"Rookies to the Stock Market"

Martin Abrahamson*

Working paper, please do not quote or publish, without permission from the author
[Version, January 2015]

This study focuses on individual investors entering the stock market "rookies". A unique data set reflecting all individual investor holdings in Swedish stocks over the sample period 2004 to 2010 is used to examine portfolio preferences. Although the average shareholder is aging, the study shows signs of rejuvenation, since rookies are attracted to the stock market. The results show that the majority of the rookies choose one well-known company for their first stock market investment. The rookie characteristics show gender differences, where female rookies have lower income, are older but hold larger portfolios than their male equivalent.

* Abrahamson is associated with Uppsala University, Sweden.
Corresponding author: martin.abrahamson@fek.uu.se

I am grateful for financial support received from *Svenska Spel*. The author would like to express his appreciation for valuable comments of previous versions of this paper by Tom Berglund and Lawrence Kryzanowski (among others) at the 15th annual SNEE meeting in Mölle 2013, and also to Mika Vaihekoski, F.Y. Eric C. Lam, Gianluca Mattarocci, Vivek Singh, Radu Taru, Alexander Kerl, Viet Cao at the 2013 Merton H. Miller doctoral seminar at the 2013 EFMA in Reading, UK. I like to express a special thanks to Adri De Ridder and seminar participants at Uppsala University. The author also acknowledges computational support from Håkan Mattsson and Mikael Segerlund at Uppsala University.

1. Introduction

Davis Evans (2009) reports that the American retail investor is dying. This study focuses on individual investors in a country outside US to examine whether they follow their American equivalent on the road to extinction. Lease et al. (1974) claim, based on Klemkosky and Scott (1973), that individuals have been net sellers of shares in the US stock market from the year 1959. Schlarbaum et al. (1978) extend the time frame to 1978 for the US market, with the same pattern of individuals being net sellers. Rydqvist et al. (2013) show that ever since World War 2 there has been a shift from direct share holdings by households/individuals to holdings through institutions. For US, the development presented in Rydqvist et al. (2013) show that for individuals/households directly held shares went from 90 percent to a third. In order for individuals not to be dying as a phenomenon, some evidence of rejuvenation has to be found. I aim at giving a piece of the puzzle of who the new investors (henceforth rookies) are. Rookies are defined as individuals investing in the stock market for their first time. The study shows characteristics and preferences of the rookies during the latter years.

Individuals are in general considered to be less sophisticated, having several biases limiting their success in trading, as oppose to institutions. Prior studies have focused on the trading behavior of the individuals (e.g. Barber and Odean 2000, 2001, Goetzmann and Kumar 2008) but knowledge on what attracts the investors to the stock market and if there is any rejuvenation amongst the individual shareholders is limited. Rather, several studies show individual investors are declining and might have a diminishing future as investors on the stock market. Therefore, the prior studies related to this study are chosen from the literature on individual and/or household stockholdings and not solely on rookies. On the other hand Guo et al. (2007) state that all individual investors are rookies, at least those in the Chinese market.

Earlier studies typically contain data from a single US brokerage house, e.g. Lease et al. (1974), among others. Therefore, the clients of that particular brokerage house will affect the sample selection. Kumar and Lee (2006) present insights of individual investors based on a large sample. Although on a large sample of investors, it is limited to holdings of a single brokerage house. However, my study uses data of all shares owned by individuals in companies listed on Swedish public exchanges e.g. NasdaqOMXS, First North i.e. a study of a country rather than a brokerage house.

De Bondt (1998) argues for 2 main reasons for studying individuals actions in the stock market and not only their beliefs. Firstly, the investment of the individual affects the wellbeing. Secondly, with ever larger responsibility for your own pension, the future wellbeing of your retirement is also seriously affected. In Sweden, as in many countries, the responsibility of the pension savings has turned more towards the individual rather than the government over the last decades. Individuals in Sweden may be investing in different kind of financial products with different time horizons. However, this study focuses on rookie investors and their entry to the stock market regardless of the time horizon of the investment. Barber et al. (2009) show that individual investors trading is driven by active own decisions and not as a reaction to institutional trading. This supports the idea of studying individuals apart from institutions and not just the net buyer (or seller) during a bull (bear) market. The focus of this paper however is on directly owned shares by individuals that are new (rookies) to the stock market.

The study of rookies shows that, although it has for decades been reported that individual investors are declining their interests in direct holdings of shares in the stock market, new investors are still attracted to the stock market. The detailed data gives the opportunity to explore the characteristics of the individual investor and their holdings in new ways. The study explicitly target rookies and explore portfolio choices made by investors

moving into the stock market for the first time. To the best of my knowledge, this is the first paper to examine this issue in a country-based enumeration study of rookies.

The remainder of the paper proceeds as follows, the next section describes this study in relation to prior studies. Section 3 describes data and methodology, Section 4 presents the results and analyses the relationship between stock ownership and individual characteristics, Section 5 concludes.

2. Previous studies

Studies of individual investors are performed with various types of data. Data availability and precision are problems that seem to have troubled the researchers. The rightful owner of the share is one essential question that has been troublesome, since brokerage houses often have one account per household. One way of dealing with this issue is to conduct surveys to known households as in e.g. Lease et al. (1974). However, a survey study has several limitations e.g. to deal with the issues of non-respondents and selection bias. I study the rightful individual owner registered to the share rather than to speculate on to whom in the household the share rightfully belongs. However, tracing the overall household economy will be out of the question since the data have no record of the composition of the households. Studying the rightful owners and not households gives the opportunity to ask other questions, such as gender balance, age and income. Lease et al. (1974) find that 80-90 percent of the investment decision-makers are men. Similar gender balance can be drawn from Cohn et al. (1975), a result that might be interesting in comparison with this study.

Odean (1998) describes traders as overconfident and therefore carrying dead weight losses, since they trade too much. An effect of these overconfident traders is according to Odean (1998) that the market underreacts to relevant information and overreacts to irrelevant information.

Barber and Odean (2001) show that men overtrade more than women. The effect of overtrade is that men loose more money than women because of transaction costs. Barber and Odean (2000) also show that individual investors as a group underperform against relevant benchmarks. The worst underperformers are the one who traded the most, which is supported by Andersson (2013).

Barber and Odean (2001) claim that individuals expect their portfolio to beat the market. Although male investors have the highest expectations, both male and female investors expect to beat the market. In their dataset they proxy the gender of the investor by identifying the name of the person who opened the household´s account. Whenever they could identify the person opening the account as a man or woman, they use that as gender of the investor. In the study undertaken in this paper the actual holder of the stock instead of holder of household account is used, and through their identification number the gender is determined with precision. Kim and Nofsinger (2007) study Japanese individual investor behavior during bear and bull market conditions. My study also includes rise and fall in the market, hence the data is from before and during the financial crisis. Burnie and De Ridder (2009) use the same source of ownership data as presented in this study and use the rise and fall in the market surrounding the IT-bubble in their study of share ownership in Sweden. They focus on institutional ownership structure in Sweden over the years 2000 to 2002 and present a picture of the total ownership development. In their study there is some support of the displacement of individual shareholders in benefit of institutions, foreign and domestic. In a recent study Rydqvist et al. (2013) present evidence of a declining proportion of individual shareholders ever since the mid of the 20th century. Considering Burnie and De Ridder (2009) together with results presented in Rydqvist et al. (2013) would suggest that the pattern in Sweden is similar to US, when considering individual shareholders as diminishing owners of companies in favor of institutions.

Graham and Kumar (2006) describe characteristics of retail investors preferring nondividend stocks. However, the individual investors' preferences for dividend yield increase by age and income according to a study of 60, 000 American households. Individual investors are described as under diversified in several studies e.g. Blume and Fried (1975), Kelly (1995), Mitton and Vorkink (2007) and Goetzmann and Kumar (2008). If taken the 30+ stocks of Statman (1987) into account, as a cut-off for diversified stock portfolios, most individuals are severely under diversified. In fact only 0, 04 percent (98 out of 243, 866 individuals) of the rookie investors hold 30 shares or more in Sweden, 95 percent of the investors have 5 shares or less in their portfolio. Goetzmann and Kumar (2008) find evidence for investors diversifying deliberately, when holding more than one share. They report investors to be able to pick a passive diversification where more than one share is held, but not necessarily successful in picking a diversified portfolio.

2.1 Earlier studies of Scandinavian investors

The Scandinavian countries have developed economies with a stock market that functions accordingly. Scandinavian studies of individuals have an advantage to American studies when it comes to data availability. The construction of the social security number combined with the openness of the governmental authorities enables the researcher access information and to decode personal information. Prior research on Swedish data e.g. Massa and Simonov (2006) on similar data as this study (from the late 1990´s), show that investors earn strong returns on holdings closely related to them, either geographically or professionally. Massa and Simonov (2006) use a data set with a representative sample (3 percent) of the Swedish population the Longitudinal Individual Data for Sweden (*LINDA*) to describe income and holdings other than shares. The same database is used in more recent Lindqvist and Vestman (2011). In comparison, in this study I use data on the governmentally

reported income statement from all shareholders in Sweden. The results of Massa and Simonov (2006) show difference between high wealth investors and low wealth investors, where the high wealth investors is more diversified which could support results showing correlation between investor sophistication and wealth. Evidence for geographic home bias to be closely linked to individual investors have support in numerous studies e.g. Huberman (2001), Grinblatt and Keloharju (2001b), Dahlquist and Robertsson (2001) and more recent by Seasholes and Zhu (2010), Levy and Levy (2014).

Calvet et al. (2007) describe the probability of a Swedish household to hold shares is dependent on sophistication of the household, where households with low education and wealth are less likely to invest in shares. They also report that sophisticated Swedish households take on larger losses due to under diversification than less sophisticated households. That might partly be explained by their result that less sophisticated households rather invest in managed funds. In a more recent study Calvet et al. (2009a) use a Swedish database, in an earlier time-period, and show that less sophisticated households sell winners and keep losers. Whilst these studies focus on households I study individuals, which naturally occasionally is the same, but since I track the actual individuals' stockholdings, the level of detail is higher and linked to a unique subject/person.

Grinblatt and Keloharju (2001a) study why investors trade, based on Finnish data. They find amongst other things a gender effect supporting previous studies e.g. Barber and Odean (2001) on difference in trading patterns between men and women. However, the differences are not as large for selling but rather for buying shares. This is explained due to men being more active traders. Grinblatt and Keloharju (2001a) find that larger portfolios are positively correlated to more active trading behavior. They also find that the propensity to buy stocks is larger for men rather than women. Grinblatt and Keloharju (2001a) show that households to a larger extent than institutions have a contrarian investment behavior. Considering Grinblatt

and Keloharju (2001a), my study use similar data exploring the holdings of Swedish men and women before and during the current financial crisis. In a recent study Andersson (2013) study trading behavior of individuals based on data from a Swedish brokerage house. Andersson (2013) reports excessive trading amongst investors with lower income, wealth, age and education, which results in turn gives the traders weaker returns due to trading losses.

2.2 The Swedish context

In Sweden the employee has to take part of the responsibility of their pension savings. The pension savings in Sweden is generally considered to be divided into three parts where parts of the pension is salary based and paid from the employer and one part is provided by the government, apart from the pension savings of the individual. This pension system (*PPM*) was introduced at the turn of the century and replaced a system where the government took responsibility of the base-level pension of the Swedish citizens. Cronqvist and Thaler (2004) describe this new system to be associated with several limitations and constraints. Swedish pension funds have also been studied by e.g. Karlsson and Nordén (2007), who present a home bias among individual investors in their pension fund investments. My study will add to these studies by focusing on the directly owned shares of individual shareholders, with a recent time frame, surrounding the current financial crisis in order to study new/first time investors turning to the stock market.

Statistics Sweden and *Swedish Financial Supervisory Authority* (2013) report that foreign owners hold between 34-38 percent of all shares in the Swedish stock market during the sample period. More than 50 percent of the shares owned by foreigners are held by owners in the US and UK according to the same report. *Statistics Sweden* and *Swedish Financial Supervisory Authority* (2013) also report the Swedish households to own 13-16 percent of the shares for the sample period. However, according their report they do not consider whether

the foreign investor is an institution or an individual. The main reason for this might be that foreign individuals are believed to be a limited part of the foreign investors, as the institutions undisputedly hold the majority of the foreign investments in Sweden. *Statistics Sweden* study individual holdings for which the number of shares in the studied company exceeds 500 shares, without considering the portfolio value. However, in this study all individual holdings are included, regardless of the number of shares in each company.

The tax situation for individuals owning and trading shares in Sweden has been simplified compared to e.g. US. Regardless of their income, individuals in Sweden pay the same fixed tax rate (30 percent) both for dividend yield and capital gain. Therefore, in that regard individuals should be indifferent to obtaining either dividend yield or capital gain. However, shortly after the sample period (January 1st 2012) individuals are given the opportunity to invest through an investment account where the owner of the account pay a percentage tax of the current market value of the account. The possible effects of the specific investor account savings are not taken into account in this paper, but should be an interesting issue to study in a few years.

In summary, previous studies are mainly focused on US investors based on survey or brokerage house data. Even though individuals are seen as less sophisticated shareholders as oppose to institutions, there are differences amongst the individuals. Differences have been connected with individual investor characteristics e.g. experience, diversification, age, income, gender, education and residence. However, the characteristics have not been tested with a population of rookie investors and their stock holdings. In this study I test a regression model explaining portfolio value based on the income, age, gender, diversification, share price level and residence. Based on prior studies, I expect income, diversification, age, residence in the capital and share price level to indicate sophistication and higher portfolio value.

3. Data and methodology

The availability of data makes it especially interesting to study the Swedish stock market from a shareholder perspective. In contrast to many other countries, there is by law only one body which monitors stock ownership in all public firms, the Central Security Registrar (*Euroclear Sweden).* This study obtains the ultimate ownership data for all listed firms, from *Euroclear Sweden.* The data is compiled as of the end of each year.

In Sweden every citizen has an identification number, either given by birth or when the residence permit is issued. Sweden is one of few countries where the information in the identification number reveals personal information, e.g. when you were born and gender.[1]

Each individual investor is tracked through the identification number enabling examination of total stock holdings across all firms as well as age, sex, place of birth and current residence. For this study the ownership database is combined with annual income[2], for all individuals within the sample, obtained from the Swedish Tax authorities, *Skatteverket.*

Foreign individual investors constitute 0.8 percent of the population. For foreign individual investors the data in this study does not provide information on income, age and gender. Therefore, foreign investors are excluded from the study. Native Swedes living (or purchasing from) abroad still enables data on the independent variables and therefore included.

3.1 Rookies

Rookies are defined as new individual investors being active on the stock market for their first time. I select the sample of rookies out of the total population of two million

[1] For citizens born between January 1947 and January 1990, it also contains information of the birth county.

[2] The annual frequency of the data for this study disables analysis of the activity around the turn of the year, as previously studied by e.g. Ritter (1988) and Rozeff and Kinney (1976). Neither can it reveal the trading activity between the reporting periods, although the advantages are in investor characteristics and the precision in the actual holdings, where previous research has been limited.

(1, 942, 523) shareholders by eliminating all individuals owning any stock during the years 1999-2003. The shareholders holding a share after January 1st 2004 and had not held a share prior are identified as "*rookies"*. Individual investors are defined as shareholders, with personal identification number, holding the stock in their name, not through corporations. Shareholders that previously held shares and later re-invested in the stock market are not considered rookies. Considering the evenly distributed number of rookies over the sample period, I believe that the methodology of excluding previous shareholders to detect the rookie is sufficient. The sample of rookie shareholders consists of 241 893 investors. Due to data limitations on annual income and the characteristics of foreign investors the sample decreased to 228 694 in the final sample. Overall the rookies correspond to 12 percent of the total number of shareholders in Sweden during the sample period.

The portfolio value is compiled by the year-end of the investors' rookie year. I use the closing price and holdings on the last day of trade to set the value of each share in the portfolio. The value of all holdings is compiled into the portfolio value for each investor. To mitigate the impact of outliers in the sample, the portfolio value and income are winsorized at 1 percent level.

In oppose to prior studies age is defined as the age of the investor and not numbers of years the account has been open, nor the oldest person within the household. Therefore, the portfolio values are comparable and the rookies can be expected to have the same possibility of holding a diversified portfolio regardless of their age.

3.3 Multivariate analysis

In examining the relationship between the portfolio value and characteristics related to the individual a multivariate analysis is used. I estimate following model, to explain the portfolio value of the rookie investor.

$$\text{Portfolio Value}_{i,t} = \alpha + \beta_1 \text{Income}_{i,t} + \beta_2 \text{Age}_{i,t} + \beta_3 \text{Gender}_i + \beta_4 \text{One share}_{i,t} + \beta_5 \text{High Price}_{i,t} + \beta_6 \text{Capital}_{i,t} + \varepsilon_{i,t} \quad (1)$$

where:

Income =annual income for each individual

Age = age for each individual

Gender = is a dummy variable with the value of 1 for male and 0 otherwise

One Share = is a dummy variable with the value of 1 if the individual hold one share/ company and 0 otherwise

High Price = is a dummy variable with the value of 1 if the average stock price in the portfolio is higher than the average stock price on the Stockholm Stock Exchange and 0 otherwise

Capital = is a dummy variable with the value of 1 if the individual resides in the capital of Sweden (the capital, Stockholm-area, based on zip codes) and 0 otherwise.

Income, age, high price and capital are used to indicate a sophisticated rookie and are expected to be positively correlated to portfolio value. Since male investors have higher income they are expected to hold larger portfolios. The one share dummy variable show lack of diversification and is expected to be negatively correlated to portfolio value[3].

[3] 62% of the rookies hold only one share in their portfolio.

3.4 Portfolio analysis

A logistic regression analysis is used in order to examine the characteristics of rookies who hold one share vis-à-vis rookies who hold more than one share in their portfolio. The one share dummy is used as dependent variable. I use income, portfolio value, age, residence, gender and share price as independent variables. The natural log of income and portfolio value are used, both are reported in SEK (Swedish krona).

$$\text{One share}_{i,t} = \alpha + \beta_1 \text{Income}_{i,t} + \beta_2 \text{Portfolio value}_{i,t} + \beta_3 \text{Age}_{i,t} + \beta_4 \text{Gender}_i + \beta_5 \text{Capital}_{i,t} + \beta_6 \text{High Price}_{i,t} + \varepsilon_{i,t} \quad (2)$$

As alternative to the high price dummy I introduce 1/price as new independent variable, which is one over the average price of a share in the rookie portfolio.

4. Results

Table 1 presents summary statistics. As reported in Panel A, across all sample years the average number of rookies is 34 559 per annum. The mean (median) age of the rookie investor is 37.3 (36) years[4] for the sample period.[5]

The last two columns in Panel A report the fraction of investors grouped by gender. There are more male than female investors in the sample, with a gender balance of 57 percent male and 43 percent female investors. This is consistent with results reported from Finnish studies e.g. Grinblatt and Keloharju (2000, 2001a). Differences compared to US studies e.g. Barber and Odean (2001), Graham and Kumar (2006) can be explained by e.g. different definitions of the owner. The studies from the Nordic countries show the rightful owner of the stock and not a proxy e.g. for household account signature. The differences in the results

[4] Despite rookies entering the stock market, the mean (median) age of the all shareholders in Sweden (not reported in the table) went from 53(54) years to 55(57) years over the sample period.

[5] 11% of the rookies are older than 65, which is the typical retirement period. This shows that some investors going in the opposite direction of realizing your savings, when retired. The insight that even retired can be rookies might be explained by the fact that people in Sweden are expected to live far beyond the retirement age or as simple as a new interest in the stock market after the end of a time-consuming work life.

between these European studies and the previous US studies stress the need for studies of actual holdings rather than proxies for households. Comparing the number of rookies entering the stock market based on gender there seems to be a rather balanced sample. Although, the largest difference between male and female rookies is in 2008/2009 when the financial crisis was ongoing and the stock markets were falling. This would rather suggest that male investors, to a larger extent, try to time the market when stock prices are falling rather than the market behaving bullish as previously shown in Finland[6].

Panel B reports statistics of four variables related to stock ownership: portfolio value, income, age and number of stocks in portfolio. As reported, portfolio value as well as income is positively skewed. The number of companies in an investor's portfolio is on average two for the sample of rookies[7], this is consistent with Barber and Odean (2001). However, the median portfolio holds only one share, consistent with results presented by Kelly (1995), that reported median holdings of US investors is just one stock, using survey data. The rookies in Sweden are even less diversified in their stock portfolios than US investors[8], although the Swedish data presented in this study does not enable diversifications as holdings in e.g. mutual funds to be considered.[9] The under diversification reported in US studies holds for rookies, but also for the entire population of Swedish individual shareholders.

In Panel C, descriptive statistics of investors based on their current residence is shown. Out of the quarter of a million rookies only 5 percent live in the capital city, compared to the overall individual shareholders in Sweden where 35 percent of the total population of individual shareholders live in the capital. The deviation from the existing shareholders is

[6] A *t*-test is used to test the difference in proportion of male rookies during the financial crisis (2008, 2009) to the overall sample, the test shows significant results of larger proportion of male rookies during the 2008-2009 period.
[7] 3 shares for the total population (not reported in the table).
[8] In the rookie sample 95% of the investors hold shares in 5 or less companies. This can be explained by unexperienced investors but also by e.g. transaction costs tempting the rookie to lower the number of transactions.
[9] Kumar and Lee (2006) show that the US investors have a concentrated stock portfolio with a mean (median) of 4 (3) shares. Dividing my sample with respect to income, this corresponds to the wealthiest Swedish investors.

significant and unexpected[10]. Investors living in the capital have higher income as expected. Portfolio value and number of shares are also higher for residence in the capital. The rookies in the capital are younger than the rookies living in other parts of Sweden.

4.2 Correlations of portfolio value and rookie characteristics

Table 2, presents a correlation matrix of tested variables. As expected, I find positive correlation between portfolio value and income (0.1299). Portfolio value is also positively correlated to number of stocks in portfolio (0.3947), age (0.1892), high price (0.1202) and residence of the capital (0.0459). As reported income show positive correlation to number of stocks in portfolio (0.0505), age (0.1830) and gender (0.0559), where men on average earn more than women.

The positive relationship between age and income is not surprising, showing that income rises with age. The positive relationships between portfolio value and age as well as between age and number of shares (diversification), supports age as being an indicator of experience and possibly also sophistication.[11] Goetzmann and Kumar (2008) show a positive relationship between age, income and diversification of the portfolio. My study supports their results but also show positive relationship of actual age of the investor and portfolio value. The positive correlation reported in Table 2, between income and passive diversification i.e. number of stocks in the portfolio support Massa and Simonov (2006) and Calvet et al. (2009b). My results show that the positive correlation between income and passive diversification also holds for rookies. However, even the wealthiest rookies hold severely under diversified portfolios.

[10] Even though the study includes a few recent years there is a remarkable difference between the rookies and overall population of shareholders. This needs to be addressed further in future research.
[11] However, the positive relationship between age and diversification is weak, as presented in table 2. Therefore, I divided the investors into 10 age groups, but investors show the same median number of shares (diversification) for all age groups.

4.5 Gender

Table 3 reports the univariate analysis of my variables sorted by gender of the rookie. The mean difference is examined using a t-test whereas a Wilcoxon rank sum test is used to examine the difference in medians. The mean annual income is higher for male rookies compared to female, the mean difference is highly statistically significant (*t*-statistics of -57.23). Unexpectedly, the female investors hold larger portfolios, despite the fact that their income is smaller. This holds for both mean and median difference (mean difference with *t*-statistics = 19.42). Prior results of individual investors hold undiversified portfolios are also evident in my sample. Specifically, the third row shows that on average female investors hold slightly more diversified portfolios than their male equivalent (2.17 compared to 2.10 respectively, with *t*-statistics = 6.84). The average price per share is higher for female investors than male, the mean difference is highly significant (*t*-statistics = 45.37). The average female rookie is older than her male equivalent (41 and 34 years respectively), the seven year difference has a *t*-statistics = 80.5. In summary, the result presented shows that female investors are older, hold larger and more diversified portfolios than their male equivalent, although the male rookie have larger income.

The gender differences reported might be explained partly with Goetzmann and Kumar (2008), since the female investors are older when they are rookies, this can explain why their portfolio value is greater. However, although the female rookies are older and have larger portfolios, their income is smaller compared to the male rookie. Taken together the relationship between portfolio value and income is affected by differences amongst the average income between male and female investors. The gender difference in portfolio value might partly be explained by overtrading as in Barber and Odean (2001), but since I only take the holdings the first year into consideration, this might not tell the whole story. Male investors are slightly overrepresented in the sample and can indicate that the threshold for

entering the market is larger for female investors than for male investors. Age can also explain some of the gender difference, since age is positively correlated with portfolio value, in my study as well as in previous studies. Although, age is normally positively correlated with income at least until retirement, but in the sample the female rookies have smaller income than the male. Taken together this would indicate that, when female rookies invest into the stock market for the first time they do so with a larger proportion of their savings than their male equivalent.

4.6 OLS-Regression

Table 4 presents the results of a multiple regression. The portfolio value is the dependent variable. The following investor characteristics are explanatory variables: income and the investor´s age. I also use dummy variables for: gender, one share in the portfolio, high price and residence of the Capital. The portfolio value and income are winsorized at 1 percent. All the explanatory variables are significant. Income, age, high price and capital are positively related to portfolio value. The gender dummy variable is negative, showing female portfolios being larger than male portfolios. This is consistent with the correlation matrix and the univariate analysis, where the gender variable is negative. As expected and consistent with previous research, the one share dummy is negative, showing that rookies with only one share hold smaller portfolio values than rookies with more diversified portfolios. The overall model explains portfolio value of the rookie investor based on characteristics of the investor (Adjusted R^2 =10, 4)[12].

[12] I also used year dummies and a financial crisis dummy (year 2008, 2009) in the regression, without major impact on the result.

4.7 *Rookie portfolio selection*

Table 5 presents results of the logistic regression, where the one share dummy is the dependent variable. I use income, portfolio value, age and along with dummy variables for residence, gender, high price. I introduce one over price in order to study nominal price effect, as an alternative to the high price dummy. Consequently I estimate two regressions, the first with one over price and the second using the high price dummy. The result of the first regression is presented in columns 2 and 3, and the second in columns 4 and 5 of table 5. Of the variables considered only one over price has no significant effect on the probability of holding only one stock in the rookie portfolio i.e. undiversified. However, the high price dummy is positive and significant. Income and portfolio value are negative and significant, which supports prior studies, where income and portfolio value can be seen as indicators of sophistication and thereby more diversified investor portfolios. Unexpectedly age is slightly positive indicating that older rookies are more likely to hold only one share. Rookies living in the capital are less likely to hold only one share, although the distribution of residence in the capital is unbalanced. Female rookies are less likely to hold only one share, showing that they are more diversified.

I analyze the share selections of the rookie holdings, in order to study what attracts the investors to the stock market. All portfolios are sorted in value order. The companies are ranked after first choice, second choice etc. until the fifth largest holding of each investor (95 percent of all holdings). I created a largest 35 dummy variable, where I calculated the firm value, ranked the firms and selected the largest 35 each year. In the rookie sample more than 99 percent held at least one share out of the largest 35 companies, i.e. all rookies held companies with the largest firm value of the stock market. Thereby, it seems like rookie investors buy shares that is well known to them (based on familiarity). The majority of the

rookies (62 percent not reported in table 1) put all the eggs in one basket and takes on all the unique risk without any diversification, since they only invest in one company.

5. Concluding remarks

I analyze stock ownership for a sample of new (*rookie*) shareholders over the period 2004 to 2010. By exploiting a unique data set of all stock holdings in Sweden, I make original contributions to the existing literature related to individual holdings as the data is sharper and more detailed than what have been used in prior studies. I find support in the Swedish data for US prophesies of individual investors having an ailing trend as shareholders. However, the stock market still attracts a substantial amount of new shareholders (in total 12 percent of all individuals during the sample period). This is evidence of rejuvenation amongst individual shareholders as part of the stock market.

A model to explain portfolio value is presented, where income, age, shares with high nominal price and residence are positively related to portfolio value. The rookie investors in Sweden are less diversified in their stock portfolios than experienced Swedish and US investors. The majority of the rookies choose one large and well known company to invest in for their first investment. The study shows that for rookies the female investor on average hold larger and more diversified portfolios than their male equivalent.

Table 1 Summary statistics

This table presents summary statistics for new individual investors ("rookies") with a long position in a stock listed on the Stockholm Stock Exchange in Sweden over the period 2004 through 2010. Panel A, reports total number of rookies sorted by age and gender for each calendar year. Panel B, reports the sample distribution of the variables used in the paper where portfolio value is defined as the total value of the portfolio for each investor. Portfolio value is calculated at the end of December for each calendar year. Data on stock ownership is obtained from the central security registrar in Sweden (*Euroclear Sweden*). Income is the annual income for each investor where data is obtained from the Swedish Tax Authorities (*Skatteverket*). To reduce the impact of outliers, portfolio value and income have been winsorized at the 1% level. Age is the age of the investor. Number of stocks in portfolio is the number of different firms in the portfolio. Panel C, shows current residence of the rookie with respect to living in capital (i.e. close to the Stockholm Stock Exchange) or not. Current residence is based on the Zip code of the address where the investor is registered. Number of observations are 241 893, but for income the sample decreases to 228 694 due to data limitations. For all panels portfolio value and income are presented in local currency SEK, *Swedish Krona* (SEK 7 corresponds to $1).

Panel A: Rookies in Sweden					
Year	Number of Rookies	Age		Proportion	
		Mean	Median	Men	Women
2004	33560	37.1	37	0.55	0.45
2005	28242	37.2	38	0.53	0.47
2006	31699	37.4	37	0.55	0.45
2007	37092	36.6	35	0.58	0.42
2008	38883	37.9	37	0.60	0.40
2009	38957	38.3	37	0.61	0.39
2010	33460	36.9	35	0.59	0.41
Average	34 556	37.3		0.57	0.43

Panel B: Sample distribution					
Variable	Mean	Std.dev.	25%	Median	75%
Portfolio value	40 994	114 737	3 285	10 472	32 319
Income	211 054	196 364	27 309	202 648	308 918
Age	37.3	21.0	23	36	53
Number of stocks	2.1	2.4	1	1	2

Panel C: Residence					
	Proportion of Rookies	Number of shares Mean (Median)	Age Mean (Median)	Portfolio Value Mean (Median)	Income Mean (Median)
Capital of Sweden	0.05	2.5 (1)	32 (31)	66 759 (14 036)	240 225 (211 816)
Other parts of Sweden	0.95	2.1 (1)	37 (37)	44 726 (10 339)	209 232 (202 348)

Table 2 Correlation matrix

The table reports correlation between studied variables. Portfolio value is defined as the total value of the portfolio for each investor. Portfolio value is calculated at the end of December for each calendar year. Data on stock ownership is obtained from the central security registrar in Sweden (*Euroclear Sweden*). Income is the annual income for each investor where data is obtained from the Swedish Tax Authorities (*Skatteverket*). To reduce the impact of outliers, portfolio value and income have been winsorized at the 1% level. Age is the age of the investor. Number of stocks in portfolio is the number of different firms in the portfolio. Gender is a dummy variable with 0 for female and 1 for male investor. High Price is a dummy where 1 refers to the investor average share price in the portfolio being higher than the average share price for the sample period. Capital is a dummy variable, coded 1 if the investor currently lives in the capital (i.e. close to the Stockholm Stock Exchange) and 0 if not. Number of observations is 228 694.

Variable	Portfolio Value	Income	Number of stocks in portfolio	Age	Gender	High Price	Capital
Portfolio Value	1.0000						
Income	0.1299	1.0000					
Number of stocks in portfolio	0.3947	0.0505	1.0000				
Age	0.1892	0.1830	0.0838	1.0000			
GENDER	-0.0374	0.0559	-0.0124	-0.1637	1.0000		
HIGH PRICE	0.1202	-0.0121	0.0566	0.0494	-0.0852	1.0000	
CAPITAL	0.0459	0.0362	0.0302	-0.0520	-0.0101	0.0228	1.0000

Table 3 Gender differences

The table reports results from univariate analysis on characteristics based on the gender of the rookie. Portfolio value is defined as the total value of the portfolio for each investor. Portfolio value is calculated at the end of December for each calendar year. Data on stock ownership is obtained from the central security registrar in Sweden (*Euroclear Sweden*). Income is the annual income for each investor where data is obtained from the Swedish Tax Authorities (*Skatteverket*). To reduce the impact of outliers, portfolio value and income have been winsorized at the 1% level. Number of stocks in portfolio is the number of different firms in the portfolio. Avg. Price/share is the average price of all shares in the investors´ portfolio, calculated as portfolio value over total number of shares across all firms. Age is the age of the investor. Difference in mean test is a t-test allowing unequal variance. Difference in median test is a Wilcoxon rank sum test. Number of observations is 241 893 (but for Income N=228 694 due to data limitations). The value is presented in local currency, *Swedish Krona*, SEK (SEK 7 corresponds to $1).

	Female		Male		Difference	Difference mean test	Difference Median test
Variable	Mean	Median	Mean	Median	Mean	*t*-stat	*p*-value
Income (SEK)	184 829	173 837	230 489	231 747	-45 660	-57.23	0.00
Portfolio Value (SEK)	50 616	11 350	41 376	9 540	9240	19.42	0.00
Number of stocks in portfolio	2.17	1	2.10	1	0.07	6.84	0.25
Average Price/share (SEK)	88.42	61.88	73.39	53.25	15.03	45.37	0.00
Age	41	43	34	33	7	80.54	0.00

Table 4 OLS-Regression results

The table reports results from regression of portfolio value dependent on income, age, gender residence. Portfolio value is defined as the total value of the portfolio for each investor. Portfolio value is calculated at the end of December for each calendar year. Data on stock ownership is obtained from the central security registrar in Sweden (*Euroclear Sweden*). Income is the annual income for each investor where data is obtained from the Swedish Tax Authorities (*Skatteverket*). To reduce the impact of outliers, portfolio value and income have been winsorized at the 1% level. Age is the age of the investor. Gender is a dummy variable with 1 for male investor and 0 for female. One Share is a dummy coded as 1 if the rookie only holds shares in only one firm, and coded 0 if shares are held in multiple firms. High Price is a dummy where 1 refers to the investor average share price in the portfolio being higher than the average share price for the sample period and 0 otherwise. Capital is a dummy variable coded 1 if the investor lives in the capital (i.e. close to the Stockholm Stock Exchange) and 0 if not. *** denote significance at the 1 % level. The *t*-statistics for the coefficient estimates are reported in parentheses.

	Portfolio Value					
Regression Model	(1)	(2)	(3)	(4)	(5)	(6)
Variable						
Constant	29 484***	4710***	7704***	44 605***	35 188***	33 369***
	(85.28)	(9.76)	(13.15)	(68.29)	(52.19)	(49.12)
Income	0.075***	0.036***	0.039***	0.036***	0.038***	0.035***
	(62.64)	(21.96)	(29.22)	(27.99)	(29.24)	(27.79)
Age		883***	857***	782***	757***	778***
		(72.68)	(68.60)	(64.31)	(62.61)	(64.13)
GENDER			-4398***	-5111***	-3317***	-3015***
			(-9.02)	(-10.79)	(-7.02)	(-6.39)
ONE SHARE				-53 782***	-52 803***	-52 451***
				(-115.28)	(-113.75)	(-113.02)
HIGH PRICE					24 805***	24 570***
					(51.90)	(51.44)
CAPITAL						20 162***
						(20.53)
Number of rookies	228 694	228 694	228 694	228 694	228 694	228 694
Adjusted R^2	0.017	0.039	0.040	0.092	0.103	0.104

Table 5 Logistic-Regression results

The table reports results from the logistic regression of the One share dummy take on the value of 1 if the investor holds only one share and 0 otherwise. Income is income of the investor for the calendar year measured in thousand SEK. The natural log of income and the portfolio value of the investor are used, measured in SEK. Portfolio value is calculated at the end of December for each calendar year. Age is the age of the investor measured in years. Gender, Capital, High price are dummy variables that take on the value of 1 for male investor, living in the capital (i.e. close to the Stockholm Stock Exchange) and average share price in investor portfolio being larger than average share price on the stock market and 0 otherwise. 1/price is calculated as 1 over the average price of a share in the investors´ portfolio, measured in SEK. Data on stock ownership is obtained from the central security registrar in Sweden (*Euroclear Sweden*). Income is the annual income for each investor where data is obtained from the Swedish Tax Authorities (*Skatteverket*). To reduce the impact of outliers, portfolio value and income have been winsorized at the 1% level. **, *** denote significance at the 5% and 1 % level respectively, using the Wald-test.

Variable	Coefficient	*P*	Coefficient	*p*
Intercept	4.771	<0.001***	4.774	<0.001***
Income (log)	-0.020	<0.001***	-0.020	<0.001***
Portfolio value (log)	-0.460	<0.001***	-0.464	<0.001***
Age	0.005	<0.001***	0.004	<0.001***
GENDER	-0.103	<0.001***	-0.098	<0.001***
CAPITAL	-0.165	<0.001***	-0.167	<0.001***
1/price	-0.004	0.031		
HIGH PRICE			0.088	<0.001***
Number of rookies	228 694			228 694
Pseudo R^2	0.0970			0.0972

References

Andersson, A., 2013. Trading and Under-Diversification. *Review of Finance,* Vol. 17, pp. 1699-1741.

Barber, M. B., T. Odean, 2000. Trading is Hazardous to Your Wealth: Common Stock Investment Performance of Individual Investors. *Journal of Finance,* Vol. 55, No. 2, pp. 773-806.

Barber, M. B., T. Odean, 2001. Boys will be Boys: Gender, Overconfidence and Common Stock Investment. *Quarterly Journal of Economics,* Vol. 116, No. 1, pp. 261-292.

Barber, M. B., T. Odean, and N. Zhu, 2009. Systematic Noise. *Journal of Financial Markets,* 2009, No. 12, pp. 547-569.

Blume, M.E., I. Fried, 1975. The Asset Structure of Individual Portfolios and Some Implications for Utility Functions. *Journal of Finance,* Vol. 30, pp. 585-603.

Burnie, D., A. De Ridder, 2009. Bear Market Behavior of Institutional Investors in Sweden. *American Journal of Business,* Vol. 24, No. 1, pp. 33-45.

Calvet, L. E., J. Y. Campbell, and P. Sodini, 2007. Down or Out: Assessing the Welfare Costs of Household Investment Mistakes. *Journal of Political Economy,* Vol. 115 No. 5, pp. 707-747.

Calvet, L. E., J. Y. Campbell, and P. Sodini, 2009a. Fight or Flight? Portfolio Rebalancing by Individual Investors. *Quarterly Journal of Economics,* Vol. 124 No. 1, pp. 301-348.

Calvet, L. E., J. Y. Campbell, and P. Sodini, 2009b. Measuring the Financial Sophistication of Households. *American Economic Review,* Vol. 99, pp. 393-398.

Cohn, R. A., W. G. Lewellen, R. C. Lease, and G. G. Schlarbaum, 1975. Individual Investor Risk Aversion and Investment Portfolio Composition. *Journal of Finance,* Vol. 30, pp. 605-620.

Cronqvist, H., R. H. Thaler, 2004. Design Choices in Privatized Social-Security Systems: Learning from the Swedish Experience. *American Economic Review,* Vol. 94, No. 2, pp. 424-428.

Dahlquist, M., G. Robertsson, 2001. Direct foreign ownership, institutional investors and firm characteristics. *Journal of Financial Economics,* Vol. 59, pp. 413-440.

Davis Evans, A., 2009. A Requiem for the Individual Investor? *Virginia Law Review,* Vol. 95 No. 4, pp. 1105-1129.

De Bondt, W. F. M., 1998. A Portrait of the Individual Investor. *European Economic Review,* Vol. 42, pp. 831-844.

Goetzmann, W. N., A. Kumar, 2008. Equity Portfolio Diversification. *Review of Finance,* Vol. 12, pp. 433-643.

Graham, J. R., A. Kumar, 2006. Do dividend Clienteles Exist? Evidence on Dividend Preferences of Retail Investors. *Journal of Finance,* Vol. 61, No. 3, pp. 1305-1336.

Grinblatt, M., M. Keloharju, 2000. The Investment Behavior and Performance of Various Investor Types: a Study of Finland´s Unique Data Set. *Journal of Financial Economics,* Vol. 55, pp. 43-67.

Grinblatt, M., M. Keloharju, 2001a. What Makes Investors Trade? *Journal of Finance,* Vol. 56, No. 2, pp. 589-616.

Grinblatt, M., M. Keloharju, 2001b. How Distance, Language and Culture Influence Stockholdings and Trades. *Journal of Finance,* Vol. 56, No. 3 pp. 1053-1073.

Guo, C., Y. Lianggyuan and K. Changwen. 2007. Understanding the Chinese Stock Market. *Journal of Corporate Accounting & Finance,* September/October, pp. 13-20.

Huberman, G., 2001. Familiarity Breeds Investment. *Review of Financial Studies,* Vol. 14, No. 3 pp. 659-680.

Karlsson, A., L. Nordén, 2007. Home Sweet Home: Home Bias and International Diversification among Individual Investors. *Journal of Banking & Finance,* Vol. 31, pp. 317-333.

Kelly, M., (1995). All the Eggs in One Basket: Portfolio Diversification of US Households. *Journal of Economic Behavior & Organization*, Vol. 27, No. 1, pp. 87-96.

Kim, K. A., J. R. Nofsinger, (2007). The Behavior of Japanese Individual Investors During Bull and Bear Markets. *Journal of Behavioral Finance*, Vol. 8, No. 3, pp. 138-153.

Klemkosky, R. C., D. F. Scott, 1973. Withdrawal of the Individual Investor from the Equity Markets. *Michigan State University Business Topics*, Spring 1973.

Kumar, A., C. M. C. Lee, 2006. Retail Investor Sentiment and Comovements. *Journal of Finance,* Vol. 61, No. 5, 2451-2486.

Lease, R. C., W. G. Wilbur, and G. G. Schlarbaum, 1974. The Individual Investor: Attributes and Attitudes. *Journal of Finance,* Vol. 29, No. 2, pp. 413-433.

Levy, H., M. Levy, 2014. The Home Bias is Here to Stay. *Journal of Banking and Finance,* forthcoming.

Lindqvist, E., R. Vestman, 2011. The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from Swedish Enlistment. *American Economic Journal: Applied Economics,* Vol. 3, pp. 101-128.

Massa, M., A. Simonov, 2006. Hedging, Familiarity and Portfolio Choice. *Review of Financial Studies,* Vol. 19, No. 2, pp. 633-685.

Mitton, T., K. Vorkink, 2007. Equilibrium Underdiversification and the Preference for Skewness. *Review of Financial Studies,* Vol. 20, No. 4, pp. 1255-1288.

Odean, T., 1998. Volume, Volatility, Price and Profit When All Traders Are Above Average. *Journal of Finance,* Vol. 53, No. 6, pp. 1887-1934.

Ritter, J. R., 1988. The Buying and Selling Behavior of Individual Investors at the Turn of the Year. *Journal of Finance,* Vol. 43, No. 3, pp.701-717.

Rozeff, M. S., W. R. Kinney, 1976. Capital Markets Seasonality: The Case of Stock Returns. *Journal of Financial Economics,* Vol. 3, pp. 379-402.

Rydqvist, K., J. Spizman, and I. Strebulaev, 2014. Government Policy and Ownership of Equity Securities. *Journal of Financial Economics,* Vol. 111, pp. 70-85.

Schlarbaum, G. G., W. G. Lewellen and R. C. Lease, 1978. Realized Returns on Common Stock Investments: The Experience of Individual Investors. *Journal of Business,* Vol. 51, No. 2, pp. 299-325.

Seasholes, M. S., N. Zhu, 2010. Individual Investors and Local Bias. *The Journal of Finance,* Vol. 56, No. 5, pp. 1987-2010.

Statman, M., 1987. How Many Stocks Make a Diversified Portfolio? *Journal of Financial and Quantitative Analysis,* Vol. 22, No. 3, pp. 353-363.