the 50 B stocks bought at price 40.

All in all, dividing payoffs by capital, Magnus total portfolio return is -0.64%. The own-benchmark return is 9.09%, which is the return on the 100 A stocks held at the beginning of the period. The turnover is measured at 108%, which means that this month Magnus bought and sold more than the value of his portfolio.

This is a simplified example where any interest with respect to the timing of the trades are unaccounted for in the return measure. The corresponding excess returns to those here could be created by the adjustments given previously in the text.

[25]The cash balance for the second trade is 10,180, obtained by adding the 12,180 in revenue for the sales to -2,000. For the third trade, the cash balance is 680.

Appendix B: GMM estimation

The regressions of the portfolio excess return for each investor i at time t can generally be specified as follows:

$$\begin{aligned} R_{i,t} &= A_0 + B_0 R_t^M + \sum_{j=1}^{J} A_j c_{i,t,j} \\ &\quad + \sum_{k=1}^{K} B_k z_{i,t,k} R_t^M + \varepsilon_{i,t}, \qquad \begin{array}{l} i = 1, \ldots, N, \\ t = 1, \ldots, T, \end{array} \; B1 \end{aligned} \tag{1}$$

where the A:s and B:s are true regression parameters, c denotes J investor characteristics, z denotes K conditional risk attributes, and ε denotes the error term. In this general form, both the characteristics and attributes can vary over time and between individuals. In all, we have $2 + J + K$ parameters and $i = 1, \ldots, N$ individual portfolio observations over time $t = 1, \ldots, T$.

Consider the following sample moment conditions:

$$\mathbf{g}_T(\theta) = \frac{1}{T} \sum_{t=1}^{T} \begin{bmatrix} \varepsilon_{1,t} \\ \vdots \\ \varepsilon_{N,t} \\ \varepsilon_{1,t} R_t^M \\ \vdots \\ \varepsilon_{N,t} R_t^M \\ \varepsilon_{1,t} c_{1,t,j} \\ \vdots \\ \varepsilon_{N,t} c_{N,t,J} \\ \varepsilon_{1,t} z_{1,t,k} R_t^M \\ \vdots \\ \varepsilon_{N,t} z_{N,t,K} R_t^M \end{bmatrix} = \frac{1}{T} \sum_{t=1}^{T} f\left(\mathbf{X}_t, \theta\right), \tag{B2}$$

where $\mathbf{X}_t$ summarizes the data and θ contains all parameters. There are $2 + J + K$ parameters and $(2 + J + K)N$ moment conditions, so the system is over-identified. We recover the parameters by forming $2 + J + K$ linear combinations of $\mathbf{g}_T(\theta)$, that is,

$$\mathbf{A}\mathbf{g}_T(\theta) = \mathbf{0},$$

where $\mathbf{A}$ is a matrix of constants. More specifically, we let $\mathbf{A}$ be of the following form

$$\mathbf{A} = \begin{bmatrix} \mathbf{1}_{1\times N} & \mathbf{0}_{1\times N} & \mathbf{0}_{1\times JN} & \mathbf{0}_{1\times KN} \\ \mathbf{0}_{1\times N} & \mathbf{1}_{1\times N} & \mathbf{0}_{1\times JN} & \mathbf{0}_{1\times KN} \\ \mathbf{0}_{J\times N} & \mathbf{0}_{J\times N} & \mathbf{I}_{J\times J} \otimes \mathbf{1}_{1\times N} & \mathbf{0}_{J\times KN} \\ \mathbf{0}_{K\times N} & \mathbf{0}_{K\times N} & \mathbf{0}_{K\times JN} & \mathbf{I}_{K\times K} \otimes \mathbf{1}_{1\times N} \end{bmatrix},$$

where $\mathbf{1}$ denotes a vector of ones and $\mathbf{I}$ is the identity matrix. We then ensure that

$$\mathbf{A}\mathbf{g}_T(\theta) = \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} \sum_{i=1}^{N} \varepsilon_{i,t} \\ \sum_{i=1}^{N} \varepsilon_{i,t} R_t^M \\ \sum_{i=1}^{N} \varepsilon_{i,t} c_{i,t,1} \\ \vdots \\ \sum_{i=1}^{N} \varepsilon_{i,t} c_{i,t,J} \\ \sum_{i=1}^{N} \varepsilon_{i,t} z_{i,t,1} R_t^M \\ \vdots \\ \sum_{i=1}^{N} \varepsilon_{i,t} z_{i,t,K} R_t^M \end{bmatrix} = \mathbf{0}_{1\times(2+J+K)}. \tag{B3}$$

Hence, the system of moment conditions in (B3) is exactly identified. It is straightforward to show that these moment conditions correspond exactly to a least square estimator of (B1).

Hansen (1982) shows that the asymptotic distribution of the parameter estimates θ_T of the true parameter vector θ_0 is given by

$$\sqrt{T}\,(\theta_T - \theta_0) \xrightarrow{d} N(\mathbf{0}, (\mathbf{A}\mathbf{D}_0)^{-1}\left(\mathbf{A}\mathbf{S}_0'\mathbf{A}\right)(\mathbf{A}\mathbf{D}_0)^{-1'}), \tag{B4}$$

where

$$\mathbf{S}_0 = \sum_{m=-\infty}^{\infty} E\left[f\left(\mathbf{X}_t, \theta_0\right) f\left(\mathbf{X}_{t-m}, \theta_0\right)'\right],$$

and $\mathbf{D}_0$ is the gradient of the moment conditions for the true parameters. The gradient is estimated from its sample counterpart and the sample variance-covariance matrix is estimated in the manner described by Newey and West (1987). It can be shown that

$$(\mathbf{A}\mathbf{D}_0)^{-1}\left(\mathbf{A}\mathbf{S}_0'\mathbf{A}\right)(\mathbf{A}\mathbf{D}_0)^{-1'} \geq \left(\mathbf{D}_0'\mathbf{S}_0^{-1}\mathbf{D}_0\right)^{-1},$$

implying loss of efficiency when the matrix $\mathbf{A}$ is constructed arbitrarily, as is the case here. However, the standard errors in (B4) are robust to heteroscedasticity and serial correlation.

When investors enter the sample at different points in time, there are missing observations for the months in which they are not observable. Bansal and Dahlquist (2000) derive results that are used to estimate pooled models with missing data, i.e., constructing a balanced panel from an unbalanced one. They define indicator variables based on data availability according to

$$I_{i,t} = \begin{cases} 1 \text{ if data are observed at } t \text{ for individual } i \\ 0 \text{ if data are not observed at } t \text{ for individual } i \end{cases}$$

The critical assumption they make is that the indicator variable is independent of $\varepsilon_{i,t}$ which implies that the data are missing randomly. When this is the case, we can form moment conditions based on the product of the previously modelled errors and the indicator variable. This implies that, for all practical purposes, we can use the same estimation approach proposed earlier on the full sample by treating missing observations in the moment conditions as zeros. See Bansal and Dahlquist (2000) for an example.

References

Bansal, Ravi, and Magnus Dahlquist, 2000, The forward premium puzzle: Different tales from developed and emerging markets, *Journal of International Economics* 51, 115–144.

Barber, Brad M., and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 60, 773–806.

——— , 2001a, Boys will be boys: Gender, overconfidence, and common stock investment, *Quarterly Journal of Economics* 116, 261–292.

——— , 2001b, The Internet and the investor, *Journal of Economic Perspectives* 15, 41–54.

——— , 2002, Online investors: Do the slow die first?, *Review of Financial Studies* 15, 455–487.

——— , 2003, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, Working Paper, University of California, Berkeley.

Benartzi, Shlomo, and Richard H. Thaler, 2001, Naive diversification strategies in defined contribution savings plans, *American Economic Review* 91, 427–465.

Bodie, Zvi, Robert C. Merton, and William F. Samuelson, 1991, Labor supply flexibility and portfolio choice in a life-cycle model, *Journal of Economics, Dynamics and Control* 16, 427–449.

Breen, William, Lawrence R. Glosten, and Ravi Jagannathan, 1989, Economic significance of predictable variations in stock index returns, *Journal of Finance* 44, 1177–1190.

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

Campbell, John Y., 1987, Stock returns and the term structure, *Journal of Financial Economics* 18, 373–399.

Coval, Joshua D., David A. Hirshleifer, and Tyler G. Shumway, 2003, Can individual investors beat the market?, Working Paper, Harvard University.

Daniel, Kent, David A. Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security under- and overreactions, *Journal of Finance* 53, 1839–1885.

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.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425–461.

Glaser, Markus, and Martin Weber, 2003, Overconfidence and trading volume, Working Paper, University of Mannheim.

Grinblatt, Mark, and Matti Keloharju, 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, Mark, and Sheridan Titman, 1993, Performance measurement without benchmarks: An examination of mutual fund returns, *Journal of Business* 66, 47–68.

Grossman, Sanford J., and Joseph E. Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393–408.

Hansen, Lars P., 1982, Large sample properties of generalized method of moments estimators, *Econometrica* 50, 1029–1054.

Heaton, John, and Deborah J. Lucas, 1996, Evaluating the effects of incomplete markets on risk sharing and asset pricing, *Journal of Political Economy* 104, 443–487.

Keim, Donald B., and Robert F. Stambaugh, 1986, Predicting returns in the stock and bond market, *Journal of Financial Economics* 17, 357–390.

Kyle, Albert S., and Albert Wang, 1997, Speculation duopoly with agreement to disagree: Can overconfidence survive the market test?, *Journal of Finance* 52, 2073–2090.

Langer, Ellen J., and Jane Roth, 1975, Heads I win, tails its chance: The illusion of control as a function of the sequence of outcomes in a purely chance task, *Journal of Personality and Social Psychology* 32, 951–955.

Levin, Irwin P., M. A. Snyder, and D.P. Chapman, 1997, The interaction of experimental and situational factors and gender in a simulated risky decision-making task, *Journal of Psychology* 122, 173–181.

Linnainmaa, Juhani, 2003, Essays on the interface of market microstructure and behavioral finance, Doctoral dissertation, Helsinki School of Economics.

McClelland, Alastair G. R., and Fergus Bolger, 1994, The calibration of subjective probabilities: Theories and models 1980-94, in George Wright, and Peter Ayton, ed.: *Subjective Probability* . pp. 453–482 (John Wiley & Sons Ltd.: New York, N.Y.).

Newey, Whitney K., and Kenneth D. West, 1987, A simple semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Odean, Terrance, 1998, Are investors reluctant to realize their losses?, *Journal of Finance* 53, 1775–1797.

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

Schlarbaum, Gary G., Wilbur G. Lewellen, and Ronald G. Lease, 1978, Realized returns on common stock investments: The experience of individual investors, *Journal of Business* 51, 299–325.

Shefrin, Hersh, and Meir Statman, 1985, The disposition to sell winners too early and ride losers too long: Theory and evidence, *Journal of Finance* 40, 777–790.

Slovic, Paul, Baruch Fischhoff, and Sarah Lichtenstein, 1982, Facts versus fears: Understanding perceived risk, in Daniel Kahneman, Paul Slovic, and Amos Tversky, ed.: *Judgement under Uncertainty: Heuristics and Biases* . pp. 463–489 (Cambridge University Press: Cambridge, M.A.).

Svensson, Ola, 1981, Are we all less risky and more skillful than our fellow drivers?, *Acta Psychologia* 47, 143–148.

Varian, Hal R., 1989, Differences of opinion in financial markets, in Courtenay C. Stone, ed.: *Financial Risk: Theory, Evidence, and Implications* . pp. 3–37 (Kluwer Academic Publishers: Boston).

Weber, Martin, and Colin F. Camerer, 1998, The disposition effect in securities trading: an experimental analysis, *Journal of Economic Behavior and Organization* 33, 167–184.

SIFR Research Report Series

All reports can be downloaded from our website www.sifr.org, under the heading Research. Reports no. 1-15 are also available in print. In order to obtain copies of printed reports, please send your request to info@sifr.org with detailed ordering information.

1. **Foreigners' Trading and Price Effects Across Firms**
   Magnus Dahlquist and Göran Robertsson, December 2001
2. **Hedging Housing Risk**
   Peter Englund, Min Hwang, and John M. Quigley, December 2001
3. **Winner's Curse in Discriminatory Price Auctions: Evidence from the Norwegian Treasury Bill Auctions**
   Geir Høidal Bjønnes, December 2001
4. **U.S. Exchange Rates and Currency Flows**
   Dagfinn Rime, December 2001
5. **Reputation and Interdealer Trading. A Microstructure Analysis of the Treasury Bond Market**
   Massimo Massa and Andrei Simonov, December 2001
6. **Term Structures in the Office Rental Market in Stockholm**
   Åke Gunnelin and Bo Söderberg, April 2002
7. **What Factors Determine International Real Estate Security Returns?**
   Foort Hamelink and Martin Hoesli, September 2002
8. **Expropriation Risk and Return in Global Equity Markets**
   Ravi Bansal and Magnus Dahlquist, November 2002
9. **The Euro Is Good After All: Corporate Evidence**
   Arturo Bris, Yrjö Koskinen, and Mattias Nilsson, November 2002
10. **Which Investors Fear Expropriation? Evidence from Investors' Stock Picking**
    Mariassunta Giannetti and Andrei Simonov, November 2002
11. **Corporate Governance and the Home Bias**
    Magnus Dahlquist, Lee Pinkowitz, René M. Stulz, and Rohan Williamson, November 2002
12. **Implicit Forward Rents as Predictors of Future Rents**
    Peter Englund, Åke Gunnelin, Martin Hoesli, and Bo Söderberg, November 2002
13. **Accounting Anomalies and Information Uncertainty**
    Jennifer Francis, Ryan LaFond, Per Olsson, and Katherine Schipper, June 2003
14. **Characteristics, Contracts and Actions: Evidence From Venture Capitalist Analyses**
    Steven N. Kaplan and Per Strömberg, June 2003

15. **Valuing Corporate Liabilities**
    Jan Ericsson and Joel Reneby, June 2003

16. **Rental Expectations and the Term Structure of Lease Rates**
    Eric Clapham and Åke Gunnelin, October 2003

17. **Dealer Behavior and Trading Systems in Foreign Exchange Markets**
    Geir Høidal Bjønnes and Dagfinn Rime, December 2003

18. **C-CAPM and the Cross-Section of Sharpe Ratios**
    Paul Söderlind, December 2003

19. **Is there Evidence of Pessimism and Doubt in Subjective Distributions? A Comment on Abel**
    Paolo Giordani and Paul Söderlind, December 2003

20. **One for the Gain, Three for the Loss**
    Anders E. S. Anderson, May 2004

21. **Hedging, Familiarity and Portfolio Choice**
    Massimo Massa and Andrei Simonov, May 2004

22. **The Market Pricing of Accruals Quality**
    Jennifer Francis, Ryan LaFond, Per Olsson, and Katherine Schipper, May 2004

23. **Privatization and Stock Market Liquidity**
    Bernardo Bortolotti, Frank de Jong, Giovanna Nicodano, and Ibolya Schindele, June 2004

24. **Pseudo Market Timing: Fact or Fiction?**
    Magnus Dahlquist and Frank de Jong, June 2004

25. **All Guts, No Glory: Trading and Diversification among Online Investors**
    Anders E. S. Anderson, June 2004

Stockholm Institute
for Financial Research