Can Individual Investors Time Bubbles?

Jussi Keppo, Tyler Shumway and Daniel Weagley*

April 4, 2017

Abstract

We document significant persistence in the ability of individual investors to time the stock market, including during periods that people describe as bubbles. Using data on all trades by individual Finnish investors over more than 14 years, we show that active investors who successfully time the market in the first half of the sample are more likely to successfully time in the second half. We further show that investors who time the market during the run-up and crash around 2000 are more likely to time the run-up and crash around 2008. Our evidence suggests that it is possible to use the trading patterns of these skilled investors to anticipate market movements, lending some credibility to the view that market bubbles are identifiable in real time.

*Keppo is at NUS Business School and the Risk Management Institute, Shumway is at the University of Michigan and Weagley is at the Georgia Institute of Technology. They can be reached at keppo@nus.edu.sg, shumway@umich.edu and daniel.weagley@scheller.gatech.edu. We thank Noah Stoffman for helping us prepare the data. We thank Justin Birru, Stephen Dimmock, Madhu Kalimipalli, Juhani Linnainmaa, Yuri Tserlukevich and seminar participants at the University of Colorado, the University of Massachusetts, the University of Michigan, National University of Singapore, the 2016 American Economic Association Meetings, the Helsinki Finance Summit on Investor Behavior, the Ben Graham Centre's Symposium on Intelligent Investing and the Finance Research Workshop, IIM Calcutta for comments. Shumway is grateful for the financial support provided by National Institute on Aging grant 2-P01-AG026571. Keppo thankfully acknowledges financial support from the Singapore Ministry of Education under grant R-314-000-097-133. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Singapore Ministry of Education.

Introduction

We have learned a great deal about the behavior of individual investors over the past decade. We know they make lots of mistakes: trading too much, holding onto losers too long, buying stocks that appear in the news, trading with stale limit orders, and many other suboptimal behaviors.[1] However, we also know they learn with experience, that some individual investors can consistently pick stocks that outperform the market even after adjusting for risk, and that smarter investors have better performance.[2] In examining individual investor performance, the literature has focused on "stock picking" ability, while "market timing" has been largely ignored.[3] This is surprising given that the ability to time the market can have both large effects on portfolio performance and important implications for the predictability of market returns. In this paper, we conduct a thorough examination of individual investors' market timing ability. We test whether any individuals are able to consistently time aggregate stock market movements, focusing on recent market crashes.

In both 2000 and 2008, stock markets declined substantially after experiencing several years of unusually high returns. Researchers debate whether these price movements are asset price bubbles.[4] A point of contention is whether or not the experienced crashes were predictable ex-ante. We look for evidence of performance persistence in the ability of individual investors to time the stock market in general and bubble periods in particular around these two crashes. Our null hypothesis is that investors are not able to persistently time market movements. However, using Finnish data covering a fourteen and a half year period from 1995 to mid-2009, we find that

[1] See Barber and Odean (2000), Odean (1998), Barber and Odean (2008), and Linnainmaa (2010).

[2] See, for example, Seru, Shumway and Stoffman (2010), Coval, Hirshleifer and Shumway (2012), Che Norli and Priestly (2009), and Grinblatt, Keloharju and Linnainmaa (2012).

[3] A number of recent papers test whether individuals as a group have any ability to forecast prices, finding evidence both for and against smart individual investors. See, for example, Dorn, Huberman and Sengmueller (2008), Hvidkjaer (2008), Kaniel, Saar and Titman (2008), Barber, Odean and Zhu (2009), Kaniel, Liu, Saar and Titman (2011), and Kelly and Tetlock (2013).

[4] See, for example, Brunnermeier and Nagel (2004), Hong and Stein (2007), Zeira (1999), and Pastor and Veronesi (2003). More generally (and in concurrent work), Goetzmann (2015) and Greenwood, Shleifer and You (2017) study the predictability of future returns after significant price run-ups in national markets and industries, respectively.

active investors who time the market successfully in the first half of the sample are particularly likely to time successfully in the second half. We also find that individuals who successfully sell before the 2000 crash are more likely than others to sell before the 2008 crash, and that those with particularly poor timing around the 2000 crash also time the 2008 crash poorly. These results provide evidence that the significant market declines were partially anticipated.

One reason to study whether individuals can time the market is to determine how much predictability there is in market returns. There is a large literature about market predictability (e.g. Goyal and Welch (2008)), but researchers still disagree about whether or not returns are truly predictable. If some individuals time the market consistently, there is clearly predictability in market returns. If all individuals fail to time the market consistently, it may be that the market is unpredictable, or it may be that the time-varying risk involved in timing strategies is a sufficient deterrent to individual investors. Whether there is substantial predictability during periods that appear to be asset price bubbles is of particular interest. If some individual investors can successfully time bubbles then recognizing and exploiting bubbles in real time cannot be extremely difficult.

Knowing whether some investors can time the market can also help us learn about whether the variation in investor ability documented in other studies is due to differences in skill or information. Successful market timing means taking more market risk when the market is going to rise. Timing is a very different investing strategy than stock picking, which means taking more stock-specific risk in stocks that are going to outperform the market. Consistently good stock picking performance might indicate investing skill, but it may also indicate access to superior information like local news, friends in the industry, or personal expertise. Successfully timing the market requires either skill or private information to predict the future of the macroeconomy. Since almost all information relevant to the macroeconomy is publicly available news, it is difficult to believe that a subset of investors has superior access to it. Showing that some investors are successful market timers implies that skill in interpreting public signals is an important determinant of active investing success.

Our approach is to measure timing performance in the first and second halves of our sample, and then to test whether timing performance has persistence, or whether performance in the first half of the sample predicts performance in the second half. One important challenge in this exercise is identifying good measures of timing performance. This challenge is particularly daunting for us since we do not observe investors' entire wealth. To deal with this challenge, we measure performance by correlating beta-weighted monthly flows into and out of investor stock portfolios with future monthly returns on a Finnish stock index. Investors with flows that predict future market returns are successful market timers in this paper. Those with flows that are inversely related to future market returns are poor timers. Since correlations are restricted to vary from -1 to 1, active traders are essentially equally-weighted in our analysis.

Timing the market is known to be difficult and prior evidence suggests it is probably rare. The efficient market hypothesis holds that timing will only be possible in the presence of significant time-varying risk. This motivates our null hypothesis that there is no persistent market timing among any individual investors. Of course, since not all trades are driven by timing considerations, the correlation-based timing measures we use contain some noise. Furthermore, since we cannot accurately measure the timing ability of investors that trade infrequently, we study a relatively small set of active Finnish investors. Given the noise, sample size and the difficulty of timing, it is not realistic to expect very strong evidence of timing in the data. Rather, any evidence that a subset of investors can consistently time the market is surprising. While our correlation measures are imperfect, there is no reason to believe that the noise in flows should create persistence in timing. Unless the noise in flows somehow systematically forecasts market returns, it should simply weaken our tests.

We find both economically and statistically significant persistence in market timing ability. We document timing performance persistence both during bubble periods and during periods of normal returns. For our main timing measure, successful timers in the first half are over 40% more likely to be in the top quintile than the bottom quintile in the second half. The persistence we document is displayed broadly across different portions of the distribution of timing performance. It is also displayed by both successful timers and unsuccessful timers. Unsuccessful timers in the first half of the sample are quite likely to be unsuccessful in the second half.

We compare the observed persistence in market timing to the persistence in stock picking ability in our sample. We find that market timing exhibits greater persistence than stock picking. This is surprising considering stock picking ability can be driven by heterogeneity in skill and access to information, while differences in market timing ability are likely determined solely by skill (it is unlikely investors have differential access to macro information). Additionally, we find little evidence that stock picking and market timing skills are correlated when measured during the same time period. Market timing and stock picking appear to be distinct skills.

We next examine if the observed performance persistence is explained by investors using a simple strategy to time the market. We report the results of two sets of analyses. First, we sort investors into quintiles based on their first-half (1995-2002) timing performance and examine the correlation between quintile group flows and other market return predictors (dividend yield, earnings-price ratio and concurrent market returns) in the second half of our sample (2002-2009). In general, the active investor flows we study appear to be contrarian. The flows of good (top 20%) and bad (bottom 20%) timers are both positively correlated with valuation ratios and negatively correlated with concurrent market returns. We also examine the difference between the top and bottom timers' flows to capture the variation in flows due to timing ability. This difference is positively correlated with concurrent market returns and the earnings-price ratio, and negatively correlated with the dividend yield.

Second, we report the results of monthly market return predictability regressions which test if information in investor flows can significantly predict future market returns. We sort investors into quintiles based on their first period performance and examine return predictability in the second-half of our sample. We find that the standardized difference in group flows between the top and bottom 20% of investors can significantly predict future market returns. The success of the flow-based measure stands in stark contrast to the prediction power of the other variables. Only past market returns can significantly predict future market returns at the 10% level during this time period, while the valuation ratios are negatively and insignificantly related to future market returns. The magnitude and significance of the flow-based measure's coefficient is unaffected by controlling for the dividend yield, earnings-price ratio and past market returns. These results suggest that investor flows have more information about future market returns than commonly used economic variables.

We further assess the economic significance of the observed market timing persistence by sorting investors into quintiles based on their first-half timing measure and examining performance across quintiles in the second-half of our sample period. First, we examine the average return of an investment strategy that invests the standardized group flow measure in each month t and earns the market return over the month t+1. The strategy based on the flow difference between the top 20% and bottom 20% earns an average annualized return of 17.19% compared to an average market return of 1.02% during this period. To adjust for total risk, we take the ratio of the strategy's average return to its standard deviation. The flow-based strategy has an annualized reward-to-risk ratio of 0.81 compared to 0.05 for the market. We also examine the ability of investor flows to predict bear markets defined as a return at least half of one standard deviation below the sample average. The unconditional probability of a bear market is 24.1% in the second half of our sample. The conditional probability of a bear market increases to 60.0% when the flow differential between good and bad timers (sorted based on first half performance) is at least 0.5 standard deviations below the mean. These results are not necessarily surprising given the ability of group flows to forecast future returns. The persistent dispersion in market timing ability appears to be economically significant.

We examine what types of investors are better at market timing. We find that middle-aged investors (from 46 to 64 years old) are better timers. Males and those living in dense zip codes are marginally better timers. However, investors that live in highly-educated zip codes, those that trade more and those with greater average flow size are more likely to be worse timers. Consistent with investors capitalizing on their skill, we find that successful market timers invest in a greater number of different securities and securities with higher betas, and they are less likely to trade in Nokia (the most traded stock during the sample period).

Shiller (2000) and Campbell and Viceira (2002) argue that individuals should be able to time the market. Dichev (2007) examines aggregate flows and finds that the dollar-weighted returns to investors are lower than buy-and-hold returns indicating that investors on average are poor market timers. We find supporting evidence that the average individual investor cannot time the market. Our main goal is to examine the cross-section of timing ability, not the average investor's ability. Most analysis of the cross-section of market timing ability has examined returns on professionally managed funds. The overwhelming majority of these studies find little evidence of market timing.[5] However, Bollen and Busse (2001) and Mamaysky, Spiegel, and Zhang (2008) find some timing by professionals. Using holdings data Elton, Gruber, and Blake (2011) find evidence that fund managers' timing attempts usually result in low returns. Kacperczyk, Van Nieuwerburgh, and Veldkamp (2013), however, examine the holdings of fund managers and find that managers have some ability to time the market, especially in recessions. In contemporaneous work, Che, Norli and Priestly (2012) use detailed Norwegian data covering all of the domestic asset holdings of their investors to show that more individuals successfully time the market than would rise by pure chance. Compared to our work, they do not look for persistence in timing ability, they do not have data over two bubble periods, and they do not analyze return predictability.

The paper proceeds as follows. In the next section we describe our data in detail and we discuss our timing performance measures. In Section 2 we present and discuss the results of our

[5]See Jagannathan and Korajczyk (2014) for a review of the returns-based measurement of market timing literature. Studies that find little evidence of timing include Henriksson (1984), Ferson and Schadt (1996), Becker, Ferson, Myers, and Schill (1999), and Goetzmann, Ingersoll, and Ivković (2000). Kacperczyk and Seru (2007) find little evidence using returns and holdings data. Daniel, Grinblatt, Titman, and Wermers (1997) using returns, and Wermers (2000) using holdings and returns, find little evidence of characteristic timing ability.

tests, and in Section 3 we conclude.

1 Data

Our main data set combines data on individual investor transactions with data on market returns. The original transaction data contains all transactions in Finnish stocks during the sample period and comes from the Nordic Central Securities Depository (NCSD). We extend the datasets used in Seru, Shumway and Stoffman (2010), Grinblatt and Keloharju (2000, 2001a, 2001b), and Kaustia and Knüpfer (2012) to cover 14.5 years of trading from January 1, 1995 to June 30, 2009. For each transaction, we are provided the number of shares transacted, the transaction price, a security identifier, an investor identifier code, and information about the investor.[6] We focus on the subset of transactions by individual investors. Individual transactions are aggregated at the individual, not account, level. This level of aggregation eases concerns that investors are not actively moving money into and out of the market, but are moving money between accounts. The dataset only provides information on investor's direct stock holdings. Investments through an intermediary are attributed to the intermediary's account. Thus, mutual funds will have their own accounts and will not be included in our analysis. Grinblatt and Keloharju (2000) find that less than 1% of the Finnish population were invested in mutual funds in the beginning of 1997. Although this proportion is likely to have grown, there is no obvious reason why excluding flows to mutual funds would affect our results.

We limit our analysis to active individual investors and, for most of our analyses, we drop transactions made by institutions. We limit our sample to individual investors for two reasons: (1) individuals are not regulated or restricted in their investment set and (2) individuals are likely investing for themselves, which eliminates any agency concerns. Further, individual investors are traditionally thought of as the least informed or skilled investor group, so finding any timing

[6]The data set contains information on the investor's zip code, gender, firm sector code, firm legal form, firm postal code, firm country, language code, and registration date in shareholder register. The data set also contains information on the type of transaction and the transaction registration basis.

ability among this sub-group of investors is particularly surprising. We use the trading records of institutions to look for evidence of timing in one test, but the test results in no evidence of timing. Given the relatively small number of institutions in the data, we do not find the lack of evidence for institutions surprising.

The trading records contain codes indicating what type of transaction corresponds to each record. About sixty percent of the records are normal transactions, or trades which individuals executed on the exchange. Other records are generated when firms merge, when they go through a stock split, or when other corporate events occur. After examining the data carefully, we choose to exclude some types of trades but we keep most of the transactions. The transactions we exclude comprise about 12 percent of the data and often have values of zero.[7] We choose to retain records associated with mergers and other corporate events because they generate trades in individuals' accounts which affect their market exposures. For robustness, we ran our analysis without these events and our results remain similar.

Since our interest lies in investors' ability to time the market, we aggregate investor flows at the investor-month level instead of analyzing trades in individual stocks. To calculate flows, we sum the euro value of all transactions by an individual within each month. With respect to nonstandard transactions, it is important for us to keep most types of transactions in our set of trades to calculate accurate flows. While it is true that an individual may not actively choose to sell a stock when a company is being taken over or liquidated, it is also true that those transactions generate a cash flow in the individual's account. Since most of our analysis considers the trades of fairly active traders, we assume that individuals are aware of such transactions, and thus it would be a mistake for us to omit them. Suppose, for example, that a company is taken over and 10,000 euros are deposited into an individual's account. If that individual then spends 5,000 euros to purchase another stock, then the correct flow for that account in that month is -5,000 euros. If we omit the takeover transaction then it appears that this individual's flow is

[7] We exclude transactions with codes of 0 (initial balance), 4 (transfer to own book-entry account), 55 (splits), 60 (creation of subscription rights), and 96 (transfer to own book-entry account to another operator).

5,000 euros this month, flipping the sign of the true flow.

The data also contain short descriptions of all the securities traded on the exchange. We use these codes to identify put options. Since put option returns are usually of the opposite sign of their underlying assets, they almost always have negative betas. Thus, when we calculate investor flows we give put option flows a negative sign. If investors sell puts we count the associated flows as positive and if they buy puts we count the flows as negative.

All of our tests of timing persistence will examine active traders, traders that have a minimum number of months with non-zero aggregate flows (active flows) in the first half of our sample. We focus on active traders to ensure we have an accurate measure of timing. The threshold to be included in our analysis depends on the market timing measure of interest. For the entire period timing measure, an investor must have active flows in at least 15 of the 87 months.[8] For the bubble period timing measure, an investor must have active flows in at least 8 months during the first bubble period. We use first half activity to determine if investors are included so there is no look ahead bias in our results. In Figure 3, we provide a histogram of active (non-zero) monthly flows in the first and second half of the sample period for the 1,386,540 investors that owned a security during the first half of our sample period.[9] The main takeaway is that there is very little activity by the typical investor. The median investor has 2 active flows (months with non-zero aggregate flows) and less than 10% of investors have greater than 10 active flows (out of 87 months) in each half of our sample. The active investor thresholds balance sample size and measure accuracy.

Summary statistics on trades and flows by active investors (i.e., investors with at least 15 non-zero flows in the first half of the sample period) are presented in Table 1. There are 68,937 investors that meet the threshold to be an active investor. We treat an individual's first trade in our sample as their first trade in the market. Once an investor makes their first trade, they

[8] Our results do not depend on the exact cut-off value chosen (e.g. with a cut-off of 12 flows, the correlation between first and second period monthly timing measures was 0.0714 and is significant at the 0.01% level). The cut-off value was chosen to optimize the trade-off between sample size and capturing active investors.

[9] 1,386,540 people is approximately 27% of the Finnish population in the year 2002 (see `http://www.stat.fi`)

remain in our sample and receive a monthly flow equal to zero in all months they do not trade. As can be seen in Column 3 of Table 1, the number of flows increases each year as new investors enter the market.[10] As a robustness check we also conduct our main tests assuming that all investors begin trading before the sample begins, and we find very similar results.

In Columns 6-8, we provide the percentage of monthly investor flows each year that are net outflows, net zero, and net inflows, respectively. An investor is assigned a monthly flow of zero if the investor was inactive or if the investor bought and sold exactly the same value of securities during the month. The overwhelming majority of monthly flows are equal to zero. The year 2000 had the lowest percentage of flows equal to zero at 52.52%, and the year 1995 had the highest percentage at 82.50%. The percentage of flows that are outflows hits a low of 5.08% in 1995 and a high of 19.91% in 2000. The percentage of flows that are inflows ranges between 8.30% in 1996 and 27.57% in 1999. In 10 out of the 15 years, a greater percentage of investor flows were inflows than outflows. This does not necessarily mean that the average flow size in euros was positive in these years. In Column 4, we present the mean flow size, which is negative in 6 out of 15 years.

To proxy for the relevant market return, we use returns on the HEX 25 Index (currently, the OMX Helsinki 25), which we obtain from Bloomberg. The HEX 25 is a value-weighted index of the 25 largest companies listed on the Helsinki Stock Exchange.[11] The cumulative return of the HEX 25 index over our sample period appears in Figure 1 and monthly returns are presented in Figure 2. As the figures clearly show, the sample period can be characterized by two episodes in which there was a market run-up and subsequent crash in prices. This remarkable pattern facilitates our test of timing around market bubbles.

In some tests, we use returns on individual securities to calculate active changes in portfolio betas, stock picking ability and to control for any effects due to the dominance of Nokia (the most traded stock in our sample) during our sample period. For these analyses we need the time

[10] The number of flows in 2009 is smaller than 2008 because our sample ends in June 2009.

[11] From November 1, 1995 to August 1, 2001 the index capped the weight of any individual stock at 20%. After August 1, 2001 the index caps the weight of any individual stock at 10%. The number of stocks capped varies over time. As an example, on November 3, 2003 there were 4 stocks capped at 10%. In the online appendix, we present results using a more general market index (HEX) and find that the results are very similar.

series of individual securities prices, which we obtain from Bloomberg for the 1,000 most traded securities in our sample.[12]

1.1 Timing Measures

A number of authors have considered the best way to measure market timing (see Jagannathan and Korajczyk (2014) for a review of the returns-based measurement literature). Most of the methods proposed by these authors involve explicit market forecasts or time-varying portfolio betas (e.g., Treynor and Mazuy (1966), Merton (1981), and Henricksson and Merton (1981)), and most use data generated by professional asset managers. In principle investors might time the market by changing the betas of their portfolios or by simply placing money in the market before relatively high returns and withdrawing money from the market before relatively low returns.

We calculate two main flow-based timing measures. The first timing measure is calculated for each investor as follows:

$$MarketTiming_i = Correlation(Flow_{it}, MonthReturn_{t+1}), \tag{1}$$

where $Flow_{it}$ is the monthly flow for investor i in month t and $MonthReturn_{t+1}$ is the cash return on the HEX 25 in month $t+1$.[13] We use cash returns so that both the flows and returns are in euros. We compare investors' timing measures in two equal length sub-periods: January 1995 to March 2002 and April 2002 to June 2009. There are 87 months in each sub-period.

[12]There are over 8,285 securities in our sample, of which 155 are stocks (identified by their existence in the Compustat Global database). Stocks account for 63.23% of the trades and 73.95% of the absolute flows. The rest of the securities are derivative instruments, bonds, and ETFs. Only a fraction of the derivatives are traded in any given period since derivatives with different expiration dates have different identifiers. As a robustness check, we run our main analysis using only equities and, as expected, the results are similar to those with all securities. The rank correlation between the first and second half monthly timing measures is 0.0285, which is significant at the 0.01% level with a p-value of 0.0000. The reason we include all the securities in our main analysis is that investors could use derivative securities or corporate bonds in their market timing strategy. In sum, the results are not sensitive to the subset of securities used.

[13]For robustness, we have re-run our main analysis using quarterly returns. This measure correlates monthly flows with the cash return on the HEX 25 over the 3 months beginning in month $t+1$ and ending in month $t+3$. These results are presented in an online appendix and are very similar to the results using the monthly market timing measure.

Using the correlation between market flows and future returns makes particular sense for our data for multiple reasons. First, the correlation measures directly if investors are able to time the market. Second, while the best possible measure of timing we might construct would utilize the fraction of each person's wealth that they allocate to risky assets, we do not observe the total wealth of the individuals in our data. Calculating the correlations of an individual's flows with future market returns essentially adjusts each individual's monthly flows by the standard deviation of their flows, which is a proxy for their total wealth. Third, while we observe all the individual stock transactions made by each individual, we do not observe their transactions in mutual funds. When individuals sell stocks, they are likely to either leave the proceeds in their investment accounts or to invest some of the proceeds in other assets. To the extent that individuals sell stocks and then use the proceeds to buy stock mutual funds, our timing measure is imperfect. However, as mentioned above, mutual fund investments are a small percentage of investor portfolios in Finland (Grinblatt and Keloharju, 2000) and their absence from our data is unlikely to affect our results.

For the second measure, we correlate a beta-adjusted flow measure with future market returns. The beta-adjusted flow is the euro value of each transaction multiplied by the security's beta (euro-value of trade of security i times beta of security i). For example, if an investor sells 100 euros of stock Y with beta equal to 1 and purchases 100 euros of stock Z with beta equal to 2, we calculate a beta-adjusted flow of $100 * 2 - 100 * 1 = 100$ beta-adjusted euros. Specifically, the measure is calculated:

$$MonthlyTimingBeta_i = Correlation\left(\sum_{j=1}^{J} \beta_j * Flow_{ijt}, MonthReturn_{t+1}\right), \qquad (2)$$

where $Flow_{ijt}$ is the monthly flow for investor i into security j in month t, β_j is the beta of stock j, J is the total number of securities in the market, and $MonthReturn_{t+1}$ is the cash return on the HEX 25 in month t+1. We calculate betas using in-sample weekly returns over a minimum of 52 weeks. We drop securities for which we are unable to calculate a beta. Our results are similar if we use the beta-adjusted or unadjusted flow measures. For the majority of the paper we use the unadjusted flow measure.

We calculate two additional timing measures that capture an investor's ability to time market bubbles. The main bubble timing measure calculates a flow-return correlation (equation (1)) only during the two bubble periods in our sample. Bubble periods are defined around the two market peaks. The peak months on the HEX 25 were March 2000 and July 2007. We treat these two months as the beginning of a market crash and calculate whether an investor performed well in the 12 months before and 12 months after the market peak. Therefore, we have 25 months of data for each bubble period.[14] The second bubble timing measure calculates the significant outflows of each investor around the market peak month. Specifically, we compare average flows during the six months ending in the market peak month to the average flows over the entire sample half. We define the measure more carefully in Section 2.

In Table 2, we present summary statistics for the entire period and bubble period timing measures for each half of the sample. In the rows labeled Entire Period, we present summary statistics for the entire period timing measure (equation (1)). There are 68,937 investors that meet the minimum number of active flows. In the second period, 2,097 investors fail to have at least two non-zero flows and thus drop out of the sample because we are unable to calculate a correlation for them. The mean of the entire period timing measure is 0.03 in the first half and 0.00 in the second half of the sample. In unreported results, we find that the mean monthly timing measure for all investors is -0.01 in each half of the sample. Frequent traders are on average better monthly timers than the entire population and these differences are statistically significant at the 1% level. A correlation of -0.01 for all investors is evidence that investors cannot time the market on average, consistent with Dichev (2007). There is significant variation in the timing measure especially in the first half of the sample. In the first half, the standard

[14] We have also examined timing in the months outside of the bubble periods, which we label "Normal Times." Normal times are defined as the 62 months in the sample half of interest that are outside of the 25-month bubble periods. We find that investors that time in normal times are more likely to time market bubbles. Results are presented in the online appendix.

deviation is 0.18, 25% of investors have a correlation less than -0.07 and 25% of investors have a correlation greater than 0.13. In the second half, the standard deviation is only 0.10 and the 25th and 75th percentiles are -0.06 and 0.07, respectively.

In the rows labeled Bubble Period, we present summary statistics for the main bubble period timing measure (equation (1)). There are 68,717 frequent traders in the first period and 51,061 in the second period. The attrition is most likely due to the short time period in which we measure bubble period performance. The mean is 0.03 in the first period and 0.00 in the second period. This is evidence that frequent traders timed the 2000 market bubble better than the 2007 bubble. The standard deviation of the bubble timing measure is larger in the first period than the second period, 0.25 versus 0.19. 25% of investors had a correlation greater than 0.19 in the first period, while the corresponding value in the second period was only 0.13. In our persistence tests, we will examine if investors that were in the top (bottom) percentiles in the first period were more likely to be in the top (bottom) percentiles in the second period.

2 Results

In Table 3, we present the results using the beta-adjusted timing measure defined in equation (2). The table is a simple cross tabulation of the first and second half timing measures. The rows of the table are sorted into quintiles based on performance in the first half, while the columns are sorted into quintiles based on performance in the second half. We present row percentages, i.e. percentages conditional on being in the relevant first half quintile. Under the null hypothesis that there is no relation between timing performance in the first and second halves of the sample we would expect to see twenty percent of the observations in each cell. The indications of statistical significance in the table are for tests of the null hypothesis.[15]

The results of Table 3 clearly show that there is some timing ability in our data. Focusing on

[15] p-values are calculated using a bootstrap procedure that clusters investors based on their geographical location and trading frequency. We discuss the boostrapping proceduce later in this section.

the top row, which corresponds to the best performers in the first half, the fraction of investors that appear in each performance quintile in the second half of the sample declines monotonically from 24.88% to 17.26%. These results indicate that the best first period timers are 44% more likely to be in the top 20% than in the bottom 20% in the second half. Looking at the last row, the fraction in each cell increases monotonically. Many of the extreme quintile pairs in these two rows are statistically significantly different from the null value of twenty percent. Looking at the first column of the table, again the fractions in each cell decline near monotonically. In the last column the fractions increase monotonically. A rank correlation between first and second half performance of 0.0692, which is significant at the 0.01% level, provides further evidence for persistent market timing ability among the investors in our sample. These results show that the best timers in the second half come disproportionately from the better quintiles in the first half, and the worst timers in the second half come disproportionately from the worst quintiles in the first half.

Table 4 reports our second set of results on timing persistence. The timing measure is the unadjusted flow-based measure from equation (1). The results are similar to the beta-adjusted timing measure results in Table 3. 24.43% of the investors in the top 20% in the first half are in the top 20% in the second half. Only 17.29% of the top 20% of investors in the first half are in the bottom 20% in the second half. Similarly, the investors in the bottom quintile in the first half are much more likely to be in the bottom quintiles in the second half. The percentages in Table 4 translate into a substantial number of investors. For example, if investors were distributed uniformly across all the cells in the table we would expect to see 2,674 investors in each cell. In the first row and column there are actually 3,266 investors, 592 more than we would expect by chance. The rank correlation between timing ability in the two halves is 0.0745 and is significant at the one-thousandths of a percent level. Because of the similarity in results between the beta-adjusted flow and unadjusted flow measures, we use the simple unadjusted flow measure for the remainder of the paper.

In the last column of Table 4, we report the average second period timing measure for each first period quintile. For the highest quintile, the average timing measure in the second half is 1.56% and is significant at the 1% level. For the lowest quintile, the average correlation is only -0.63%. The difference between the highest and lowest quintile is 2.21% and is significant at the 1% level. The results of the table imply that the correlation is not driven by the tails of the distribution, and it is not driven primarily by either very unsuccessful or very successful timers. Rather, there is a considerable amount of persistence in good and bad timing ability across the entire distribution. In subsection 2.4, we calculate the correlation between first half group level aggregate flows and future returns and find even more economically significant dispersion in performance in the second half.

If investors make correlated investment decisions due to factors beyond individual market timing ability, like regional shocks to wealth or a common financial advisor, then simple tests of statistical significance may be misspecified. To account for any influence of such clustering of individuals, we calculate p-values in the main tests of market timing using a bootstrap procedure. We first sort individuals into one of nine geographical regions based on their zip codes. Then within each region we sort individuals into terciles based on trading frequency, generating a total of 27 region/frequency groups. We match each investor with a randomly chosen (with replacement) investor from the same region/frequency group 1,000 times. Finally, we use the distribution of the matched samples to determine the significance level of each cell in the cross-tabulation. We find that our bootstrapped significance tests are extremely close to the simple p-values calculated using OLS methods. This gives us confidence that the standard errors are correct in some of our robustness analyses which do not use a bootstrap procedure.

2.1 Bubble Analysis

In Table 5, we report the results of persistence tests using our bubble period market timing measure. This measure assesses investors' ability to time the market bubbles of 2000 and 2007.

Since the bubble timing measures are based on just 25 months of trades right around the market crash, they are likely to be weaker than the results reported in Table 4. Looking at the first row, the percentages in the cells decrease nearly monotonically from 21.96% to 19.39%. Looking at the final row of the table, except for the bottom 80-100% column, the percentages are monotonically increasing. The rank correlation between the 2000 and 2007 bubble timing measures is 0.021, significantly lower than the correlations of our full period measure, but still statistically different from zero at the thousandths level of confidence. Overall, there appears to be significant persistence in investors ability to time bubbles over monthly horizons.

In Table 6, we report the persistence tests for significant outflows around the market peaks of 2000 and 2007. Our significant outflow measure is the difference between average flow over the six months ending with the market peak month and the average flow over the entire sample half divided by the standard deviation of flow. Thus, the measure is calculated as follows: $-\frac{\frac{1}{6}\sum_{m=PeakMonth-5}^{PeakMonth}[flow_{im}]-\overline{flow_i}}{s_i}$, where $flow_{im}$ is the flow of investor i in month m, $\overline{flow_i}$ is the average flow for investor i over the sample half in which the bubble occurs, s_i is the standard deviation of flows for investor i over the sample half in which the bubble occurs, $PeakMonth$ is the month the market reached its apex during the sample half. We use the negative of the standardized difference, so that investors with larger outflows are considered better timers.

The persistence pattern is similar to the pattern of the monthly flow measure in Table 5, although it is weaker. The weaker persistence is likely due to noise because skill is estimated using fewer data points. Investors that were in the top quintile of timers of the 2000 peak were more likely to be in the top quintile than in the bottom quintile of timers of the 2007 peak. Similarly, investors that were in the bottom quintile in 2000 were significantly more likely to be in the bottom quintile in 2007. The rank correlation is 0.015 and is significant at the 0.01% level. In an unreported test, we calculate the standardized flow measure over the 11 months centered at the monthly peak. The results are similar. Investors that were more likely to have significant outflows around the market peak in 2000 were more likely to have significant (out)flows around the market peak in 2007.

2.2 Market Timing versus Stock Picking

There have been many studies that document persistence in stock picking ability across individual investors (e.g., Seru, Shumway and Stoffman (2010), Coval, Hirshleifer and Shumway (2012), Che Norli and Priestly (2009), and Grinblatt, Keloharju and Linnainmaa (2012)). How does the observed market timing persistence compare to the stock picking persistence in our sample? To answer this question, we create a measure for an investor's stock selection ability following Seru, Shumway, and Stoffman (2009). For each investor, we calculate the average 30-day return on their stock purchases minus the market return. If the investor sells the security before 30 days, we use the return during the holding period. We calculate a performance measure for each half and present the cross-tabulation in Table 7. Confirming previous findings, we find significant persistence in stock picking ability. The spearman rank correlation is 0.024 and is significant at the 0.01% level. Surprisingly, the persistence in the main market timing measure is as strong or stronger than the stock picking persistence. These results show that market timing persistence is likely as economically significant as stock picking persistence. This is especially surprising considering market timing persistence likely comes from heterogeneity in skill, whereas stock picking persistence can be driven by heterogeneity in access to information and skill.

Are market timers better stock pickers? We calculate the stock selection measure and the monthly market timing measure over the entire sample period (14.5 years) for investors with at least 75 trades and 15 non-zero monthly flows in the first half of the sample. We find a Spearman rank correlation between the two measures of -0.0124 (p-value of 0.05) and a pairwise correlation of 0.0014 (p-value of 0.82). There is little evidence that good market timers are more likely to be good stock pickers when ability is calculated over a long time span. The lack of a positive correlation between skills could be due to noise in our measure of stock picking ability or due to investors specializing in one of the two skills. Kacperczyk, Van Nieuwerburgh, and Veldkamp

(2013) provide evidence that skilled fund managers will focus on one of the two skills conditional on the business cycle. They find that fund managers that are good stock pickers in expansions are more likely to be good timers in recessions. In the online appendix, we provide evidence that investors that are better timers in the bubble periods are also better timers during normal times. We do not measure stock picking or market timing ability during different periods of the business cycle, however. Instead, we show that, unconditionally, stock picking and market timing are relatively uncorrelated.

2.3 Flow Analysis and Return Predictability

In this section, we explore what drives the persistence in market timing. The observed persistence in market timing ability could be due to investors trading on many different types of information: valuation ratios, past market returns, news flow, etc. If we find that a simple strategy explains the observed persistence (e.g., asset allocation is based on past market returns), then our paper is a noisy test of return predictability using this simple strategy. If we find little evidence supporting a simple strategy, this indicates that our market timing measure captures individual investor skill, and there is additional information in investor flows than what can be gleaned from simple predictive regressions using economic variables.

Our first analysis compares investor flows to three predictors of market returns: earnings-price ratio, dividend yield, and the concurrent market return. To do this, we sort investors based on their first period monthly market timing measure into quintiles and create a monthly quintile group flow measure for each month in the second half of the sample. Our group flow is defined as follows:

$$AggFlow_{gt} = \sum_{i \in g} \frac{Flow_{it} - \overline{Flow_{it(H2)}}}{\sigma(Flow_{it(H1)})},$$

$$GroupFlow_{gt} = \frac{AggFlow_{gt} - \overline{AggFlow_{gt(H2)}}}{\sigma(AggFlow_{gt(H2)})}, \tag{3}$$

where the subscript g indexes quintile groups, $(H1)$ and $(H2)$ indicate sample averages or standard deviations taken over the first or second half of the sample, respectively.[16] By the first normalization above, the normalized flows of different investors are weighted equally, and the second normalization guarantees that each group flow has a mean of zero and a standard deviation of one, allowing for comparison across groups.

Additionally, we create another standardized flow measure from the difference in $AggFlow_{gt}$ for the top 20% of timers and the bottom 20% of timers based on first half performance. The measure is calculated as follows:

$$StdTopBottomFlow_t = \frac{TopBottomFlow_t - \overline{TopBottomFlow_{t(H2)}}}{\sigma(TopBottomFlow_{t(H2)})}, \tag{4}$$

where $TopBottomFlow_t = (AggFlow_{top,t} - AggFlow_{bottom,t})$, *top* and *bottom* are the groups of the top 20% and bottom 20% timers, $\overline{TopBottomFlow_{t(H2)}}$ is the mean of $TopBottomFlow_t$ during the second period, $\sigma(TopBottomFlow_{t(H2)})$ is the standard deviation of $TopBottomFlow_t$ flow during the second period. Like the group flow measure in equation (3), the Top-Bottom strategy in equation (4) has a mean of zero and a standard deviation of one.

In Table 8, we present correlations between the top 20% group flow, the bottom 20% group flow, the Top-Bottom strategy (Top 20-Bottom 20) and the three predictor variables (the earnings-price ratio, dividend yield and the lagged excess return of the HEX25). Examining Columns (2) and (3), we see that the group flows are contrarian. Both the top 20% and bottom 20% group flows are positively correlated with the valuation ratios (earnings-price ratio and the dividend yield) and negatively correlated with lagged market returns. The Top-Bottom strategy (presented in Column (1)), is significantly positively correlated with the earnings-price ratio and past market returns, and insignificantly negatively correlated with the dividend yield. These correlations are consistent with the "timing" portion of flows capturing momentum trading (correlation with past market returns), but also trading in a contrarian manner (correlation with the

[16] We winsorize the standardized flows of the investors at the 1% and 99% level to minimize the impact of outliers.

earnings-price ratio). The R^2 from regressing the top 20% group flow, bottom 20% group flow, and top 20% minus bottom 20% group flow on the three predictor variables is 18%, 37%, and 17%, respectively. Based on these results the persistence in timing that we capture cannot be fully explained by the combination of these simple strategies.

We next turn to return predictability regressions to examine the relationship between flows and future returns after controlling for other economic variables. If the group flow measure in equation (3) can improve predictions of future returns beyond the three economic variables (past returns, dividend yield and earnings-price ratio), then it is unlikely that a simple strategy can explain our results. We run regressions of the form:

$$HEX25_{t+1} - r_{f,t+1} = \beta_g \times GroupFlow_{gt} + \beta_x \times X_t + \alpha + \epsilon_t,$$

where $HEX25_{t+1}$ is the return on the HEX 25 over month $t+1$, $r_{f,t+1}$ is the one-month Euribor rate, X_t is the vector of excess returns on the HEX25 in month t, the logarithm of dividend yield and the logarithm of earnings-price ratio of the HEX 25 over month t, and $GroupFlow_{gt}$ is calculated for month t according to equation (3). We focus on the flows of the top 20% and bottom 20% groups and the Top-Bottom strategy.

Results are presented in Table 9. In Columns 1-3 we run univariate OLS regressions using the flow-based measures. The top 20% flow is positively related to future returns with a coefficient of 0.0133, which is insignificant. The bottom 20% flow has a negative coefficient of -0.0032 and is insignificant. The coefficients have the expected signs, but the coefficients and R^2 are small in magnitude. In Column 1, the independent variable is the Top-Bottom strategy, calculated according to equation (4). This measure can significantly predict future market returns at the 5% level. The coefficient is 0.0145. This implies that a one-standard deviation increase in the flow measure is associated with an increase in the market return of 1.45%. The R^2 is 5% which is much larger than any of the other univariate regressions. Thus, the Top-Bottom strategy has some ability to predict returns before controlling for the other economic variables.

In Columns 4-6, we regress the market return in month $t+1$ on the other predictor variables individually. The dividend yield and earnings-price ratio show little forecasting power. They both have negative and insignificant coefficients. We would expect a positive coefficient on these two valuation ratios. The coefficient on past market return is 0.200 and is significant at the 10% level. The R^2 is 4%, which is of reasonable magnitude. This indicates there is some auto-correlation in the HEX25 returns during this period. In Column 7, we include all the three economic variables. The R^2 is 5%, which is the same as the R^2 of the Top-Bottom strategy alone. In Column 8, we include the Top-Bottom strategy with the other economic variables. The R^2 almost doubles to 10.1%. The Top-Bottom strategy remains significant at the 5% level and the coefficient actually increases to 0.0160.[17] In the online appendix, we create a market timing measure that is orthogonal to a strategy based on the auto-correlation in monthly returns since monthly returns is the most significant of the other predictor variables. We find similar persistence in this test as in our main test. Thus, the difference in performance across good and bad timers is not captured by the three economic variables (past returns, dividend yield, and earnings-price ratio) and the results suggest that including information in individual investor flows improves the performance of predictability regressions.

The observed market timing persistence and return predictability is especially surprising given the different economic drivers behind the run up and crash in the 1995-2002 time period and the run up and crash in the 2002-2009 time period. The relative outperformance of investor flows to predict future returns could be due to investors dynamically adjusting their models to different economic environments and synthesizing many public signals beyond the three economic variables. The above results and the different economic environments across the two time periods ease concerns that performance persistence can be explained by investors following a simple strategy.

[17]The significance of the coefficient increases dramatically if Newey-West standard errors with a 1-month lag are used. We report the uncorrected standard errors to be conservative.

2.4 Economic Significance

There is clear persistence in timing ability and information in investor flows is correlated with future market returns. In this section, we examine further the economic significance of the observed timing persistence. We calculate the performance of strategies that mimic successful and unsuccessful timers (equation (3)) as well as the Top-Bottom strategy (equation (4)). We look at the performance for both the entire period and bubble period timing measure.

We calculate three performance metrics. First, we calculate the correlation between the quintile group flow in month t and the excess return on the HEX25 (minus 1-month Euribor) in month $t + 1$. Second, we calculate the average return to a flow-weighted return strategy. The strategy weights each month's return by the previous month's group flow. Specifically,

$$FlowWeightedReturn_{g,t+1} = GroupFlow_{gt} * ExcessReturn_{t+1}, \quad (5)$$

where $ExcessReturn_{t+1}$ is the excess return on the HEX 25 index in month $t + 1$, $GroupFlow_{gt}$ is the monthly group flow. The third measure accounts for the risk in the flow-weighted return strategy. We deflate the average return from the flow-weighted strategy by the standard deviation of the strategy. This is similar to a Sharpe ratio.

Results are presented in Table 10. Panel A of Table 10 shows the correlations between group flows and market excess returns. Consistent with previous results, successful timers in the first half of our sample are more likely to be successful in the second half of our sample. The top performing groups (Q1) have correlations of 0.10 and 0.08 for the entire and bubble period measures, respectively. The correlations decrease monotonically from the top to bottom quintiles. For the worst timers, the correlations are -0.03 for both timing measures. Note that the passive buy-and-hold strategy has a zero correlation since it is analogous to investing one euro into the index in the first month and then cumulating that investment in subsequent months. The correlations for the Top-Bottom strategy are 0.22 and 0.21, and are significantly different from zero. These results provide further evidence of the persistence in market timing.[18]

We report the average annualized flow-weighted returns along with the returns generated by a passive buy-and-hold strategy in Panel B. Consistent with our earlier results, past successful timers outperform unsuccessful timers by 10.18% per year using the entire period measure and by 8.37% using the bubble period measure. A t-test of the difference produces p-values of 0.02 for both measures. Past successful timers outperform a passive benchmark by 6.73% and 5.42% using the entire and bubble period measures, respectively. The differences are not statistically significant. The corresponding returns of the Top-Bottom strategy are 17.19% and 16.44%, i.e., about 17 times the passive return. The Top-Bottom strategy standardizes the difference in the group flows, hence the discrepancy in the return on the Top-Bottom strategy and the difference between the Q1 and Q5 strategy returns. The dispersion in returns are economically large and the monotonic decrease in performance from top to bottom quintiles is unlikely to be due to chance.

In Panel C of Table 10 we report the total risk-adjusted performance measure (i.e., Flow-Weighted Return-Volatility Ratio). The successful timers' ratios are 5 to 6 times the ratio of the passive buy-and-hold strategy. The ratios of the worst timers are -0.11 and -0.09 and the ratios of the Top-Bottom strategy are 0.81 for both measures. Again our results suggest that the market timing ability of our investors is economically large and important.

Our tests, so far, have examined a linear relationship between investor flows and future returns. We are also interested in whether investor flows can improve our predictions of "bear" markets - defined here as a return at least half of one standard deviation below the sample average. We calculate the increase in probability of a "bear" market the next period if there is a large (negative) difference between good and bad timers flows. We consider a difference in flows "large" if it is at least half of a standard deviation below its mean (labeled "Low Flow"). The

[18]The correlations between adjusted group flows and future returns are much larger (in absolute magnitude) than the average correlation of individual investor flows and returns within each group presented in Table 4. This is not surprising considering the aggregated flows effectively diversify away idiosyncratic drivers of flows and more efficiently capture the timing component. This is analogous to individual securities having lower average correlations with the market than a portfolio of securities.

discrepancy in flows is calculated in the same way as the Top-Bottom strategy in equation (4).

In Table 11, we calculate the probability of a "bear" market, the probability of a "Low Flow", the probability of a "Low Flow" in month t given a "bear" market in month $t+1$, and the probability of a "bear" market in month $t+1$ given a "Low Flow" in month t. Perhaps unsurprisingly, when good timers have lower flows than bad timers, this is a pretty good predictor of poor market returns. For both measures, a "Low Flow" gives a higher probability of negative market returns in the next month than the unconditional probability. For the entire period measure, the probability almost doubles. Because we only have 87 months over which to calculate the probabilities and very few months with significant outflows, these numbers should be considered only as suggestive evidence.

2.5 Investor Skills and Characteristics

In this subsection we analyze which types of investors are better at market timing. Further, we study if successful market timers are also successful in stock picking. We analyze investors along many dimensions: sex, age, education, other demographics, and trading behavior. We estimate three separate regressions where the dependent variable for investor skill is based on the investor's monthly timing measure calculated over the entire 14.5 year period of interest.

Regression results are reported in Table 12. In Column 1, the dependent variable is a dummy variable equal to 1 if an investor is in the top 20% of investors. In Column 2, the dependent variable is a dummy variable equal to one if the investor is in the bottom 20% of investors. In Column 3, the variable is equal to 5 if the investor is in the Top 20%, 4 if the investor is in the Top 20-40%, etc. The regressions in Columns 1 and 2 are probits. The regression in Column 3 is an ordinary least squares regression.

There is no simple explanation of which traders are good timers and which are poor timers. We find that men are more likely to be in the top of the distribution. Their coefficient in the overall timing skill regression is positive, but insignificant. Middle age investors (45-64 in

1995) are better timers than younger and older investors, and they are the most likely to be in the top of the distribution. Retirement-age (65+) investors are the least likely to be in the bottom of the distribution. Young investors are the worst timers. The age results are consistent with investors learning with experience as they age. After a certain point performance decays, perhaps as cognitive abilities decline. These results are consistent with studies on how individual productivity varies by age. Skirbekk (2004) surveys the literature on age and productivity and finds that workers experience productivity reductions after age 50 and the reductions are greatest in tasks where problem solving, learning, and speed are needed. Market timing likely requires all three of these qualities.

Examining the characteristics of the traders zip code, we see that investors from more dense (urban) zip codes are marginally better timers but their overall ability coefficient is not significantly different from zero. Surprisingly, living around highly educated people (% of individuals in a zip-code with a University degree) significantly reduces the likelihood of being in the top 20%. Investors that speak the Finnish language (instead of Swedish) are marginally worse timers.

We include an indicator variable for whether an individual ever trades an option (following Seru, Stoffman, and Shumway (2009)), expecting this to measure sophistication. We find that investors that trade options are actually worse timers than others. This may be explained by the fact that many of the options traded on the Helsinki Exchange are given to corporate executives as compensation. Expectations of future aggregate market returns are unlikely to be an important driver of executives' flows as compared to firm-specific private information.[19]

We also find that timers trade less. This could be due to macro information arriving at a slower frequency than firm-level information or, possibly, good market timers are less overconfident (Barber and Odean (2001)). Investors with larger absolute flows (a proxy for wealth) are less likely to be in the top 20%, but are also less likely to be in the bottom 20% of investors. Flow size is negatively correlated with timing ability in general. It is not obvious whether or not wealthier

[19]For example, Yermack (1997) shows that options are awarded before the revelation of positive firm announcements.

investors should be better or worse market timers. They have a greater monetary incentive to gather information, but they may have larger opportunity costs. Additionally, wealthier investors are unlikely to have better information about the aggregate market than other investors.

We hypothesize that good timers are more likely to diversify their holdings allowing them to take advantage of their timing ability while minimizing idiosyncratic risk. We include two dummy variables based on the number of different securities traded by the investor during the sample period, one variable equal to one if the investor traded between 26 and 40 securities and the other variable equal to one if the investor traded more than 40 securities. Using dummy variables allows us to better capture any non-linearity in the relationship between diversification and performance and minimizes concerns that investors that trade an extreme number of securities are driving the results. Consistent with our hypothesis, timing skill is positively correlated with the number of securities the investor trades. Investors that trade more than 40 securities are the most likely to be in the top 20% and least likely to be in the bottom 20%. We also find that investors with a larger portion of their transactions in the most traded stock (in % of euro value), Nokia, are less likely to be good at market timing. Good market timers appear to diversify their holdings, but, surprisingly, they are not more likely to use the OMX ETF - a low-cost strategy to time the market. There are two possible explanations for this result. We may not have enough variation across investors as the OMX ETF accounts for only 0.04% of transactions or unsophisticated investors could be using the OMX ETF to achieve low-cost diversification. In summary, a stereotype of a good timer is a middle age man living in a less educated area. He has enough wealth to invest in the stock market, but is not very wealthy. He trades infrequently, does not trade options and diversifies his portfolio.

2.6 Survivorship

In previous tests we have shown that investors exhibit persistence in their market timing ability and this timing ability is large and economically significant. In this subsection, we ask, do investors learn about their abilities? To answer this question we examine how first period market timing performance affects the probability of an investor stopping active participation in the market. In Table 13, we calculate the probability an investor becoming inactive (zero absolute monthly flow in the second period) for each first period performance quintile. We use the entire period timing measure as the first period performance measure. The results are mixed. Investors in the worst market timing quintile are about 42% more likely to drop out of the sample than investors in the top market timing quintile and the difference is significant at the 0.00 level. Investors in the middle quintiles are more likely to drop out than either the top or bottom quintile. Thus, the relationship is nonlinear in a way that investors with average timing performance are most likely to drop out.

The nonlinearity eases concerns that survivorship may be affecting our results. Furthermore, we examine whether our sample selection process affects our inferences by performing the same analysis with investors that are active (having at least 15 months with nonzero flows) in both sample halves instead of classifying based on just first half activity. The results, presented in the online appendix, are very similar to those in Table 4, so we conclude that our sample selection procedure is reasonable and survivorship is not driving the observed market timing persistence.

2.7 Additional Tests

In this subsection, we examine several possible explanations for our results besides heterogeneity in investor market timing ability. We adjust our timing measure in various ways to rule out alternative theories for the observed persistence and show the robustness of our results. We also examine the ability (or inability) of financial institutions to persistently time the market in our sample.

Finland is a relatively unique market in that Nokia makes up approximately 50% of the market capitalization during our sample period. Although the market weight of Nokia in our index is capped at 20% until August 1, 2001 and then capped at 10% thereafter, one possible explanation for our results is that investors are just timing movements in Nokia and this is driving our results. To address this concern we run two tests. First, we test whether investors can persistently time Nokia, by correlating investor monthly flows into and out of the market with Nokia returns over the next month. The results are presented in the online appendix. The persistence is similar to the results using market returns, with a nearly monotonic relationship between first and second period performance. Because Nokia's returns are correlated with market returns, this result may not be surprising and does not necessarily mean Nokia is driving our results. To determine whether investors are timing the market, not solely Nokia, we run a similar test, but omit flows into and out of Nokia and exclude Nokia's returns from the index. The detailed results are in the online appendix. Once again, we see a near monotonic relationship between first and second period performance and very significant departures from the null of no timing, so it is highly unlikely Nokia alone is driving the timing persistence we observe.

Investors may time the market at various frequencies: daily, monthly, quarterly, or longer. We focus on monthly timing for two reasons. First, we cannot reliably estimate timing over longer frequencies due to the length of our sample. Second, a very small percentage of individuals trade on a daily basis in our sample and trading daily would be a costly activity for most individual investors. Goetzmann, Ingersoll, and Ivković (2000) show that estimating timing ability of daily timers at a monthly frequency will create downward bias in the estimation of timing skill when using a Henrikson and Merton (1981) returns-based measure. It is unclear if a similar bias exists with our timing measure and if the rank-order of investors would change significantly if we changed the frequency. As a robustness check, we re-estimate our main sets of analysis with a quarterly timing measure. This measure is calculated following equation (1) except we replace the cash return in month $t+1$ with the cash return from the start of month $t+1$ to the end of month $t+3$. The results are presented in the online appendix and are very similar to those for the monthly measure. The monthly and quarterly measures are highly correlated across investors with a correlation of 0.75 (0.66) in the first period (second period). This gives us confidence that our results are robust to measuring timing over different frequencies.

In Table 14, we look for persistent timing ability among financial institutions. We perform the same kind of test for institutions that we perform for individuals. However, our test is somewhat limited by the relatively small number of active institutions in our data; there are only 324 institutions that trade sufficiently in both halves of the sample to be included in our analysis. The results in the tables reveal that there is no clear evidence of timing for these firms. This is not too surprising given the low sample size of our tests, the objectives of most financial institutions and the lack of control institutions have over their customers' inflows and outflows. We would not expect, for example, for market makers, index funds, or standard equity funds to display any timing ability (as captured by our measure) for these reasons.

3 Conclusion

We document significant persistence in the ability of individual investors to time the stock market, both in general and during market bubble periods. We find that some investors consistently time the market while others consistently mis-time the market. The results are especially surprising considering the two market episodes examined have very different economic drivers. The observed persistence implies that there must be some predictability in returns and there must be variation in the skills of individuals in processing news about markets and the economy. We find that information in investor flows is a better predictor of future market returns than commonly used economic variables. A trading strategy based on investor flows has a reward to risk ratio of 0.81, which is over sixteen times the market ratio of 0.05. We also show that the probability of a bear market goes from 24% to 60% when good timers flows are relatively low compared to bad timers flows. Our evidence is consistent with models that feature both rational arbitrageurs and noise traders, such as DeLong, Shleifer, Summers and Waldman (1990) and Abreu and Brunnermeier (2003).

The fact that some investors consistently time the market has important implications for the way we model markets and investor behavior. It means, of course, that the market cannot be perfectly informationally efficient. If there is a lot of dispersion in the skill of investors then people may rationally incur significant costs to improve their skills. They may trade in an experimental fashion to learn, or they may use financial products that do not make sense in a world characterized by efficient markets. If companies can successfully time the market they may want to issue shares at times when prices are abnormally high (see Baker and Wurgler (2000, 2002)). They may also want to time their share repurchases or their granting of executive stock options. We leave a detailed exploration of the implications of market timing to future research.

References

Abreu, Dilip and Markus Brunnermeier, 2003, Bubbles and Crashes, *Econometrica* 71, 173-204.

Baker, Malcolm and Jeffrey Wurgler, 2000, The Equity Share in New Issues and Aggregate Stock Returns, *Journal of Finance* 55, 2219-57.

Baker, Malcolm and Jeffrey Wurgler, 2002, Market Timing and Capital Structure, *Journal of Finance* 57, 1-32.

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

Barber, Brad, and Terrance Odean, 2001, Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment, *The Quarterly Journal of Economics* 116, 261-292.

Barber, Brad, and Terrance Odean, 2008, All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors, *Review of Financial Studies* 21, 785-818.

Barber, Brad, Terrance Odean and Ning Zhu, 2009, Do Retail Trades Move Markets? *Review of Financial Studies* 22, 151-186.

Becker, Connie, Wayne Ferson, David H. Myers and Michael J. Schill, 1999, Conditional Market Timing with Benchmark Investors, *Journal of Financial Economics* 52, 119-148

Bollen, Nicolas P.B. and Jeffrey A. Busse, 2001, On the Timing Ability of Mutual Fund Managers, *The Journal of Finance* 56, 1075-1094

Brunnermeier, Markus and Stefan Nagel, 2004, Hedge Funds and the Technology Bubble, *Journal of Finance* 59, 2007-2034.

Campbell, John Y. and Luis M. Viceira, 2002, *Strategic Asset Allocation*, Oxford University Press.

Che, Limei, Oyvind Norli, and Richard Priestly, 2009, Performance Persistence of Individual Investors, working paper, Norwegian Business School.

Che, Limei, Oyvind Norli, and Richard Priestly, 2012, Market Timing by Individual Investors, working paper, Norwegian Business School.

Coval, Joshua D., David Hirshleifer and Tyler Shumway, 2012, Can Individual Investors Beat the Market?, working paper, University of Michigan.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring Mutual Fund Performance with Characteristic Based Benchmarks, *Journal of Finance* 52, 1035-1058.

Delong, J. Bradford, Andrei Shleifer, Lawrence Summers and Robert Waldmann, 1990, Positive Feedback Investment Strategies and Destabilizing Rational Speculation, *Journal of Finance* 45, 375-395.

Dichev, Ilia D., 2007, What Are Stock Investors' Actual Historical Returns? Evidence from Dollar-Weighted Returns, *American Economic Review* 97, 386-402.

Dorn, Daniel, Gur Huberman and Paul Sengmueller, 2008, Correlated Trading and Returns, *Journal of Finance* 58, 885-920.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 2011, An Examination of Mutual Fund Timing Ability Using Monthly Holdings Data, *Review of Finance* 16, 619-645.

Ferson, Wayne E. and Rudi W. Schadt, 1996, Measuring Fund Strategy and Performance in Changing Economic Conditions , *The Journal of Finance* 51, 425-461

Goetzmann, William N., 2015, Bubble Investing: Learning From History, No. w21693, National Bureau of Economic Research

Goetzmann, William N., Jonathan Ingersoll Jr. and Zoran Ivković, 2000, Monthly Measurement of Daily Timers, *The Journal of Financial and Quantitative Analysis* 35, 257-290

Goyal, Amit and Ivo Welch, 2008, A Comprehensive Look at the Empirical Performance of Equity Premium Prediction, *Review of Financial Studies* 21, 1455-1508.

Greenwood, Robin, Andrei Shleifer and Yang You, 2017, Bubbles for Fama, No. 23191, National Bureau of Economic Research

Grinblatt, Mark, 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, Matti Keloharju, 2001a, What Makes Investors Trade? *Journal of Finance* 56, 589-616.

Grinblatt, Mark, Matti Keloharju, 2001b, How Distance, Language, and Culture Influence Stockholdings and Trades, *Journal of Finance* 56, 1053-1073.

Grinblatt, Mark, Matti Keloharju and Juhani T. Linnainmaa, 2012, IQ, Trading Behavior, and Performance, *Journal of Financial Economics* 104, 339-362.

Henriksson, Roy D. and Robert C. Merton, 1981, On Market Timing and Investment Performance II. Statistical Procedures for Evaluating Forecasting Skills, *Journal of Business* 54, 513-533.

Henriksson, Roy D., 1984, Market Timing and Mutual Fund Performance: An Empirical Investigation, *The Journal of Business* 57, 73-96

Hong, Harrison and Jeremy Stein, 2007, Disagreement and the Stock Market, *Journal of Economic Perspectives* 21, 109-128.

Hvidkjaer, Soeren, 2008, Small Trades and the Cross-Section of Stock Returns, *Review of Financial Studies* 21, 1123-1151.

Jagannathan, Ravi, Robert A. Korajczyk, 2014, Market Timing, Working Paper, Northwestern University.

Kacperczyk, Marcin T. and Amit Seru, 2007, Fund Manager Use of Public Information: New Evidence on Managerial Skills, *The Journal of Finance* 62, 485-528.

Kacperczyk, Marcin T., Stijn Van Nieuwerburgh and Laura Veldkamp, 2014, Time-Varying Fund Manager Skill, *Journal of Finance* 69, 1414-1441.

Kaniel, Ron, Shuming Liu, Gideon Saar and Sheridan Titman, 2012, Individual Investor Trading and Return Patterns around Earnings Announcements, *Journal of Finance* 67, 639-680.

Kaniel, Ron, Gideon Saar and Sheridan Titman, 2009, Individual Investor Trading and Stock Returns, *Journal of Finance* 63, 273-310.

Kaustia, Markku, and Samuli Knüpfer, 2012, Peer Performance and Stock Market Entry, *Journal of Financial Economics* 104, 321-338.

Kelley, Eric K. and Paul C. Tetlock, 2013, How Wise are Crowds? Insights from Retail Orders and Stock Returns *Journal of Finance* 68, 1229-1265.

Linnainmaa, Juhani, 2010, Do Limit Orders Alter Inferences about Investor Performance and Behavior? *Journal of Finance* 65, 1473-1506.

Mamaysky, Harry, Matthew Spiegel and Hong Zhang, 2008, Estimating the Dynamics of Mutual Fund Alphas and Betas, *Review of Financial Studies* 21, 233-264.

Merton, Robert C., 1981, On Market Timing and Investment Performance I. An Equilibrium Theory of Value for Market Forecasts, *Journal of Business* 54, 363-406.

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

Pástor, Ľuboš and Pietro Veronesi, 2003, Technological Revolutions and Stock Prices, *American Economic Review* 99, 1451-1483.

Seru, Amit, Tyler Shumway and Noah Stoffman, 2010, Learning by Trading, *Review of Financial Studies* 23, 705-739.

Skirbekk, Vegard, 2004, Age and Individual Productivity: A Literature Survey, *Vienna Yearbook of Population Research* 2, 133-153.

Shiller, Robert J., 2000, *Irrational Exuberance,* Princeton University Press.

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

Yermack, David, 1997, Good Timing: CEO Stock Option Awards and Company News Announcements, *The Journal of Finance* 52, 449-476.

Zeira, Jozeph, 1999, Information Overshooting, Booms and Crashes, *Journal of Monetary Economics* 43, 237-257.

Table 1: Summary Statistics of Investor Monthly Flows

Year	# Trades	# Flows	Mean	Std. Dev.	Outflows	Flow= 0	Inflows
1995	269,670	581,630	119.48	18,307.01	5.08%	82.50%	12.42%
1996	362,778	622,318	-722.73	28,609.67	10.37%	81.33%	8.30%
1997	504,473	656,227	268.13	39,415.36	9.06%	73.74%	17.19%
1998	746,684	725,097	322.13	36,271.59	11.01%	68.37%	20.62%
1999	1,416,253	793,227	818.65	116,432.90	16.50%	57.21%	26.28%
2000	2,142,002	823,453	-2,306.98	210,042.00	19.91%	52.52%	27.57%
2001	1,709,249	827,244	85.79	71,351.88	14.78%	63.60%	21.63%
2002	1,277,208	827,244	-431.95	124,862.10	17.05%	67.99%	14.96%
2003	1,089,789	827,244	459.57	49,009.56	10.18%	74.39%	15.43%
2004	1,290,402	827,244	1,381.05	88,048.21	10.31%	71.31%	18.38%
2005	1,624,510	827,244	-941.24	142,472.60	17.25%	68.81%	13.95%
2006	1,808,329	827,244	-1,687.51	290,959.70	13.29%	74.34%	12.37%
2007	2,323,850	827,244	-551.95	121,005.80	13.09%	75.48%	11.43%
2008	2,390,962	827,244	974.97	59,321.56	6.67%	79.54%	13.80%
2009	1,399,064	413,622	713.16	32,935.14	9.28%	70.39%	20.34%

This table displays summary statistics of monthly investor flows into and out of securities on the Helsinki Stock Exchange. Our sample contains all transactions by individual investors from January 1995 to June 2009. Statistics are presented for the active traders in our sample. To be considered an active trader, investors must have monthly absolute flows greater than zero in a minimum of 15 months during the first half. **# Trades** is the number of trades made during the year. **# Flows** is the number of monthly flows aggregated at the investor-month level. **Mean** is the average investor-month flow size in euros. **Std. Dev.** is the standard deviation of investor-month flows. **Outflows** is the percentage of flows that are outflows during the year. **Flow**= 0 is the percentage of flows equal to 0 during the year. **Inflows** is the percentage of flows that are inflows during the year. We drop all trades with 0 value and all cancelled trades from our original transaction data.
*The 2009 values are for the first six months of the year.

Table 2: Summary Statistics of Market Timing Measures

Time Period	Flow Freq.	Return Freq.	Mean	Std. Dev.	25th	Median	75th	N
Entire Period								
1995-2002	Monthly	Monthly	0.03	0.18	-0.07	0.01	0.13	68,937
2002-2009	Monthly	Monthly	0.00	0.10	-0.06	0.00	0.07	66,840
Bubble Period								
2000 Bubble	Monthly	Monthly	0.03	0.25	-0.12	0.01	0.19	68,717
2007 Bubble	Monthly	Monthly	0.00	0.19	-0.13	0.00	0.13	51,061

This table gives the summary statistics (mean, standard deviation, 25th percentile, median and 75th percentile) for the market timing measure. The monthly measure is calculated using equation (1). We separate the sample into two equal length sub-periods: January 1995 - March 2002 and April 2002 - June 2009. Timing measures calculated for each half are presented under the "Entire Period" heading. For the "Bubble Period" measures, we center our analysis around the market peak during the relevant half of our sample (February 2000 and October 2007). We calculate each investor's performance during the time period from 12 months before the peak to 12 months after (e.g., for the 2000 bubble, the performance period is February 1999 to February 2001). Statistics are presented for the active traders in our sample. To be considered an active trader, investors must have monthly absolute flows greater than zero in a minimum number of months during the first half. For the entire period ("Bubble Period") measure, the minimum is 15 (8) months. N is the number of investors that meet the criteria and have enough flows to calculate a correlation.

Table 3: Two Period Cross-Tab of the Entire Period Monthly *Beta-Adjusted* Timing Measure

	Second Period					
First Period	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Total**
Q1	24.88%***	21.24%***	18.66%***	17.95%***	17.26%***	100%
Q2	19.44%*	19.90%	20.34%	20.03%	20.30%	100%
Q3	19.46%*	20.37%	20.39%	20.33%	19.45%*	100%
Q4	19.12%**	19.31%**	20.58%*	19.95%	21.04%***	100%
Q5	17.01%***	19.15%**	20.07%	21.77%***	22.00%***	100%
Total	100%	100%	100%	100%	100%	

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Null: Cell%=20%.

This table provides frequencies of investors sorted and grouped by their monthly timing beta-adjusted measure in each of the two sample sub-periods (January 1995 - March 2002 and April 2002 - June 2009). The beta-adjusted timing measure is calculated using equation (2). The January 1995 - March 2002 (April 2002 - June 2009) percentile rank is along the vertical (horizontal) axis. The timing measures are grouped into quintiles. **Q1** is the top performance quintile. We present row percentages. If the two periods were independent, we would expect row percentages of 20% in each cell. p-values are calculated using bootstrapping procedures conducted as follows: each first period investor is matched with a second period investor based on their geographic region (9 regions in Finland) and their within-region tercile ranking of the number of trades. We match first and second period investors 1,000 times. Significance levels are based on two-sided tests of significance. Significance levels are nearly identical if basic OLS standard errors are used to calculate significance. The quintile cut-off values for the first period are: -0.107, -0.020, 0.067, 0.241. For the second period, the cut-off values are: -0.087, -0.027, 0.025, 0.085. The sample size is 66,596 investors.

The pairwise correlation between the first and second period monthly timing measures is 0.0711 and is significant at the 0.01% level. The spearman rank correlation coefficient is 0.0692 and is significant at the 0.01% level.

Table 4: Two Period Cross-Tab of the Entire Period Monthly Timing Measure

	Second Period						Average Second
First Period	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Total**	**Period Timing**
Q1	24.43%***	20.57%*	19.35%**	18.36%***	17.29%***	100%	1.56%***
Q2	20.20%	20.31%	19.93%	20.03%	19.53%*	100%	0.55%
Q3	19.42%*	20.31%	20.70%**	19.78%	19.79%	100%	0.46%
Q4	18.74%***	20.05%	20.25%	20.27%	20.69%**	100%	0.14%
Q5	17.16%***	18.76%***	19.78%	21.58%***	22.71%***	100%	-0.63%
Total	100%	100%	100%	100%	100%		Q1-Q5 = 2.21%***

*** p<0.01, ** p<0.05, * p<0.1. Null: Cell%=20%.

This table provides frequencies of investors sorted and grouped by their monthly timing measure in each of the two sample sub-periods (January 1995 - March 2002 and April 2002 - June 2009). The monthly timing measure is calculated using equation (1). The January 1995 - March 2002 (April 2002 - June 2009) percentile rank is along the vertical (horizontal) axis. The timing measures are grouped into quintiles. **Q1** is the top performance quintile. We present row percentages. If the two periods were independent, we would expect row percentages of 20% in each cell. p-values are calculated using bootstrapping procedures conducted as follows: each first period investor is matched with a second period investor based on their geographic region (9 regions in Finland) and their within-region tercile ranking of the number of trades. We match first and second period investors 1,000 times. Significance levels are based on two-sided tests of significance. Significance levels are nearly identical if basic OLS standard errors are used to calculate significance. The quintile cut-off values for the first period are: -0.098, -0.019, 0.051 & 0.173. For the second period, the cut-off values are: -0.079, -0.022, 0.028 & 0.088. Avg. Second Period Timing provides the average second half timing measure (April 2002 - June 2009) of investors sorted and grouped by their monthly timing measure in the first half of the sample (January 1995 - March 2002). The sample size is 66,840 investors.

The pairwise correlation between the first and second period monthly timing measures is 0.0745 and is significant at the 0.01% level. The spearman rank correlation coefficient is 0.0726 and is significant at the 0.01% level.

Table 5: Two Period Cross-Tab of the Bubble Period Monthly Timing Measure

	2007 Bubble					
2000 Bubble	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Total**
Q1	21.96%***	19.95%	19.20%**	19.50%	19.39%*	100%
Q2	19.80%	20.11%	19.76%	20.31%	20.03%	100%
Q3	19.78%	19.80%	19.66%	20.53%*	20.23%	100%
Q4	19.41%**	20.26%	19.47%	20.87%***	19.98%	100%
Q5	18.91%***	19.97%	20.01%	20.73%***	20.38%	100%
Total	100%	100%	100%	100%	100%	

*** p<0.01, ** p<0.05, * p<0.1. Null: Cell%=20%

This table provides frequencies of investors sorted and grouped by their 2000 and 2007 market bubble monthly timing measure. The monthly timing measure is calculated using equation (1). The 2000 (2007) percentile rank is along the vertical (horizontal) axis. The timing measures are grouped into quintiles. **Q1** is the top performance quintile. We present row percentages. If the two periods were independent, we would expect row percentages of 20%. p-values are calculated using bootstrapping procedures conducted as follows: each first period investor is matched with a second period investor based on their geographic region (9 regions in Finland) and their within-region tercile ranking of the number of trades. We match first and second period investors 1,000 times. Significance levels are based on two-sided tests of significance. Significance levels are nearly identical if basic OLS standard errors are used to calculate significance. The quintile cut-off values for the first period are: -0.161, -0.039, 0.070, 0.239. For the second period, the cut-off values are: -0.155, -0.052, 0.047, 0.156. The sample size is 51,061 investors.

The pairwise correlation between the first and second period monthly timing measures is 0.0250 and is significant at the 0.01% level. The spearman rank correlation coefficient is 0.0212 and is significant at the 0.01% level.

Table 6: Two Period Cross-Tab of Significant Outflows Around the Market Peak

	2007 Bubble					
2000 Bubble	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Total**
Q1	19.82%	21.70%**	18.57%***	20.75%**	19.16%**	100%
Q2	20.10%	19.69%	19.91%	20.49%	19.81%	100%
Q3	19.96%	20.09%	20.58%*	19.92%	19.45%	100%
Q4	20.31%	19.79%	21.26%***	18.89%***	19.74%	100%
Q5	19.85%	18.67%***	19.49%	20.00%	22.00%***	100%
Total	100%	100%	100%	100%	100%	

*** p<0.01, ** p<0.05, * p<0.1. Null: Cell%=20%

This table provides frequencies of investors sorted and grouped by their flows around the market peaks in 2000 and 2007. The significant flow for investor i is calculated as follows: $-\frac{1}{6}\sum_{PeakMonth-5}^{PeakMonth}\frac{flow_{im}-\overline{flow_i}}{s_i}$, where $flow_{im}$ is the flow of investor i in month m, $\overline{flow_i}$ is the average flow for investor i over the sample half in which the bubble occurs, s_i is the standard deviation of flows for investor i over the sample half in which the bubble occurs, $PeakMonth$ is the month the market reached its apex during the sample half. The 2000 (2007) percentile rank is along the vertical (horizontal) axis. The flow measures are grouped into quintiles. **Q1** is the top performance quintile. We present row percentages. If the two periods were independent, we would expect row percentages of 20%. p-values are calculated using bootstrapping procedures conducted as follows: each first period investor is matched with a second period investor based on their geographic region (9 regions in Finland) and their within-region tercile ranking of the number of trades. We match first and second period investors 1,000 times. Significance levels are based on two-sided tests of significance. Significance levels are nearly identical if basic OLS standard errors are used to calculate significance. The quintile cut-off values for the first period are: -0.420, -0.025, 0.257, 0.684. For the second period, the cut-off values are: -0.158, -0.074, 0.034, 0.237. The sample size is 65,370 investors.

The pairwise correlation between the first and second period monthly timing measures is 0.0108 and is significant at the 1% level. The spearman rank correlation coefficient is 0.0150 and is significant at the 0.01% level.

Table 7: Two Period Cross-Tab of the Stock Picking Measure

	Second Period					
First Period	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Total**
Q1	22.14%***	20.64%	18.43%***	17.51%***	21.28%**	100%
Q2	19.30%	21.13%**	22.41%***	19.97%	17.19%***	100%
Q3	17.57%***	20.30%	21.84%***	21.58%***	18.71%**	100%
Q4	19.81%	19.20%	20.35%	20.90%*	19.74%	100%
Q5	21.19%**	18.76%**	16.92%***	20.02%	23.11%***	100%
Total	100%	100%	100%	100%	100%	

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Null: Cell%=20%.

This table provides frequencies of investors sorted and grouped by their stock picking measure in each of the two sample sub-periods (January 1995 - March 2002 and April 2002 - June 2009). The stock picking measure is calculated according to Seru, Stoffman and Shumway (2009). For all stock purchases, we calculate the return over the next 30 days less the market return. If the investor sells the stock before 30 days, we use the holding period return less the market return. The stock picking measure is the average of these returns over all stock purchases. The January 1995 - March 2002 (April 2002 - June 2009) percentile rank is along the vertical (horizontal) axis. The stock picking measures are grouped into quintiles. **Q1** is the top performance quintile. We present row percentages. If the two periods were independent, we would expect row percentages of 20% in each cell. The p-values are calculated using OLS regressions. The quintile cut-off values for the first period are: -3.0%, -1.3%, -0.0% & 1.5%. For the second period, the cut-off values are: -1.5%, -0.4%, 0.4% & 1.4%. The sample size is 24,521 investors. Investors are included if they trade at least 75 times in the first half of the sample.

The pairwise correlation between the first and second period monthly timing measures is 0.0057 and is insignificant. The spearman rank correlation coefficient is 0.0239 and is significant at the 0.01% level.

Table 8: Cross-correlation Between Market Timing Group Flows and Return Predictors

Variables	Top 20-Bottom 20	Top 20 Flow	Bottom 20 Flow	Log(EP ratio)	Log(Div. Yield)	$HEX25_t$
Top 20-Bottom 20	1.000					
Top 20 Flow	0.696	1.000				
	(0.000)					
Bottom 20 Flow	0.193	0.839	1.000			
	(0.074)	(0.000)				
Log(EP ratio)	0.253	0.352	0.290	1.000		
	(0.018)	(0.001)	(0.007)			
Log(Div. Yield)	-0.011	0.346	0.482	0.369	1.000	
	(0.917)	(0.001)	(0.000)	(0.000)		
$HEX25_t$	0.259	-0.224	-0.502	-0.243	-0.339	1.000
	(0.015)	(0.037)	(0.000)	(0.023)	(0.001)	

This table presents correlations between group flows in equation (3) and various return predictors over the time period April 2002 to June 2009. Top 20 Flow (Bottom 20 Flow) is the month t standardized group flow from for the top (bottom) quintile of timers sorted by first half performance measured using equation (1). Top 20-Bottom 20 which is the month t standardized difference in group flows between the top and bottom quintile of timers sorted by first half performance. Log(EP ratio) is the logarithm of the earnings-price ratio of the HEX25 at the end of month t less the monthly Euribor rate. Log(Div. Yield) is the logarithm of the dividend yield of the HEX25 at the end of month t. $HEX25_t$ is the return on the HEX25 in month t. p-values are reported in parentheses.

Table 9: In-Sample Predictability Regressions

VARIABLES	(1) $HEX25_{t+1}$	(2) $HEX25_{t+1}$	(3) $HEX25_{t+1}$	(4) $HEX25_{t+1}$	(5) $HEX25_{t+1}$	(6) $HEX25_{t+1}$	(7) $HEX25_{t+1}$	(8) $HEX25_{t+1}$
Top 20-Bottom 20	0.0145**							0.0160**
	(0.00687)							(0.00749)
Top 20 Flow		0.00654						
		(0.00701)						
Bottom 20 Flow			-0.00205					
			(0.00705)					
Log(EP ratio)				-0.0337			-0.0270	-0.0471
				(0.0256)			(0.0277)	(0.0287)
Log(Div. Yield)					-0.0149		0.0120	0.0135
					(0.0277)		(0.0306)	(0.0300)
$HEX25_t$						0.200*	0.188	0.105
						(0.107)	(0.115)	(0.120)
Observations	87	87	87	87	87	87	87	87
R-squared	0.050	0.010	0.001	0.020	0.003	0.040	0.050	0.101

Standard Errors in Parentheses
*** p<0.01 ** p<0.05 * p<0.1

This table presents predictive OLS regressions for the market excess return over the time period April 2002 to June 2009. The dependent variable is the monthly HEX25 return in month t+1 less the monthly Euribor rate ($HEX25_{t+1}$). Top 20-Bottom 20 is the month t standardized difference in group flows between the top and bottom quintile of timers sorted by first half performance (see equation 3). Top 20 Flow (Bottom 20 Flow) is the month t standardized group flow from for the top (bottom) quintile of timers sorted by first half performance (see equation 4). Log(EP ratio) is the logarithm of the earnings-price ratio of the HEX25 at the end of month t. Log(Div. Yield) is the logarithm of the dividend yield of the HEX25 at the end of month t. $HEX25_t$ is the return on the HEX25 in month t less the monthly Euribor rate. R-squared is the unadjusted R^2.

Table 10: Second Half Performance Measures

Timing	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Top 20%-Bot. 20%**	**Passive**
Panel A: Correlation between flow$_t$ and HEX25 Excess Return$_{t+1}$							
Entire Period	0.10	0.06	0.04	0.03	-0.03	0.22**	0.00
Bubble Period	0.08	0.04	0.01	0.01	-0.03	0.21**	0.00
Panel B: Average Flow-Weighted Excess Return							
Entire Period	7.75	4.30	3.29	2.00	-2.43	17.19**	1.02
Bubble Period	6.44	2.95	1.01	0.61	-1.93	16.44**	1.02
Panel C: Flow-Weighted Return-Volatility Ratio							
Entire Period	0.32	0.18	0.14	0.08	-0.10	0.81	0.05
Bubble Period	0.27	0.12	0.04	0.03	-0.08	0.81	0.05

This table provides measures of performance in the second half of our sample for investors grouped by first period performance. We present results sorting by the two main timing measures. We by market timing calculated during the entire period (**Entire Period**) and by market timing calculated during the Bubble Period (**Bubble Period**). Panel A presents the correlations between quintile group flows in month t and market returns in month $t+1$. Panel B presents the average flow-weighted return, labeled Average Flow-Weighted Return, calculated according to equation (5). This measure multiplies the return in month $t+1$ by the group flow in month t. Panel C presents the ratio of the average flow-weighted return to the standard deviation of the flow-weighted return, labeled Flow-Weighted Return-Volatility Ratio. The quintile group flows are calculated according to equation (3). The Top 20% - Bot. 20% flow is calculated according to equation (4).

Table 11: Predicting Negative Returns With Difference Between Good and Bad Timers' Flows

Timing Measure	P(Bear Mkt)	P(Low Flow)	P(Low Flow |Bear Mkt)	P(Bear Mkt |Low Flow)
Entire Period	24.1%	11.5%	28.6%	**60.0%**
Bubble Period	24.1%	16.1%	28.6%	**42.9%**

This table presents the probability of a "large" negative excess return in month $t+1$ conditional on a large negative difference in flows between the good and bad timers in month t. We examine 2 timing measures: market timing calculated over the entire period (**Entire Period**) and market timing calculated in the Bubble Period (**Bubble-Monthly**). The timing measures are calculated in the relevant months of the first half of our sample (January 1995 - March 2002) whereas all the probabilities are calculated in the second half of our sample (April 2002 - June 2009). There are 87 months in the second half of our sample. In column 1, we provide the unconditional probability of a "bear" market. A "bear" market is defined as a monthly excess return (HEX25 minus 1m-Euribor) that is at least half of one standard deviation below the mean excess return. In column 2, we provide the unconditional probability of a "Low Flow" defined as a difference between the top and bottom timers flows that is at least half of one standard deviation below the average. The difference between the top and bottom group flows is calculated according to equation (4). In column 3, we provide the probability of "Low Flow" given the next month is a "bear" market. In column 4, we provide the probability of a "bear" market in month $t+1$ given investors have a significant outflow in month t.

Table 12: Timing Investors Characteristic Regressions

VARIABLES	(1) Top 20	(2) Bottom 20	(3) Timing Skill
Male	0.0250*	0.00746	0.0132
	(0.0142)	(0.0141)	(0.0133)
Age 25-45	0.0327	0.00920	0.00931
	(0.0205)	(0.0200)	(0.0194)
Age 46-64	0.128***	-0.0541***	0.112***
	(0.0211)	(0.0207)	(0.0199)
Age 65+	0.0905***	-0.104***	0.104***
	(0.0280)	(0.0277)	(0.0258)
Density	0.161	-0.496*	0.193
	(0.292)	(0.287)	(0.269)
University %	-0.00230***	-0.000493	-0.000932
	(0.000664)	(0.000647)	(0.000612)
Finance %	-0.000508	-0.000160	-0.000510
	(0.00210)	(0.00206)	(0.00193)
Finnish	-0.000962	0.0169	-0.0403**
	(0.0202)	(0.0195)	(0.0183)
Option	-0.169***	0.0117	-0.109***
	(0.0292)	(0.0250)	(0.0251)
OMX ETF	-0.00333	0.0245	-0.0363
	(0.0464)	(0.0421)	(0.0415)
Nokia Flow %	-0.893***	0.727***	-1.041***
	(0.0378)	(0.0321)	(0.0320)
Avg. Beta	1.341***	0.250***	0.693***
	(0.0359)	(0.0350)	(0.0334)
Log(Trades)	-0.318***	0.156***	-0.261***
	(0.0111)	(0.00911)	(0.00902)
Log(Flow Size)	-0.111***	-0.0487***	-0.0236***
	(0.00766)	(0.00718)	(0.00687)
26-40 Securities	0.317***	-0.162***	0.282***
	(0.0158)	(0.0155)	(0.0147)
41+ Securities	0.466***	-0.229***	0.418***
	(0.0216)	(0.0202)	(0.0192)
Observations	62,879	62,879	62,879
R-squared			0.049

Standard Errors in Parentheses
*** p<0.01 ** p<0.05 * p<0.1

This table presents results for regressions of timing ability on investor characteristics. We use investors monthly timing ability measured over the entire period for these tests. We run the regressions for 3 skill measures: Top 20, Bottom 20 and Timing Skill. Top 20 is a dummy variable equal to 1 if the investor is in the top 20% of investors. Bottom 20 is a dummy variable equal to 1 if the investor is in the bottom 20% of investors. Timing Skill is a variable equal to 5 if the investor is in the top 20%, 4 if the investor is in the Top 20-40%, 3 if the investor is in the Top 40-60%, etc. The regressions in columns 1 and 2 are probit regressions. The regression in column 3 is an OLS regression. Male is a dummy variable equal to 1 if the investor is a male. Age 25-45 is a dummy variable equal to 1 if the investor is 25 to 45 years old at the end of 1995. Age 46-64 is a dummy variable equal to 1 if the investor is 46 to 64 years old at the end of 1995. Age 65+ is a dummy variable equal to 1 if the investor is 65 years old or older at the end of 1995. Density is the population density of the investor's zip code times 10^{-5}. University % is the percentage of persons in the investor's zip code with a university degree. Finance Profession % is the percentage of persons in the investor's zip code working in the finance industry. Finnish is a dummy equal to one if the individual's primary language is Finnish. Option is an indicator variable equal to 1 if the investor ever trades an option. OMX ETF is an indicator variable equal to 1 if the investor transacted in the OMX ETF during the sample period. Nokia Flow % is the value percentage of absolute flows in Nokia. Avg. Beta is the average beta of all securities the investor traded. Log(Trades) is the logarithm of the total number of transactions placed by the investor over the entire sample. Log(Flow Size) is the logarithm of the investor's mean absolute monthly flow over the entire sample. 26-40 Securities (41+ Securities) is a dummy variable equal to one if the number of securities (unique CUSIPS) the investor transacted in during the sample period is between 26 and 40 securities (41+ securities). The R^2 reported for the probit regressions in columns 1 and 2 are Pseudo-R^2. The reported R^2 from the OLS regression in column 3 is the adjusted-R^2. We report the number of observations in the last row and z and t-stats are reported in parentheses.

Table 13: Investor Survivorship

First Period Performance	**% 2nd Half Inactives**
Q1 (Best)	2.07 %
Q2	3.34 %
Q3	3.75 %
Q4	3.12 %
Q5 (Worst)	2.93 %
t-test (Top-Bottom) p-value	0.00

This table presents the percentage of investors within each first period performance quintile that do not move money in or out of the market in the second half of the sample (Inactives). First Period Performance is the first period performance quintile, measured using the monthly timing measure. 2nd Half Inactives is the percentage of investors within the quintile that did not move money in or out of the market during the second half of the sample period (April 2002 - June 2009). The last row presents the p-value for a student's t-test for a difference in means between the best (Q1) and worst (Q5) investor groups.

Table 14: Institutions Entire Period Monthly Timing Measure

	Second Period					
First Period	**Q1**	**Q2**	**Q3**	**Q4**	**Q5**	**Total**
Q1	14.75%	21.31%	24.59%	18.03%	21.31%	100%
Q2	32.14%**	19.64%	10.71%*	16.07%	21.43%	100%
Q3	15.38%	23.08%	26.15%	13.85%	21.54%	100%
Q4	18.31%	18.31%	25.35%	23.94%	14.08%	100%
Q5	19.72%	18.31%	12.68%*	26.76%	22.54%	100%
Total	100%	100%	100%	100%	100%	

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Null: Cell%=20%.

This table provides frequencies of institutions sorted and grouped by their monthly timing measure in each of the two sample sub-periods (January 1995 - March 2002 and April 2002 - June 2009). The monthly timing measure is calculated using equation (1). The January 1995 - March 2002 (April 2002 - June 2009) percentile rank is along the vertical (horizontal) axis. The timing measures are grouped into quintiles. **Q1** is the top performance quintile. We present row percentages. If the two periods were independent, we would expect row percentages of 20% in each cell. p-values are calculated using OLS standard errors. There are 369 institutions that have non-zero flows in at least 15 months during the first period, 324 of these institutions have at least 2 months of non-zero flows in the second period.

The pairwise correlation between the first and second period monthly timing measures is 0.0399 and is insignificant with a p-value of 0.4746. The spearman rank correlation coefficient is 0.0449 and is insignificant with a p-value of 0.4206.

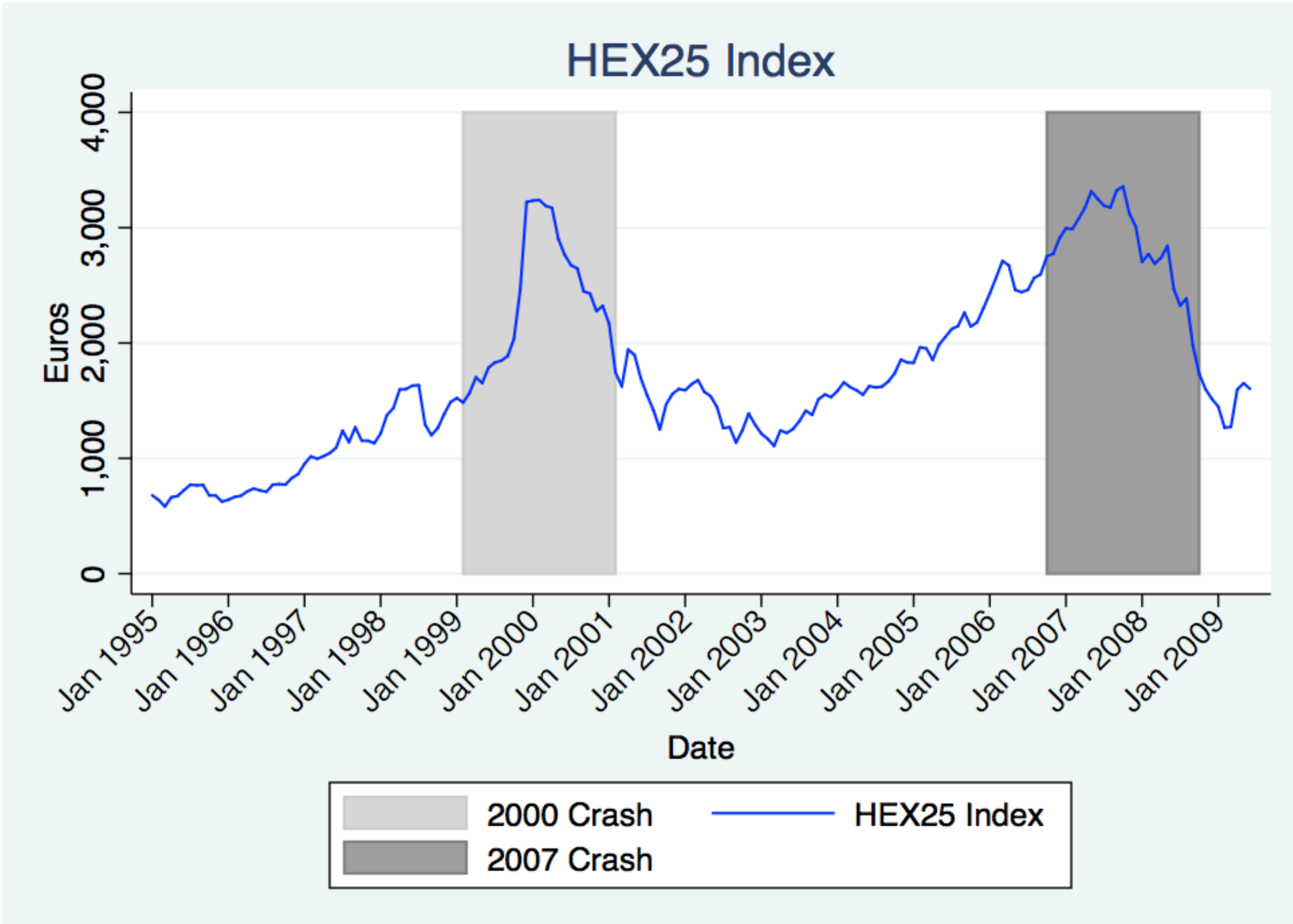

Figure 1: HEX25 Cumulative Returns

This figure displays the growth of the OMX Helsinki 25 index (formerly the HEX25 index). The HEX25 is a stock index of the 25 most traded shares on the NASDAQ OMS Helsinki exchange. The index is value weighted with a maximum weight on an individual security of 10 percent. We present the value of the OMX Helsinki 25 index for our sample period: January 1995 to June 2009. The shaded areas are the periods in which we examine performance around a market bubble.

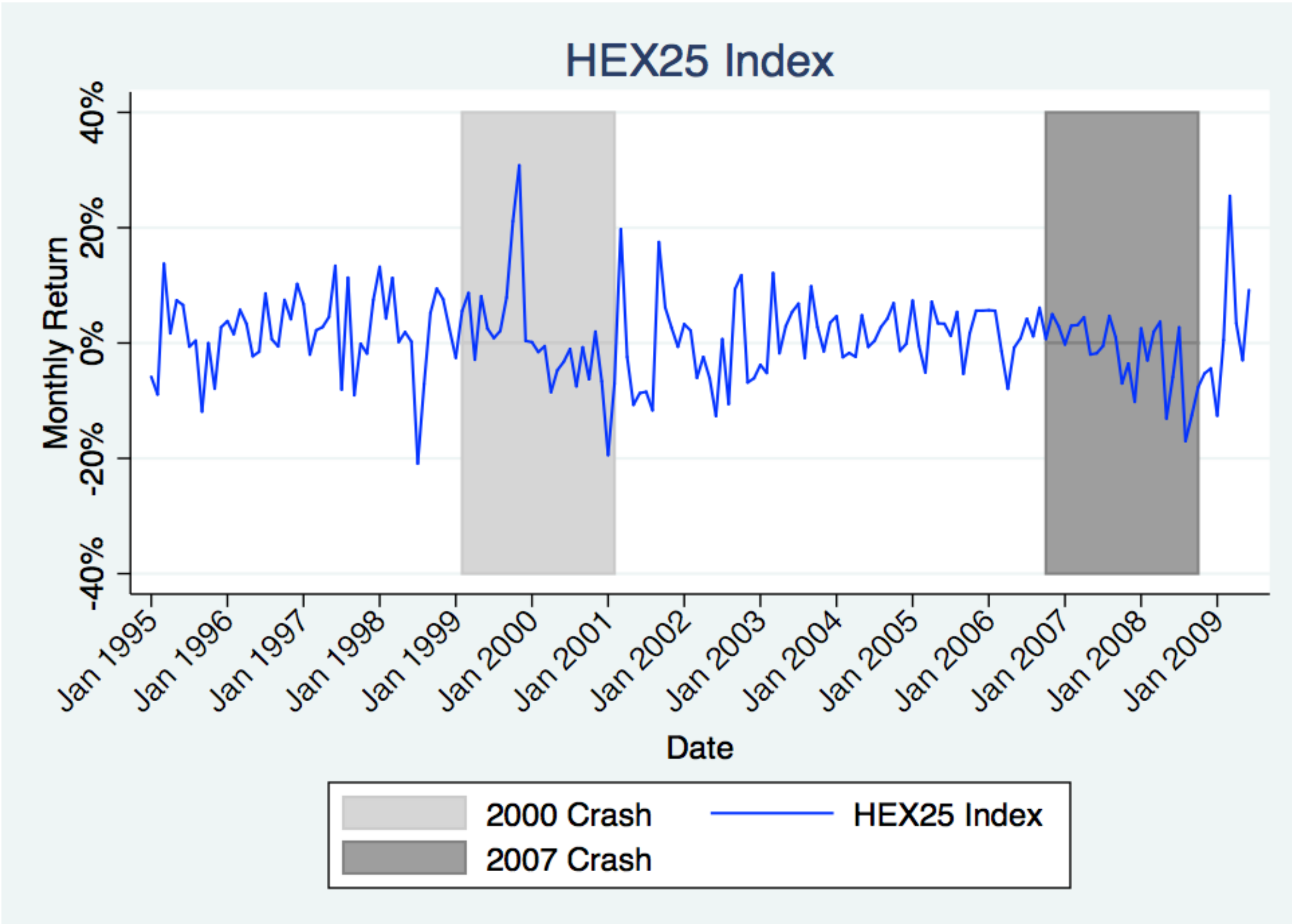

Figure 2: HEX25 Monthly Returns

This figure displays the monthly returns of the OMX Helsinki 25 index (formerly the HEX25 index). The HEX25 is a stock index of the 25 most traded shares on the NASDAQ OMS Helsinki exchange. The index is value weighted with a maximum weight on an individual security of 10 percent. We present the monthly returns of the OMX Helsinki 25 index for our sample period: January 1995 to June 2009. The shaded areas are the periods in which we examine performance around a market bubble.

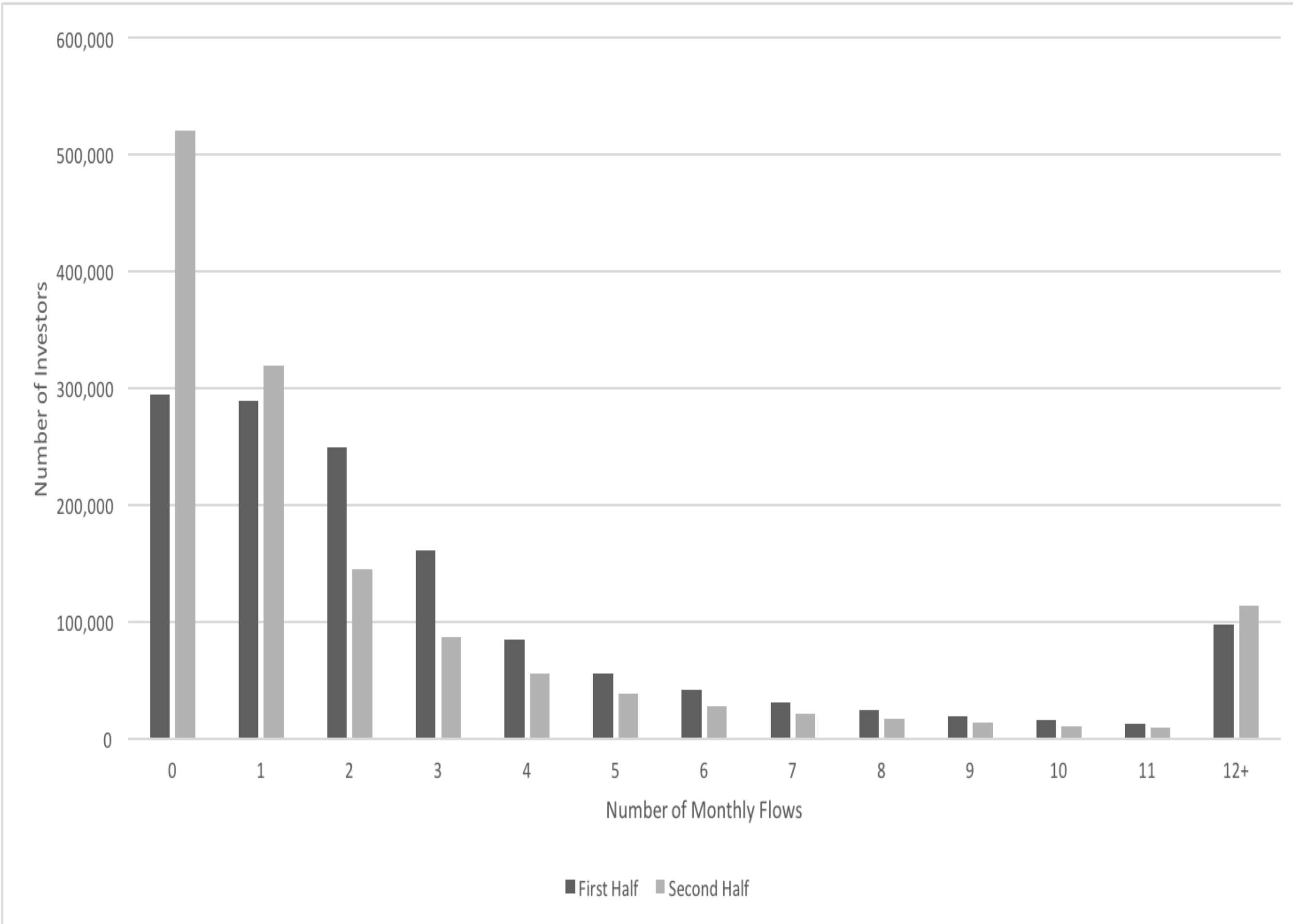

Figure 3: Histogram of Non-Zero Monthly Flows

This figure displays a histogram of the number of months of non-zero monthly flows per investor in the first and second half of the sample.