Investment Decisions and Risk Preferences among Non-Professional Investors

Anders Karlsson

Stockholm University

ISBN 978-91-7155-426-0

Cover by Anders Karlsson, insprired by René Magritte “The Son of Man”
1964
Printed in Sweden by Intellecta AB, Sweden 2007
Distributor: Stockholm University School of Business

- What are you doing?
- I don’t know…

Viktor (3 years old) 2007

Acknowledgements

I entered the PhD program at Stockholm University School of Business confident of my knowledge in corporate finance and financial economics. I had earned my masters from the same university and had worked a season as a financial analyst, putting my knowledge into practice and polishing my skills. I entered the PhD program knowing exactly what I wanted to research and after 6 months I was completely at a loss. In the PhD program a new world of knowledge was presented before me and it felt like entering an enormous pitch black room, fumbling with objects I didn't understand nor knowing whether they would be of any use.

Gradually, this new world with its new knowledge started to sink in, relevant questions began to arise and I was able to define the unknown world around me. I found a great interest in how people make decisions and felt like I struck gold when I came across investment decisions made by the entire Swedish population in the recently reformed pension system. This data was coupled with extensive demographic data which in all formed a unique, world class data set, in many ways perfectly suited for research in "choice under uncertainty".

I worked this "gold-mine" for some time and began cooperating with some top grade researchers who brought the potential of this work to new levels. I owe a great deal of thanks to my co-researchers; Lars Nordén, who is also my supervisor, Andrei Simonov who has given me much encouragement and feedback, Grant McQueen who not only has inspired and helped my writing but also gave me an opportunity to teach at Brigham Young University and finally Massimo Massa who seems to be a never ending source of energy.

As a PhD student one can't survive without a little help from ones friends. Both senior and junior PhD students have inspired and helped me, probably more than they realize. We have fought our way through stochastic calculus, the letter *m*, hoards of essays and a number of seminars and conferences. Jens Lusua, Jonas Råsbrandt, Fredrik Berchtold, Therese Strand, Bengt Pramborg, Håkan Thorsell, Vita and John Andrews and all my other friends in the PhD program have made this experience so much richer.

Financially, I owe a great deal of thanks to Stockholm University for hiring me and at times letting me teach a number of courses for extra pay. Olle Högberg, Jens Lindberg, Jan Smolarski, Mohammed Tahir and Lars Nordén allowed me to teach or lead student seminars on their courses. It has not only helped me financially but also allowed me to polish my teaching skills and win the "Ballerina" prize, the students' award for best teacher at Stockholm University School of Business. I am also grateful to The Wallander & Hedelius foundation, the Tore Browaldh foundation and SNS for financial support during my time at Stockholm Institute of Transition Economics at Stockholm School of Economics.

Not the least, I thank my loving wife Veronica and my four children; Ebba, Gustav, Viktor and Sofia who help me keep my feet on the ground. Veronica's ability to "keep it real" and simplify complex problems has been an invaluable asset for my research at times when I would get stuck. Also, her encouragement and faith in my ability to complete this thesis has fuelled me and kept me on track. This thesis is her accomplishment as much as it is mine.

Anders Karlsson
Stockholm, April 2007

Contents

Introduction

In this dissertation I primarily study risk preferences and investment decisions of non professional investors. I analyze a large number of investment decisions based on theories that have been developed and formalized over the past 50 years. Understanding how individuals handle choice under uncertainty when they invest money is of great interest for economic policy makers as well as for academic research. One reason for this great interest is trying to foresee the effects of redesigning existing pension systems, an effect which will be felt by many and consequently affect the economy as a whole.

As the majority of countries in western civilization experience a shift in demographics with an aging population, pension systems are under review. Since the mid 1980's in the US, there has been a major shift from defined benefit, to defined contribution plans, also known as 401(k) plans. In a defined benefit plan, workers are promised a certain amount at retirement. In other words, the benefit, or promised pension, is pre-defined. A "defined contribution" plan means that the portion which is contributed towards ones pension is defined, but the outcome of that portion is uncertain meaning that individuals, instead of receiving a pre-defined benefit or pension, now receive a defined contribution towards their pension savings, which they can invest at their own discretion, not always knowing the outcome at the time of retirement. In Europe, a number of countries have also initiated pension schemes where the beneficiaries are required to take partial responsibility for how their pension savings are invested. In late 2000, all working citizens in Sweden were invited to invest a portion of their wage towards their pension. This would be the start of a new type of pension system in Sweden.

Early in the PhD program I read Terrence Odean's studies of investor overconfidence, Shlomo Benartzis and Richard Thalers studies of naive diversification rules, Mark Grinblatts and Matti Keloharjus studies on investor sophistication and John Campbells and Louis Viceiras work on strategic asset allocation. Their work unveils a number of biases which affect ones choices when the outcome is uncertain, as is the case in pension savings. In my thesis I find evidence of these already known biases and focus on finding rational explanations for this behaviour. But, more importantly, I introduce two unexplored biases; the homeboy bias and menu bias.

The four main issues of the six chapters in this dissertation are;

First, in chapter one I study the propensity of taking more risk if the investment horizon is longer. This is a praxis known as time diversification.

Second, in chapters two and three, the home bias or, the aptitude of choosing investment alternatives that are close to home. I investigate the home bias in chapter two and in chapter three, the aptitude of choosing domestic fund managers over foreign, the "Homeboy Bias".

Third, in chapter four, I introduce the Menu Bias, or the practice of letting the manner in which the investment alternatives are presented guide ones investment decision.

Finally, in chapters five and six, I study the Swedish Pension system itself and identify a number of lessons that can be learned from this experience. In chapter five I find that the Swedish Pension reform has an educational effect on it's participants who are much more likely to participate in the stock market after having made investment decisions in the pension system. In the final chapter I analyze the performance of all participants.

In summary, sophisticated investors are generally less subject to the biases mentioned above, which would have caused them to make sub-optimal investment decisions, in an economic context. Since pension schemes in many nations are shifting towards defined contribution schemes, investors' investment decisions and risk preferences will be of great consequence to their personal economy and ability to consume as retirees, affecting the economy in general. It is therefore of great importance that policy makers do all that they can in increasing investors sophistication and creating a playing field which facilitates economically sound investing.

Data

The data used in this dissertation is nothing short of a gold mine. It can be seen as a nation wide experiment in choice under uncertainty. In the year 2000 Sweden redesigned its pension system where all working individuals were invited to allocate a portion of their pension within an array of mutual funds listed in a prospectus.

Early research on investment behaviour was based on experimental data where participants made investment decisions based on hypothetical situations with no financial consequences. Recently, researchers have studied individuals' actual investment decisions in 401(k) plans. One drawback of exclusively using data from 401(k) plans is that participants often may be employees of the same firm, meaning that the population is more or less homogenous. A second drawback is the difficulty in linking reliable demographic data to the actual individuals making the investment decisions. The Swedish pension system offers a unique opportunity to study how an entire population handles choice under uncertainty.

The data represent a cross section of the Swedish work force. The first pension investments in the new pension system, in autumn 2000, involved 4.4 million

individuals. Their investment choices are linked with individual demographic data collected by Statistics Sweden for the year 2000.[1] The same year, Statistics Sweden surveyed 15,000 households on household economy and other demographics. This survey represents a cross section of the Swedish population. This compiled data set makes it possible to study investment behaviour in great detail. For each individual there is information on the amount invested, which funds and how many funds the individual has invested in. Also, the age, gender, education, occupation, disposable income and net wealth for the same individual is included in the data set. From the 15,651 individuals with complete individual information in the data set, 10,375 individuals (66.4%) made an active investment decision. For these individuals it is possible to investigate the exact allocation of assets in their portfolios. The remaining 5,276 individuals (33.7%) did not make an active investment decision. Instead, they are assigned to the default alternative; the Seventh Swedish Pension Fund, which is an equity fund run by the government. We treat the default alternative as an entirely passive choice. Even if an individual considered the default fund to be the optimal choice, and acted accordingly, he/she still shows up as making a passive choice in the data set[2].

Six features make these data unique and suitable for the questions asked in this dissertation. First, they include choices made by an entire population. Even though not all citizens made a choice, all were offered to make a choice and the data is so extensive in nature, that techniques can be used to correct for any selection bias. Second, the investment choice involved a non trivial amount of real money. Third, this money can only be invested on ones pension and not be passed on to any third party. Fourth, all investment decisions were made during the same brief time period in late 2000. Fifth, all investors receive the same

[1] Data sources from Statistics Sweden are, HEK 2000; a report on household economy, IoF 2000; income report and SUN 2000; educational status. These three reports are for the total population in Sweden. They are linked to a survey on 15,000 households reporting in-depth wealth statistics.
[2] For a detailed analysis over default investors see Engström and Westerberg 2003

information for a specific investment universe. The original prospectus that was sent to all working citizens contains basic information on historical return, risk, fee and a short description of investment strategy for over 400 funds. Finally, thanks to Statistics Sweden, demographic conditions like wage, wealth, occupation, education, marital status and much more are known for all investors.

Summary of the Thesis

Chapter 1. Time diversification in pension savings[3]

Anders Karlsson

The purpose of this chapter is to investigate a practice often recommended by financial planners; to reduce the risk of ones investments as investment horizon decreases. I test whether investors follow this practice and whether this practice can be explained by differences in human capital or if it can be attributed to investors believing in what is commonly referred to as "time diversification". Time diversification is the belief that the risk of stocks decline as investment horizon increases.

For this purpose I use a unique Swedish dataset and investigate how a great number of investors choose to allocate a portion of their pension. The first round of choices made in the new Swedish pension system in the year 2000 has many characteristics of a laboratory experiment. A unique and very important feature of the data is that the investment horizon is known for each investor, since the portfolios are for retirement purposes only and all investments are locked in until time of retirement. Also, all investors are provided with the same information, including risk, at the same time for the available funds. By fiat, the entire work force, those with short and long

[3] Winner of the Best Paper Award at the 13th Conference on the Theories and Practices of Securities and Financial Markets

investment horizons, in the Swedish population construct portfolios using an equal proportion of their wage. I couple this individual portfolio choice data with an extensive database of individual demographic and economic variables and attempt to calculate each investor's human capital. Finally I test whether investors with longer horizons take more risk and to which degree this can be explained by their human capital.

The results show that investors with long horizons choose higher risk portfolios than investors with short horizons and the results remain robust even when controlling for differences in human capital. I also find it difficult to reject this finding in an economic context. My conclusion is therefore that investors seem to believe in time diversification.

Chapter 2. Home sweet home: Home bias and international diversification among individual investors[4]

Anders Karlsson and Lars Norden

In this chapter we investigate differences in home bias for individuals by studying portfolios formed within the new defined contribution pension plan in Sweden, and consider both rational and behavioral explanations for individuals' relative home bias. Specifically, we analyze whether home bias is related to investors' desire to hedge against inflation, level of investor sophistication, and overconfidence. Our analysis contributes to previous research in several ways. First, we study home bias in individuals' portfolio choices, whereas most previous work studies home bias on an aggregate level. Second, the inclusion of assets in the pension portfolios is made entirely in the interest of future returns. Hence, the investments contain no corporate control dimension that can be of significant importance when investigating the home

[4] Published in the Journal of Banking & Finance 31 (2007) 317-333

bias issue in a stock market.[5] Third, we analyze the choices of a large representative sample of a nation's entire working force with a well-defined investment set. All portfolios are formed at the same time, and investors are provided with the same information brochure regarding available mutual funds. Hence, the set up is close to a laboratory experiment in terms of choice under uncertainty. Fourth, after controlling for several reasons for home bias found in previous research, our detailed data enable a categorization of individuals based on demographic and socioeconomic features such as gender, age, education level, wealth and income. Thus, besides analyzing established theories for home bias, we study relationships between investor demographics and the likelihood of being home-biased.

We find that the typical profile for a home biased investor is a not so sophisticated man, who has a high level of job security, and seems to be somewhat overconfident.

Chapter 3. The homeboy bias: Evidence for and determinants of

Anders Karlsson and Grant McQueen

In this chapter, we document and study a new investor bias that we call the homeboy bias. Homeboy is a slang term, from the Hip-Hop culture of North America, for a close friend from ones' own neighborhood, hometown, or region. For our purposes, the homeboy bias refers to the tendency of investors to invest with financial institutions from their own country.

At least since Levy and Sarnat (1970) and clearly since French and Poterba (1991), finance researchers have known that investors have a preference for assets in their own country. Coval and Moskowitz (1999) and Huberman (2001) extend this preference for domestic assets to local assets. Massa and

[5] Dahlquist et al. (2002) argue that in a country like Sweden, where controlling shareholders are economically important, a large home bias is expected on an aggregate level.

Simonov (2006) suggest that investors are biased toward familiar assets because of they have better information about them; in contrast, Grinblatt and Keloharju (2001) suggest that the bias is based on a familiar language or culture. Whereas the *home* bias is a preference for domestic or local assets, the *homeboy* bias is a preference for domestic fund managers. We show that the homeboy bias is empirically distinct from the home bias. For example, investors prefer international funds offered by local institutions to international funds offered by foreign institutions.

We posit five potential explanations for the homeboy bias; the first three are economic explanations and the last two are behavioural explanations. First, the preference for domestic managers could be the consequence of an information advantage they have about domestic assets. Second, the selection of domestic managers could be driven by a preference for funds denominated in the domestic currency. Third, the choice of local institutions could be associated with benefits to the local economy in terms of employment and tax revenue. Fourth, the homeboy bias could be another manifestation of the general bias toward the familiar (see Zajonc (1969) and Huberman (2001)). Fifth, the bias toward domestic managers could be evidence of behaviours such as national identity or pride and xenophobia (see Muller-Peters (1998) and Rygren (2004)). Our tests contradict the first three economic explanations and support the latter two behavioural explanations.

Chapter 4. Portfolio choice and menu exposure[6]

Anders Karlsson, Andrei Simonov and Massimo Massa

Are investors more likely to invest in a growth fund if fund companies offer more growth funds? The goal of this chapter is to address this question by investigating how investment is affected by the available alternatives. We study

[6] Winner of the Best Paper Award in Behavioural Finance, European Finance Association 2006

whether investors choose to invest in assets that are proportionally more represented in the menu of choices available to them and how an exogenous change in the menu representation affects investor decision.

We show that investors choose funds according to the funds' representation in the menu. Categories that offer more funds are chosen proportionally more. Moreover, if, for some exogenous reason, the number of funds of a specific category rises, the investors rebalance their portfolios increasing their investment in that category. This is the case for both the new funds being offered and the existing funds belonging to the category (i.e. growth) in which the number of funds has increased.

We define this phenomenon – the tendency to invest in the stocks that are more heavily represented in the menu – as "menu exposure". It posits that a bigger representation in the menu conveys the idea that the particular category is better, regardless of its intrinsic merit.

Our findings contribute to two main literatures. The first is the literature on "style investing" and the second is the literature on the "home bias". In the literature on style investing, Mullanaithan (2002) and Mullanaithan and Shleifer (2006) argue that investors have a tendency to think through categories. Barberis and Shleifer (2003) and Barberis et al. (2005) show that investors classify risky assets into different styles and that news about one style can affect the prices of other apparently unrelated styles. Stocks returns are therefore affected because they belong to a particular style (Teo and Woo, 2003).

We show that an exogenous change in the way a category is represented in the menu does indeed affect portfolio choice. An increase in the number of funds of a category raises the investment in the funds already existing in the category. In particular, an increase of the weight of a category in the menu by 1% induces the investors to increase their investment in this category by 0.38%.

This supports the interpretation that style-based demand is created and not just catered to.

Chapter 5. Stock market participation and pension reform[7]

Anders Karlsson, Andrei Simonov and Massimo Massa

This chapter examines how the introduction of a funded defined contribution retirement system affects participants' propensity to participate in the stock market. This is an interesting topic, in view of the transition to funded individual account plans in many nations around the world. Our main question is whether individuals perceive their investments in the pension scheme as a substitute for direct investments, and whether allocating more equities in their pension accounts induces participants to reduce or increase their directly-held equity investments.

Standard portfolio choice theory suggests that investors would choose an optimal overall market exposure taking into account both direct and indirect (e.g., through a pension scheme) investment in equity. That is, pension and non-pension equity holdings would be seen as substitutes, at least to a first order. By contrast, behavioural finance theorists instead hypothesize that investors tend to categorize their investments, along with associated gains and losses, according to narrow categories. This approach, labelled "mental accounting" or "narrow framing," predicts that investors apply such mental accounting to stock holdings and react separately to gains and losses for different stocks (Barberis and Huang, 2001).

The results show that investors do not perceive direct investment in the equity market as a close substitute for their retirement accounts, suggesting that an individual account system does not crowd out direct equity market

[7] Published Oct 2007 by the Pension Research Council at the Wharton School, as a chapter in: "Transitioning to Retirement: How Will Boomers Fare?"

investment. The new Swedish system may actually help educate investors of the benefits of stock market participation, increasing participation and therefore, indirectly, boosting saving.

Chapter 6. Benefits of contribution: Individual asset allocation, diversification and welfare in a defined contribution pension system

Anders Karlsson and Lars Nordén

In this chapter we analyze the Swedish pension system in three steps. In the first step, a factor analysis is performed in order to explore the latent factors, or actual asset classes, that are driving the returns of the mutual funds. The findings show that although there appears to be a wide range of choices available to individuals, the large amount of mutual funds can be represented with only a few orthogonal factors or distinct asset classes. By using the ten most important factors, more than 90 percent of the total variance of returns for the original set of 465 mutual funds can be accounted for. In other words, allowing for roughly ten percent noise, about ten orthogonal asset classes are available for investors to choose among in the initial round of investment in the defined contribution part of the Swedish pension system.

After having identified the factors within the Swedish defined pension contribution system the analysis is turned to individuals' asset allocation. Using a sample of individuals taking part in the defined contribution pension portfolio formation in the year 2000, the individuals' loadings and communalities with respect to the different factors are analyzed. Moreover, the individuals' factor exposures are also related to a number of demographic and socio-economic variables in order to find out whether asset allocation differs with respect to individual characteristics. The results show that more sophisticated individuals have a higher probability of making an active choice,

and the result of this activity is to reduce the loadings on the overall market factor and domestic Swedish bond factors.

Finally, the performance of the individuals' portfolios is investigated over the first four years since the introduction of the new Swedish pension system. Jensen's alpha from a regression of an individual's monthly excess return on excess returns on a set of market indices, related to the orthogonal factors, is used to measure performance. The results indicate that sophisticated investors are more likely to choose worse performing funds than non-sophisticated investors. Also, men and older investors are more likely to pick well performing funds than women and younger investors.

Chapter 1. Time Diversification in Pension Savings[1]

Anders Karlsson

1. Introduction

The purpose of this chapter is to investigate a practice often recommended by financial planners; to reduce the risk of ones investments as investment horizon decreases. I test whether investors follow this practice and whether this practice can be explained by differences in human capital or if it can be attributed to investors believing in what is commonly referred to as "time diversification". Time diversification is the belief that the risk of stocks decline as investment horizon increases.

For this purpose I use a unique Swedish dataset and investigate how a great number of investors choose to allocate a portion of their pension. The first round of choices made in the new Swedish pension system in the year 2000 has many characteristics of a laboratory experiment. A unique and very important feature of the data is that the investment horizon is known for each investor, since the portfolios are for retirement purposes only and all investments are locked in until time of retirement. Also, all investors are provided with the same information, including risk, at the same time for the available funds. By fiat, the entire work force, those with short and long investment horizons, in the Swedish population construct portfolios using an equal proportion of their wage. I couple this individual portfolio choice data with an extensive database of individual demographic and economic variables and attempt to calculate each

[1] Winner of the Best Paper Award at the 13th Conference on the Theories and Practices of Securities and Financial Markets

investor's human capital. Finally I test whether investors with longer horizons take more risk and to which degree this can be explained by their human capital.

The results show that investors with long horizons choose higher risk portfolios than investors with short horizons and the results remain robust even when controlling for differences in human capital. I also find it difficult to reject this finding in an economic context. My conclusion is therefore that investors seem to believe in time diversification.

In the following sections I first review theory concerning time diversification. In section III I present my data and describe the Swedish pension system focusing on the defined contribution portion and discuss the methodology. In section IV I present my results and I offer my concluding remarks in Section V.

2. Theory

The basic idea of time-diversification is that above-average returns tend to offset below-average returns over long horizons. Formally, if returns are lognormal and independent over time then the average return will increase linearly with time while the standard deviation will increase by the square root of time. Consequently, the risk of getting a lower return than the risk free rate, alternatively, the risk of loosing money, approaches zero as time moves towards infinity. The other side of this argument is that although the risk of losing money may decrease with time, the amount that can be lost increases proportionally thus canceling out any increase in utility that a longer horizon can offer.

Table 1. Factors that explain horizons effect on risk

Factors	According to:	zero horizon coefficient if	positive horizon coefficient if	negative horizon coefficient if
1. Non-tradable assets (human capital)	P. Samuelson 1994 Bodie, Merton, W. Samuelson 1992	Investors have no human capital	Human capital decreases with time	Human capital increases with time
2. Covariance between earning power and market performance	Bodie, Merton, W. Samuelson 1992	Investors human capital is not correlated with market performance	Investors human capital has a negative covariance with market performance	Investors human capital has a positive covariance with market performance
3. Frequency of required withdrawals from portfolio	P. Samuelson 1989	Risk less component of savings is excluded	One future withdrawal is required	Periodic withdrawals are required to finance everyday expenditures
4. Risky asset return process	Expected Utility	Security returns are IID	Security prices display mean reversion	Security prices display mean aversion
5. Investors relative risk aversion	Expected Utility	Utility displays constant relative risk aversion	Utility displays decreasing relative risk aversion	Utility displays increasing relative risk aversion

In table 1 I review five possible explanations why horizon could affect risk taking, three against time diversification and two in favor of time diversification.

Arguments against Time Diversification

While critics agree that the ratio of expected return to standard deviation (reward-to-risk ratio) increases with time, they also point out that the size of an investor's potential loss increases in the same proportion as the expected

returns, thus reducing the attractiveness of the higher reward-to-risk ratio. Although one is less likely to lose money over a long horizon than over a short horizon, the magnitude of the potential loss increases with the duration of the investment horizon. Kritzman makes a comparison with cross-sectional diversification. If an investor is unwilling to invest $10,000 in a risky project based on his level of risk aversion, then that same investor would not agree to invest in ten independent but equally risky projects which require $10,000 each. Although the investors' risk of losing money is reduced when investing in the ten independent projects, the exposure and, therefore, the amount the investor risks losing is ten times as great. The only way to reduce risk while keeping the exposure constant is if the investor instead is able to invest $1,000 in each of the independent projects. Kritzman further explains that whether an investor has a utility function equal to the logarithm of wealth, or if an individual is even more risk averse and has a utility function equal to negative one divided by wealth, the utility of the risky venture will remain unchanged over time, meaning that an increase in time horizon will not affect an investors tolerance towards risk (see Kritzman 1994 for details). The critics of time diversification are well aware of the studies that show that as investors grow older or as the investment horizon decreases, investors tend to carry lower risk in their portfolios. However, the critics claim that the observed relationship between risk and horizon is not driven by time diversification. I will briefly review three of these alternative explanations.

Non tradable assets (i.e. human capital) may have an impact on the risk of an investor's portfolio. The prediction of utility theory assuming constant relative risk aversion is that the fraction of equities in proportion to true total wealth is unchanged over time. True total wealth is defined as human capital plus liquid capital. Samuelson (1994) illustrates the confounding effect of

human capital with a young professional with future non-security earnings. Since the human capital prospects can not be capitalized or borrowed on, to keep the portion of equities at a proper fraction of true total wealth, the young professional should keep a relatively large fraction of his liquid wealth in equities. Later in life, as human capital is converged into liquid capital, the fractional holding of equities appears to decrease when compared to liquid capital, whereas, in fact, the fraction has remained unchanged when compared to true total wealth. Developing this argument further, Bodie, Merton and W. Samuelson (1992) propose that an investor with the ability to work a little harder or postpone consumption is more likely to take a higher risk whereas an older professional does not have the same option.

Another aspect of human capital expressed by Bodie, Merton and W. Samuelson (1992) concerns ones ability to work in times of poor risky-asset returns. I call this the covariance between ones earnings and market performance. If two investors with the same investment horizon differ in covariance, then the investor with a higher covariance is expected to have a lower portfolio risk. The reason being that, if covariance is high, earning power is reduced in times of poor market performance resulting in a major shortfall. A longer investment horizon allows for more of these extreme shortfalls. A "high covariance" investor is more cautious of taking risky investments. If this is true, then investment horizons effect on risk is related to the covariance of ones earning power with market performance.

The third factor in table 1 concerns the frequency of required withdrawals from ones portfolio. Samuelson (1989) suggests that this can explain why risk may decrease as horizon decreases. If an investor requires a minimum amount at a future date i.e. for future pension payments, the investor makes a series of low risk investments to meet this requirement. As the future date approaches,

the low risk fund increases in size compared to high risk investments. Such an investment strategy creates the illusion that the investor believes in time diversification. The opposite is true if periodic withdrawals are needed for everyday consumption i.e. if the investor is a pensioner and the investment horizon goes from 65 years of age to death. Then the low risk fund will shrink in size gradually compared to the risky investments as time goes by which may give the appearance of an investment strategy which is opposite to time diversification.

Arguments in support of Time Diversification

Researchers who criticize the notion of time-diversification commonly assume that returns are Independent and Identically Distributed (IID) and that investors have constant relative risk aversion (CRRA), meaning that they will allocate the same proportion of their wealth towards risky assets regardless of their absolute level of wealth. The following two arguments challenge the assumptions of CRRA investors and IID returns. If either of these two assumptions do not hold, then economists argue that time does diversify risk.

First, Thorley (1995) shows mathematically that given an expected utility setting and that investors have a decreasing relative risk aversion towards what they perceive to be serially uncorrelated returns, time is in fact a factor in investment decisions and that the allocation towards risk should increase with an increase in investment horizon. Several techniques are used in estimating the nature of investors' relative risk aversion. The various techniques lead to conflicting conclusions. J. Pratt (1964) and K. Arrow (1965) formalize measures of risk aversion and suggest that investors generally have an increasing relative risk aversion (IRRA). K. Arrow referred to IRRA's ability

to explain observed economic behavior with respect to holding cash (see Selden 1956, Friedman 1959, Latane 1963 and Meltzer 1963). The measurement of relative risk aversion has shown to be sensitive to what measure of wealth is used. Research since the mid 1960's have shown evidence of increasing- (Siegel and Hoban 1982, Eisenhauer and Halek 1999), constant- (Szpiro 1986) and decreasing- relative risk aversion (Levy 1994).

Second, There is a great deal of research on exploring how mean reverting prices affect risk with respect to investment horizon. Campbell and Viceira (2002) find that prices over the past 100 years are mean reverting and therefore time diversifies risk. According to Campbell and Viceira, because of mean reversion, there is a degree of predictability in stock prices. Measured over long horizons the risk in stock returns is lower than when measured over shorter horizons. Bull markets tend to follow bear markets. Short term investors will seek assets with higher mean reversion, namely bonds, while investors with longer investment horizons accept assets that are less mean reverting since over time the long term risk is lower than the short term risk. This is suggested to be true regardless of an investors relative risk aversion. Hansson and Persson (2000) also conclude that optimal weights using US data from 1990-1997 suggest that time diversification exists and that allocation decision seems to be independent of the utility function of the investor.

The idea of mean reverting prices is not uncontested. Brown, William and Ross (1995) argue that if an equity market survives, average returns in the beginning of a time period is higher than average return near the end of that time period. For this reason, statistical measures of long-term dependence are typically

biased towards rejection of a random walk. On another note, ex post evidence of mean reversion does not guarantee mean reverting prices in the future.

3. Data and Methodology

Data

The data used in this study represent a cross section of the Swedish work force. The first pension investments in the new pension system, in autumn 2000, involved 4.4 million individuals. Each investor is given a brochure including information on all investment alternatives. Figure 1 is a representation of how the funds are presented to the investor and what information is made available to all investors. In total 464 funds are available[2].

Figure 1: Extract from the PPM Investor Information Folder, fund example

Fund number	Fund name, Management company	Information regarding the funds	Fund fee (%)	Percentage return 99-12-31 (after fees)						Total risk
				In the year					Last 5 years	(last 3 years)
				95	96	97	98	99		
191080	Baring Global Emerging Markets Baring International Fund Managers (Ireland) Ltd	Emerging markets' equity and equity related assets	1.59	-32	10	25	-25	77	25.3	32 (Red)

The percentage return for the last five years equals the compounded annual growth rate of return for the years 1995 through 1999. The total risk corresponds to an annualised percentage standard deviation of three-year monthly historical fund returns. The total risk is also categorised into five different classes, and colours, with respect to standard deviation; Class 1: very low risk, dark green, percentage standard deviation in the range 0-2; Class 2: low risk, light green, 3-7; Class 3: average risk, yellow, 8-17; Class 4: high risk, orange, 18-24; Class 5: very high risk, red, 25-.

Source: Premiepensionsmyndigheten (2000).

[2] 455 funds were included in the original brochure. Before the first choice was completed some were added and some were taken away resulting in a total of 464 funds.

The funds are divided into four major categories; Equity, mixed, generation and bond funds. Equity funds consist of 100 percent equity and are in turn divided into several subgroups i.e. separate regions or sectors. Mixed and generation funds contain both equity and bonds and normally report the level of equity. Generation funds are pre-set mix funds designed to have a high level of equity if the potential client has a long time to retirement and vice versa. Consequently, an investor who picks a generation fund corresponding to his age will act in a time diversifying manner. Bond funds are a mixture of different bonds. All investment choices in the pension system are linked with individual demographic data collected by Statistics Sweden for the year 2000.[3] Statistics Sweden surveys 15,000 households which represent a cross section of the whole population in Sweden. This compiled data set makes it possible to study investment behaviour in great detail. For each individual there is information on the amount invested, which funds and how many funds the individual has invested in. Also, the age, gender, education, occupation, income and net wealth for the same individual is included in the data set. From the 15,651 individuals with complete individual information in the data set, 10,375 individuals (66.4%) made an active investment decision. For these individuals it is possible to investigate the exact allocation of assets in their portfolios. The remaining 5,276 individuals (33.7%) did not make an active investment decision. Instead, they are assigned to the default alternative; the Seventh Swedish Pension Fund, which is an equity fund run by the government. I treat the default alternative as an entirely passive choice. Even if an individual considered the default fund to

[3] Data sources from Statistics Sweden are, HEK 2000; a report on household economy, IoF 2000; income report and SUN 2000; educational status. These three reports are for the total population in Sweden. They are linked to a survey on 15,000 households reporting in-depth wealth statistics.

be the optimal choice, and acted accordingly, he/she still shows up as making a passive choice in the data set[4].

Our data set has five advantages. First, I know the approximate investment horizon for each investor since the investment can be used for retirement purposes only and can not be passed on to a third party my. Second, all investors choose from the same investment universe and were given the same information on that investment universe. The information includes a risk measure on most funds which provides us with reasonable proxies for perceived risk.[5] Third, the data is a representative sample of an entire country's population . Fourth, all investment decisions were made during the same brief time period..[6]Fifth, I have a number of variables stating the demographic and economic situation of the individual investor. These data have been gathered for the same time period that the initial investment choices were made. The Swedish pension system offers a unique opportunity to study how an entire population handles choice under uncertainty.

Dependent variable, Portfolio risk

Four risk measures with varying degrees of sophistication are associated with each investor.

First, I use proportion of equity in ones portfolio. The available funds are divided into four groups: equity-, mixed-, generation- and bond funds. For mixed- and generation funds, information concerning equity proportion is generally available. If no information is available for one particular fund I

[4] For a detailed analysis over default investors see Engström and Westerberg 2003
[5] We use five different risk measures.
[6] Under certain circumstances the retirement account could be passed on to a spouse, but not to any other third party.

assign it the equity proportion of its peer funds in the same subgroup[7] in the brochure.

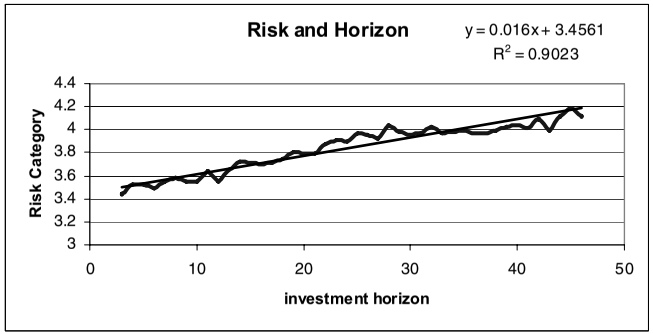
Second, all funds with three years of history or more are assigned a risk category from 1-5 represented by a colored graph. An illustration where a green flat line represents the lowest risk category (1) and a jagged red line represents the highest risk category (5). Funds with less history than three years are assigned the average risk category of its peers in the same subgroup in the brochure. I use the weighted average risk category of the funds in each portfolio as my second risk measure.

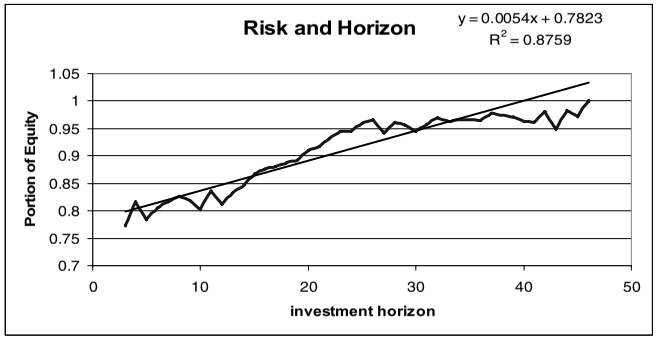
Third and fourth, next to the risk category illustration there is a number representing the annual standard deviation calculated using returns for the past 36 months. I use this number to construct two risk measures: the weighted average standard deviation of the funds in a portfolio with and without considering the covariance between these funds. Although the covariance between funds is not included in the general information given to all investors there is no stopping them in gathering this information on their own. Also, I can not rule out the possibility that investors have a certain feeling for correlation between sectors. I therefore use both measures in my tests. All funds do not have 36 months of history, wherefore I extrapolate a risk measure for these funds by assigning it the average 36 month standard deviation of the funds in its subgroup.

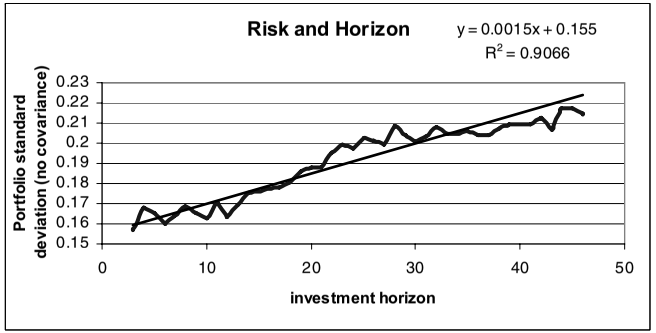
In figure 2 displays four graphs relating the average level of risk to investment horizon (65-age). We see a clear correlation between risk and horizon with R-squared between 0.8538 and 0.9066.

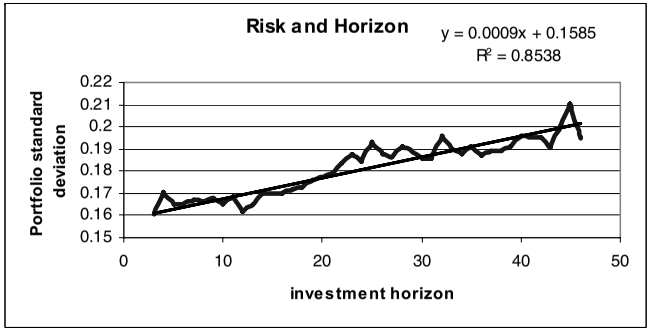
[7] The brochure assigns each fund to a subgroup consisting of funds with similar allocation objectives (i.e., Swedish growth stocks or European value stocks)

Figure 2: Relationship between risk and investment horizon









Our proxy for risk includes four variables; amount invested in equity / total invested amount (equity), the average risk category according to the information in the brochure (risk category), weighted average standard deviation without considering covariance (stdev (no cov)) and portfolio standard deviation including the covariance (stdev). Investment horizon is the amount of years left to retirement (65-age) since this particular investment must be discontinued at the age of retirement.

Independent variables

My primary focus is to investigate to what extent investment horizon affects asset allocation with regards to risk. The investment horizon in this particular investment is 65 minus the investors age since this investment is for retirement purposes only and can not be passed on to any third party. In my model I will use the logarithm of investment horizon, since the difference between 5 and 4 years left to retirement represents a 20 percent reduction in time; whereas, the difference between 35 and 34 years is only a 2.5 percent reduction in time left to retirement. I mentioned 5 main theories as to why investment horizon affects or appears to affect risk. In the following paragraphs I explain the data used in attempt to control for these theories.

First, Samuelson states that one must control for total wealth, meaning liquid wealth + human capital. I use the logarithm of net-wealth assuming a concave utility function. Net wealth is made up of four components: market value of low risk assets plus market value of risky assets plus market value of real estate less debt. The survey used to calculate this particular data (HEK 2000) includes foreign as well as domestic assets and debt. The market value of a house is estimated as the tax assessment value times the ratio of market price to assessed value using data from recent sales prices of houses in the same area. In Sweden, condominiums are not assessed for taxation purposes. The market value of condominiums is estimated as the average value of the recently sold condominiums in the immediate area. According to Samuelson, net wealth or, liquid capital as he calls it, is only one part of total wealth. The other part is human capital. Human Capital can be defined as a discounted present value of

expected income[8] (see Halek and Eisenhauer 1999, Poterba et al 2003 and Cocco and Gomes 2005).

$$E(income_i) = \alpha(age_i) + \gamma\ Z_i + \varepsilon_i \qquad (1)$$

Expected income can be seen as a function of age and a vector of other individual characteristics (Z_i). I regress income on the individuals' age, occupation, education, whether he lives in metropolitan area, a smaller town or on the country side, his gender, marital status and number of children. The estimated coefficients are then used when I derive expected income for each year until retirement. I do this by allowing age to increase one year at a time while holding all other variables constant multiplying them with the parameter estimates from equation (1). The discounted value of expected income represents an individuals non-tradable asset; human capital.

$$Human\ \ Capital_i = \sum_{t=1}^{65-age} \frac{E(income_i)_t}{(1+r_i)^t} \qquad (2)$$

When estimating the present value of future income the discount factor used (r_i) should correspond to occupational risk. My main assumption is that individuals within the same occupational cohort have approximately the same occupational risk. Following this argument, the discount factor will also be similar for these individuals. For example, I estimate expected income for all employees in a specific sector. Once I have these estimates I calculate human capital using the same discount factor for all employees within that specific sector. I use the same

[8] We have also calculated human capital as the present value of disposable income but due to the tax system in Sweden, demographic variables such as education and occupation are better at explaining income than disposable income.

discount factor as Halek and Eisenhauer 1999, namely 2% corresponding to the inflation adjusted risk free rate, for all four separate groups.

Our primary interest lies in controlling for differences in human capital between individuals. I argue that using the same discount factor for individuals with similar occupational risk is sufficient in revealing these differences. I therefore estimate expected income and its present value separately first for twelve occupational cohorts but find similar results when only using four occupational cohorts. I use employed in the public sector (occ1), employed in the private sector (occ2), self employed (occ3) and employment unknown or un employed (occ4). The results from the different cohorts are not comparable, but they serve as robustness checks to verify whether the coefficient signs and levels of significance tell the same story. I also estimate human capital using the same discount factor for the whole population regardless of occupation. Consequently I have five series of estimates of human capital, one for each of the four occupational cohorts and one for the entire sample. Albeit a noisy measure, I argue that it captures the essential portion of human capital.

Bodie's argument is that one can work more if investments go bad. A risky investment gone poorly can be compensated by working harder or consuming less over the remaining investment horizon. A longer investment horizon will then allow higher risk. Since human capital is related to age I expect a high level of multicollinearity in my model. I therefore orthogonalize the horizon variable with human capital and use the error term from the regression:

$$\ln(65 - age_i) = \alpha + \beta_1 human\ capital_i + \varepsilon_i \qquad (3)$$

thus using the portion of investment horizon which is not explained by my human capital related variable.

Second, I consider the covariance between expected income and market performance. It is argued that horizons influence on risk will differ depending on the covariance between an investors earning power and market performance. I use my four occupational cohorts mentioned earlier as dummies. I need not make any assumptions on the differences in market covariance between the different cohorts, only that market covariance is similar for all individuals within the same cohort. In the four equations where the occupational cohorts are estimated separately, the aspect of covariance between market performance and earning power is taken into account.

Third, Samuelson raises the issue of how frequent withdrawals need to be made from the investor's portfolio. The basis for his argument is that the present value of a minimum level of required wealth at retirement is invested in a risk less fund and that this investment will become an increasing portion in ones portfolio. Consequently, the risk less portion of pension savings resembles a time diversification strategy. Therefore, I need to focus on pension savings in excess of the portion required for minimum wealth at retirement to see if horizon affects risk. The investments observed in my dataset are in excess of the portion required for minimum wealth at retirement and is therefore suitable to use in this context. So, by default this aspect is taken into consideration by the nature of this data.

Factors four and five for explaining horizons effect on risk,deal with the investors' relative risk aversion and whether returns are IID or not. Based on classic economics I assume that investors believe returns are random walk and have constant relative risk aversion. By so doing, any significance in the horizon coefficient after controlling for human capital indicates that investment horizons effect on ones appetite for risk is explained by something other than human capital. The literature suggests that such an investor believes in time

diversification in the sense that he either has a decreasing relative risk aversion or believes that asset prices are mean reverting.

The Menu

Investors' choices are affected by how the alternatives are presented (Benartzi and Thaler 2001). As mentioned earlier, each investor is given a brochure including information on all investment alternatives. In my analysis, I add all aspects of how the investment alternatives are presented in an attempt to control for the effect they may have on risk.

Portfolio optimization primarily concerns risk and return. I therefore control for historical return reported in the brochure with regards to the well documented momentum effect. I also control for the number of years of historical return reported in the brochure and whether the fund is new. In each asset class (equity-, mixed-, generation- and bond funds) there are subgroups (e.g. medical funds, European funds etc). The order in which the subgroups are presented in the broschure and the alphabetical order within the subgroups are quantified into a number between 0 and 1 to control for the order in which the funds are presented. A fund in the first subgroup in the brochure and starting with the letter "A" will consequently receive values close to 0. In the fund information in the brochure the investor gets information concerning the portion of domestic/foreign assets in the fund. Since home bias is a known issue in asset allocation I control for this aspect. Market cap and fee are also presumed to impact investors' choice and are therefore included in the regression. Since my intention is to measure the time-diversification phenomenon, I need to control for a category of funds called "generation funds". Generation funds are similar to the suggested "pre-set mix" fund in the U.S. The investors that have chosen

generation funds could be seen as time-diversifiers by default if they choose the "correct" fund for their investment horizon. Since the data provides the details of all investment choices I can control for those who have chosen generation funds, whether they have chosen a "correct" mix with regards to their investment horizon or not. For example, an investor with 10 years to retirement who choses a generation fund labeled "20 years to retirement" will be assigned a 1 for the dummy "longer" in the sense that he chose a fund with a longer investment horizon than the "correct" one.

Method

The purpose of this paper is to test whether I can empirically discard the belief in time diversification. The data I have provides information on the risk level of a specific investment bearing economic consequence and the corresponding horizon of this investment, namely time to retirement. Because my sample suffers from selection bias, in the sense that one third of my sample ended up investing in the default fund with unknown risk, I estimate my parameters with the two step Heckman procedure (Heckman 1979) where first the likelihood of investing is estimated from a probit model. The method may be described by the following two equations:

$$risk_i = x_{1i}\beta_1 + \varepsilon_{1i} \ , \qquad (4)$$

$$e_i^* = x_{2i}\beta_2 + \varepsilon_{2i} \ . \qquad (5)$$

Equation (4) determines the individual's risk, whereas (5) is a "participation equation" describing the individual's propensity to work. Thus, $risk_i$ is the observed risk for individual i if she participates in the pension system and e_i^* is a latent variable that captures the propensity to participate in the pension system; x_{1i} and x_{2i} are vectors of observed explanatory variables, such as age and

education; ε_{1i} and ε_{2i} are mean-zero stochastic errors representing the influence of unobserved variables affecting $risk_i$. The vectors of interest are β_1 and β_2.

Although the latent variable e_i^* is unobserved, I can define a dummy variable $e_i = 1$ if $e_i^* \geq 0$ and $e_i = 0$ otherwise; I thus observe individual risk only if $e_i = 1$, i.e. only if the individual makes an observable choice in the pension system. It is possible that the unobserved terms ε_{1i} and ε_{2i} are positively correlated; individuals that are more likely to take a higher risk, given x_{1i} and x_{2i}, might also be more likely to participate in the pension system. If so, the sample of individuals that participate in the pension system will not accurately represent the underlying population.

Heckman suggests the following method to deal with this selection problem. Note that the conditional mean of ε_{1i} can be written as:

$$E(\varepsilon_{1i} \mid e_i^* \geq 0) = E(\varepsilon_{1i} \mid \varepsilon_{2i} \geq -x_{2i}\beta_2), \tag{6}$$

and hence

$$E(risk_i \mid x_{1i}, e_i = 1) = x_{1i}\beta_1 + E(\varepsilon_{1i} \mid \varepsilon_{2i} \geq -x_{2i}\beta_2). \tag{7}$$

Thus, the regression equation on the selected sample depends on both x_{1i} and x_{2i}. Omitting the conditional mean of ε_{1i} will cause the estimates of β_1 to be biased (unless ε_{1i} and ε_{2i} are uncorrelated, in which case the conditional mean of ε_{1i} is zero).

Under the assumption that ε_{1i} and ε_{2i} are drawn from a bivariate normal distribution, I can derive the regression equation:

$$E(risk_i \mid x_{1i}, e_i = 1) = x_{1i}\beta_1 + \rho\sigma_1\lambda_i \ , \qquad (8)$$

In (8) ρ is the correlation coefficient between ε_{1i} and ε_{2i}, σ_1 is the standard deviation of ε_{1i}, and λ_i, which is the inverse mills ratio, is given by

$$\lambda_i = \frac{\phi(x_{2i}\beta_2 / \sigma_2)}{\Phi(x_{2i}\beta_2 / \sigma_2)} \ , \qquad (9)$$

where ϕ and Φ are the density and distribution functions of the standard normal distribution and σ_2 is the standard deviation of ε_{2i}.

Heckman shows how to estimate (8) in a two step procedure. The first step involves estimating the parameters in (5) by the probit method, using the entire sample. These estimates can then be used to compute λ_i for each individual in the sample. Once λ_i is computed, I can estimate (8) over the sample of working individuals with an OLS regression, treating $\rho\sigma_1$ as the regression coefficient for λ_i. When modelling the likelihood of making an active investment choice I use experience with risky assets and the amount invested in this specific investment[9] as explanatory variables. The inverse mills ratio from the probit estimation is used in the second step estimation with risk as the dependent variable and investment horizon being the key explanatory variable and proxies for the factors which are known to affect risk, mentioned in the previous section. The database I use provides suitable proxies for three of these factors:

[9] We also estimate the Heckman model with a maximum likelihood procedure and retrieve very similar results.

human capital, the covariance between earning power and market performance and the frequency of required withdrawals from ones portfolio.

I run heckman estimations for each occupational cohort and one for the entire sample where I include dummies for occupation. I do this for four different risk measures which in total provide us with 25 estimates for the horizon coefficient.

4. Empirical Results

Estimating expected income

I estimate the present value of expected income in accordance with equation (1). Ln(income) is explained by age, age^2, education level and major, occupation, whether you live in a city, town or village, gender, marital status and number of children. In table 2 I report the coefficients and t-statistics used in estimating expected income. In the general equation where all individuals' income is estimated, the R-squared is 68% which is relatively large. On the other hand, when estimating income for the self employed the R-squared is just above 2% and for the self employed, hardly any of the coefficients are significantly different from zero. R-squared for the other three occupational cohorts are around 10%. Examining the coefficients in my five regressions I first see that age has a positive effect on income which seems to turn negative as age reaches a certain level (since age^2 is negative). Secondly, in most cases, education level seems to have a reasonable effect on income. Post high school education yields a higher income and less than high school yields a lower income.

Table 2, results from OLS regressions on ln(income)

	All	OCC 1	OCC 2	OCC 3	OCC 4
age	0.059	0.130	0.089	0.193	-0.272
	4.35	*9.31*	*9.46*	*1.49*	*-3.79*
age2	-0.001	-0.001	-0.001	-0.002	0.002
	-4.89	*-8.40*	*-8.35*	*-1.45*	*2.56*
edl1	-0.229	-0.142	-0.081	-0.170	-1.101
	-4.03	*-1.72*	*-1.97*	*-0.61*	*-3.42*
edl3	0.458	0.322	0.281	-0.003	1.517
	12.82	*9.23*	*10.49*	*-0.01*	*5.13*
edm2	-0.068	-0.077	-0.032	-0.148	0.024
	-1.51	*-1.08*	*-1.06*	*-0.38*	*0.07*
edm3	0.020	0.038	-0.180	-0.337	0.641
	0.40	*1.03*	*-3.04*	*-0.58*	*1.66*
edm4	-0.011	-0.080	-0.120	-0.126	0.822
	-0.21	*-1.43*	*-3.19*	*-0.35*	*2.32*
occ2	0.160				
	5.54				
occ3	-1.034				
	-9.61				
occ4	-8.451				
	-75.69				
urban	0.104	0.082	0.141	0.259	0.145
	2.87	*2.25*	*5.27*	*0.98*	*0.58*
rural	-0.073	-0.048	-0.032	-0.051	-0.148
	-1.95	*-1.16*	*-1.12*	*-0.20*	*-0.57*
gender	0.475	0.472	0.480	0.281	0.448
	13.28	*13.55*	*17.69*	*1.23*	*1.89*
married	0.129	0.031	0.084	0.681	0.380
	2.97	*0.71*	*3.33*	*1.88*	*1.56*
children	-0.100	-0.101	-0.072	0.023	-0.169
	-5.45	*-4.62*	*-4.51*	*0.17*	*-1.35*
constant	10.634	8.847	9.932	6.055	9.943
	39.79	*31.38*	*56.71*	*2.22*	*6.96*
n obs	15651	4486	8736	745	1648
mean ln(inc)	11.38	11.94	10.93	10.93	3.38
prob > F	0.000	0.000	0.000	0.202	0.000
adj. R-sq	0.679	0.109	0.095	0.021	0.116

Age and age^2 relate to the individuals age 31 Dec 2000. Edl1 = less than high school education, edl2 (default) = high school education, edl3 = more than high school education. Edm1 (default) = social science major, edm2 = technical engineer major, edm3 = major in medicine and edm4 = unknown major. Occ1 (default) = employed in public sector, occ2 = employed in private sector, occ3 = self employed and occ4 = sector unknown. Urban = investor lives in one of Sweden's three major cities, town (default) = investor lives in a town, rural = investor lives on the county side. Gender; 1 = man, 0 = woman. Married or cohabitant = 1, single = 0. Children, refers to number of children. N.obs is the number of observation in the entire sample and in each occupational cohort. Mean ln(inc) = the average ln(income) for the entire sample and each occupational cohort. Prob>F = the probability that all coefficients are collectively indistinguishable from zero. Adj. R-sq = the adjusted R-squared of the regressions. The t-statistics are reported under each coefficient in italics.

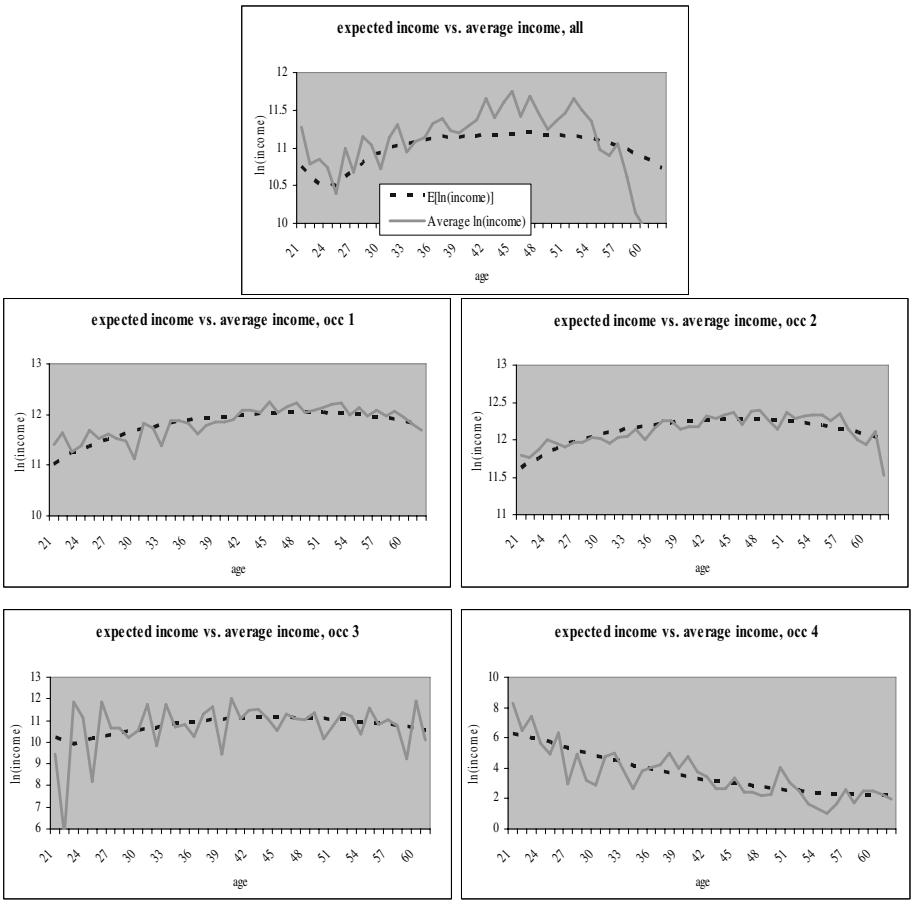
The level of education seems to matter more than education major also, an education major within social science (edm1) tends to result in a higher income than an education within medicine (edm3)[10]. Third, income is higher for those employed in the private sector (occ2) compared to employees in the public sector and lower for the self employed (occ3) and lower for those that are un-employed or have an un-known employment (occ4). Fourth, those living in a metropolitan area (urban)[11] have a higher income than those living in smaller towns (town = default) or on the country side (rural). Finally, Men tend to have higher income than women and married individuals tend to have higher income while those with children on average have lover income. A possible explanation for the coefficient on children could be that if one parent stays home or works part time, he or she will have a relatively low income thus decreasing the average income for individuals with children.

All parameters mentioned above are used when I calculate expected ln(income). When so doing, all variables are held constant except age. A string of expected ln(income) from current age to retirement is converted to expected income and then discounted to a present value in accordance with equation (2). Figure 3 displays the estimated values for ln(income) for ages between 21 and 62 compared to income averages for the corresponding age group. The hump-shaped estimated curve corresponds well with the actual income averages in my samples.

As I discussed earlier, when calculating the present value, the discount factor should vary in accordance with the risk of ones occupation.

[10] Edm4 = unknown education major which makes this coefficient difficult to interpret.
[11] Here, urban is a dummy for living in one of the three major cities in Sweden: Stockholm, Göteborg or Malmö. Rural is if you live on the country side.

Figure 3. Expected income vs. average income



Expected income is the predicted value of ln(income) which is calculated for each individual separately using the coefficients in table 2, the dotted line is the average of the separate outcomes. Average income is the average ln(income) for ages 21 to 62 in my sample. The top figure includes all individuals. In the other four figures I divide the individuals on basis of their occupation. Occ 1= employed in the public sector, occ 2 = employed in the private sector, occ 3 = self employed and occ 4 = unknown employment or unemployed.

I address this issue by estimating the present value of expected income separately for each occupational cohort. By doing this, I assume the same occupational risk for all individuals within the same cohort. I therefore let the occupational cohorts in themselves act as proxies for covariance between

market performance and earning power. I then estimate the present value of expected income separately for each group using the same discount factor for the four groups.

Descriptive statistics of all variables that will be used in my main regressions are reported in table 3. As for my risk measures, I note that the demand for equity is fairly high, 90.3% equity on average. The high proportion of equity is reflected in the average portfolio standard deviation of ca: 18%. Overall, the average risk level is fairly high for a pension plan. Two explanations for high level of risk are that first, previous to this particular investment decision the investors had just witnessed one of the largest bull markets in recorded history which may have increased their appetite for risk. Second, this investment is in excess of the minimum amount required for their retirement. They are guaranteed a certain amount at retirement meaning that all investors can afford to take a higher risk in this particular investment than if their entire pension depended on it. Since market conditions are the same for all investors I assume this will not skew my results.

According to my variables on individual characteristics the average age is 43. Investment horizons in the dataset span from ln(3 years) to ln(46 years) with an average of ln(21.97 years). These numbers reflect time to retirement and capture a representative sample of the working force in Sweden. Human capital is lower for individuals in the public sector than for individuals in the private sector. Also, the standard deviation for the human capital estimates is larger for individuals in the private sector than those in the public sector. I find these results to be reasonable. There is a large variation in ln(net wealth) with a minimum of 0[12] and a maximum ln(MSEK 354).

[12] For simplicity, investors with slightly negative or zero net wealth are assigned a net wealth of SEK 1

Table 3. Summary statistics, average values

Risk measures from the brochure

Variable	Obs	Mean	Std. Dev.	Min	Max
equity	10375	0.903	0.169	0	1
riskcat	10375	3.811	0.574	1	5
stdev (no cov)	10375	0.188	0.050	0	0.53
stdev	10375	0.179	0.043	0.002	0.766

Individual characteristics

Variable	Obs	Mean	Std. Dev.	Min	Max
age	15651	43.028	11.111	19	62
horizon	15651	2.922	0.634	1.099	3.829
human capital (all)	15651	2 495 358	1 983 527	36	11 600 000
human capital (occ1)	4486	2 489 360	1 362 728	225 995	8 207 950
human capital (occ2)	8736	3 528 756	1 872 326	247 734	10 000 000
human capital (occ3)	745	838 310	508 538	45 794	2 426 976
human capital (occ4)	1684	134 384	90 682	12 106	358 252
Ln(net wealth)	15651	8.290	6.041	0	19.686

Information from the brochure

Variable	Obs	Mean	Std. Dev.	Min	Max
compounded return	10375	1.429	0.519	-0.241	6.705
domestic	10375	0.342	0.246	0	1
correct	2729	n.a.	n.a.	0	1
shorter	499	n.a.	n.a.	0	1
longer	283	n.a.	n.a.	0	1
market cap	10375	3418.6	11178.0	0	208666.3
Fee (percent)	10375	0.734	0.300	0	1.98
years of history	10375	1.971	1.495	0	5
new	10375	0.602	0.360	0	1
subgroup	10375	0.557	0.242	0.053	1
alphabet	10375	0.536	0.219	0.036	1

Selection variables for probit estimation in heckman model

Variable	Obs	Mean	Std. Dev.	Min	Max
experience with risky assets	9507	n.a.	n.a.	0	1
amount invested	15 651	12 858	6 647	210	26 202

Full sample = 15 651 observations, sub-sample including only those who made an active portfolio choice = 10 375 observations. When estimating human capital for the four occupational cohorts I only use individuals in that cohort.

Risk measures include four variables; amount invested in equity / total invested amount (equity), the average risk category according to the information in the brochure (risk category), weighted average standard deviation without considering covariance (stdev (no cov)) and portfolio standard deviation including the covariance (stdev).
Individual Characteristic consists of four main variables: age is the investor's age, horizon is ln(65-age). In the heckman regression, I use horizon orthogonalized with regards to human capital in accordance with equation (3). Human capital is the present value of estimated ln(income) estimated for all investors and for each occupational cohort separately reported in SEK. Occ1=employed in public sector, occ2=employed in private sector, occ3=self employed and occ4=unknown employment or unemployed. Net wealth is financial wealth + real estate – debt.
Information from the brochure include 11 variables which are found in the brochure given to each investor. Compounded return is the total return from Jan 1997 to Dec 1999, domestic is the proportion of investment alternatives with investments in the Swedish market, correct, shorter and longer corresponds to the number of individuals who have chosen generation funds with a correct, shorter or longer investment horizon than is prescribed by that particular fund. A generation fund can be advertised as a fund tailored for investors with, for example, 20 years to retirement. If an investor with 10 years to retirement chooses this funds, the dummy variable "longer" =1 for this investor. Market cap is the average fund market capitalization measured in MSEK. Fee is the average fund fee reported in percent. Years of history is the average number of years of recorded fund performance listed in the brochure where the maximum number of years is 5. A fund is labeled as new if it has less than three years of performance data listed in the brochure. Subgroup and alphabet deal with the order in which the funds are presented. In each asset class (equity-, mixed-, generation- and bond funds) there are subgroups (e.g. medical funds, European funds etc). The order of the subgroups and the alphabetical order within the subgroups are quantified into a number between 0 and 1.
Selection variables for probit estimation in heckman model include compounded three year return, a dummy variable for previous experience with risky assets, meaning equity or equity funds. 61 % of the population was exposed to risky assets prior to this investment decision. Amount invested represents the SEK amount invested in this particular investment.

From the information in the brochure I find that the average compounded three year return is 143%, which is exceptionally high in comparison with historical figures. This high figure reflects an unusually positive development for equity markets. This among other factors may explain the comparatively high rate of participation in the pension system (only one third default investors) and the demand for equity as noted earlier. Funds investing in domestic markets represent 34.2% of all investment alternatives, which is relatively high since the proportion of Swedish assets on global markets is closer to 3%. Correct, shorter and longer corresponds to the generation funds available. 33 percent of all active investors (3511) invested in generation funds where 77 percent of them (2729) chose funds with "correct" horizon. The average market cap is MSEK 3,419 and the fee is on average 0.734 percent. In the brochure each fund reports at most 5 years of history where the average fund has ca: 2 years of history. A fund is considered new if there is less than two years of history since there then isn't enough data to calculate the funds standard deviation based on 36 months annualized return. Subgroup and alphabet simply refers to the order in which the funds are presented. All funds are listed in separate subgroups and in alphabetical order. A mean above 0.5 indicates that funds in the latter half of the individual subgroups were on average more popular than the rest.

Finally, I have my two variables related to the fist step selection equation in the Heckman procedure. Approximately 61% of the population or 9507 individuals in my sample have previous experience with risky assets, meaning equity or equity funds. This variable and the amount invested in the pension system have proven to be of importance in explaining pension system participation (see Karlsson & Nordén 2004 and Engström & Westerberg 2003).

Human capital is per definition correlated with age wherefore I orthogonalize the horizon variable so it reflects the portion of horizon not explained by human capital. From the regressions in table 4 I observe large values for adjusted R-squared as expected. There is however still a portion of horizon which is not explained by my proxies related to human capital.

Table 4. Results from regression explaining horizon with human capital

	all	occ1	occ2	occ3	occ4
Human Capital	1.9E-07	3.2E-07	2.5E-07	9.2E-07	8.2E-06
constant	2.439	2.042	2.117	2.023	1.704
n.obs	15651	4486	8736	745	1684
Prob > F	0.0000	0.0000	0.0000	0.0000	0.0000
adj. R-sq	0.37	0.54	0.59	0.63	0.83

Horizon is defined as Ln(65-age). Human capital is the present value of expected income. N.obs is the number of observation in the entire sample and in each occupational cohort. Prob>F = the probability that all coefficients are collectively indistinguishable from zero. Adj. R-sq = the adjusted R-squared. Occ 1 represents employees in the public sector, occ2 represent employees in the private sector, occ3 represent self employed and occ4 are of unknown employment.

I then regress risk on the proportion of investment horizon not explained by human capital as the main independent variable, using a Heckman procedure. In the regression I also include human capital, ln(net wealth), 11 variables on fund characteristics and how the funds are presented in the broschure (reported in table 3) and two variables, experience with risky assets and amount invested, used in the first stage probit estimations. For robustness I use four different risk measures as the dependent variable and redo my estimations five times, one for each occupational cohort and one for the entire population where I include occupational dummies. Since I am primarily interested in the horizon coefficient I report only them in table 5 while the full results can be found in the appendix.

In table 5 I see that out of the four occupational cohorts, occ4 (employment sector unknown) have the weakest results where the horizon coefficient is not significantly different from zero in three cases out of four. This corresponds to the fact that the R-squared in table 4 for occ4 is 0.83 which indicates that most of the investment horizon for those in occ4 is inseparable from their human capital. In most other cases, the horizon coefficient is significantly positive on the 1% level.

Table 5. Horizon coefficients

	all	occ1	occ2	occ3	occ4
equity	0.078	0.089	0.077	0.059	0.060
t-value	*27.3*	*15.2*	*18.29*	*3.13*	*3.48*
change 5 - 45 yr horizon	*11%*	*13%*	*11%*	*10%*	*9%*
riskcat	0.175	0.180	0.161	0.171	(0.116)
t-value	*22.7*	*11.63*	*14*	*2.94*	*2.41*
change 5 - 45 yr horizon	*6%*	*7%*	*6%*	*6%*	*5%*
stdev (no cov)	0.016	0.017	0.015	0.016	(0.010)
t-value	*23.14*	*12.15*	*14.29*	*2.84*	*2.34*
change 5 - 45 yr horizon	*11%*	*12%*	*10%*	*11%*	*7%*
stdev	0.007	0.006	0.006	(0.007)	(0.006)
t-value	*10.13*	*5.3*	*5.31*	*0.83*	*1.36*
change 5 - 45 yr horizon	*6%*	*5%*	*5%*	*5%*	*4%*

Coefficients in parenthesis are not significantly separate from zero on a 1% level

I regress risk on the portion of investment horizon (ln(65-age) not explained by human capital as the main independent variable, using a Heckman procedure. In the regression I also include human capital, ln(net wealth), 11 variables on how the funds are presented in the broschure (reported in table 3) and two variables, experience with risky assets and amount invested, used in the first stage probit estimation.

Table 5 only reports coefficients for the portion of investment horizon (ln(65-age) not explained by human capital. The full results are reported in the Appendix (A1-A5) As a robustness check I run the regressions using four separate risk measures; amount invested in equity / total invested amount (equity), the average risk category according to the information in the brochure (risk cat), weighted average standard deviation without considering covariance (stdev (no cov)) and portfolio standard deviation including the covariance (stdev). The Heckman regression is estimated separately for each occupational cohort and once for the entire population. Occ 1 represents employees in the public sector, occ2 represent employees in the private sector, occ3 represent self employed and occ4 are of unknown employment. The t-values are reported under each coefficient in italics. *Change 5 – 45 yr horizon* represents the change in the risk measure if the investment horizon changes from 5 to 45 years all else held equal. When the risk measure is estimated, all variables, except horizon, are held at average values.

Although the t-values are large, the relevance of the horizon coefficient is not altogether clear. I therefore measure the difference in estimated risk between a 5 year and 45 year investment horizon, all other variables held equal. In this case the differences in risk range between 13% - 5% which in my view is too large to ignore.

The results state that investment horizon matters even when I include my proxy for human capital. As I have discussed in the paper, this could be an effect of investors believing in time diversification in the sense that they either have decreasing relative risk aversion or believe that prices are mean reverting. Another explanation is that investors are simply inconsistent. In line with the example by Kritzman earlier, maybe investors feel differently towards 5% risk of loosing $100 and 0.05% risk of loosing $10,000 although the expected value is the same.

Another possible explanation is that age might be a factor in and of itself. Maybe older investors take less risk than younger investors. However, Calvet, Campbell and Sodini (2006) investigated the portfolio structure of the entire Swedish population in 2002 and found virtually no age effect. Older investors had little or no difference in risk compared to younger investors. There was however no indication of the investment horizon of these portfolios. These individuals could be saving for their grand children, a car or a vacation trip next month.

In summary, my results can not discard the beleif in time diversification in pension savings among non professional investors, at least not in a pure statistical context. Also, I find that changes in risk caused by changing the horizon coefficient are too large to dismiss.

5. Conclusion

Previous research offers compelling arguments for and against time diversification. Arguments against time diversification are that if returns are

IID, investors have no human capital and have constant relative risk aversion then, time should not diversify risk. Arguments related to human capital in some way or another are used to explain why many studies show a positive relationship between risk and investment horizon. These arguments stress that it may be rational to increase risk as investment horizon increases but time in and of itself does not decrease risk. Arguments for time diversification attack the assumptions of IID returns and constant relative risk aversion. If investors have a decreasing relative risk aversion or if asset prices are mean reverting, then investment horizon will affect risk.

The purpose of this analysis is to study whether investment horizon affects risk. In so doing, I attempt to control for three of the explanations offered by economists which would explain such behavior; Investors' human capital, the covariance between their earning power and the market and the frequency of required withdrawals. In accordance with classic finance literature I assume that returns are IID and all investors' utility display constant relative risk aversion.

I use investment data from the first round of the Swedish pension system. my data represents the entire population in Sweden, all investments are made during the same brief time period, and investors chose from the same investment universe and are given the same information on that investment universe. Also, I know the approximate investment horizon for all investors and have four reasonable measures for risk, proportion of equity, average risk category of the portfolio and two measures of portfolio standard deviation, one measure considering the covariance between funds and one without the covariance.

Having linked all investment data to detailed demographic data I calculate estimates for human capital by regressing income on age, education, occupation, gender, marital status and number of children. The present value of expected income is my human capital measure. I use the investors' occupation in an attempt to control for effects on risk caused by the covariance between earning power and market fluctuations. The effects on risk caused by the frequency of required withdrawals are controlled for by the construction of the pension system since no funds can be withdrawn before retirement. Net wealth and historical returns are also included in the model since they also are known to affect risk. Finally, I add control variables to control for confounding effects that may be caused by how the investment alternatives are presented.

One third of the population did not make an observable investment choice. Their money was invested in the government run equity fund which is the default alternative. Since I can not tell which of these investors actively chose not to make a choice or simply didn't care, I treat all default investors as non-active investors. I use the method introduced by Heckman in 1979 to address the selection bias.

We regress risk on three major variables, investment horizon human capital and net wealth. Since human capital and investment horizon are correlated I use the portion of investment horizon which can not explain human capital in my Heckman procedure. My results indicate that investors believe in time diversification. Whether this is due to mean reverting prices, investors having decreasing relative risk aversion or the mere fact that much of the advice in media propagates strategies resembling time diversification, I can not say.

Our proxies for human capital receive coefficient signs in accordance with theory whereas the covariance of ones earning power with market fluctuations, i.e. my occupation dummies, have either coefficients that are indistinguishable from zero or coefficients so small, they have very little economic significance. A full account of all coefficients is found in the appendix.

I conclude that while focusing on risk measures which are available to the average investor, I can not reject the impact of investment horizon on risk, even when including my measure for human capital.

Appendix, variable explanation for tables A1-A4 found after table A4

EQUITY	all	occ1	occ2	occ3	occ4
Horizon	0.0775	0.0886	0.0773	0.0588	0.0604
	27.3	*15.2*	*18.29*	*3.13*	*3.48*
Compounded return	0.1406	0.1781	0.1332	0.0920	0.1148
	51.34	*31.31*	*38.12*	*8.07*	*13.42*
Human capital	1.1E-08	2.3E-08	1.5E-08	5.2E-08	5.7E-07
	14.39	*12.72*	*16.29*	*3.63*	*9.49*
Ln(net wealth)	3.2E-04	-1.4E-04	3.4E-04	1.6E-03	2.2E-03
	1.4	*-0.34*	*1.17*	*1.4*	*2.57*
Employed private sector	0.0066				
	2.34				
Self employed	-0.0082				
	-1.23				
Employment sector unknown	-0.0091				
	-1.71				
Proportion domestic funds	-0.0897	-0.1166	-0.0814	0.0383	-0.0899
	-13.72	*-9.78*	*-9.51*	*1.29*	*-3.91*
correct	-0.0227	-0.0344	-0.0158	-0.1014	-0.0041
	-6.7	*-5.73*	*-3.64*	*-5.49*	*-0.3*
shorter	-0.1078	-0.1017	-0.0905	-0.2336	-0.1547
	-17.78	*-10.05*	*-11.29*	*-9.31*	*-5.12*
longer	-0.0331	-0.0095	-0.0440	-0.0692	-0.0225
	-4.22	*-0.64*	*-4.51*	*-1.77*	*-0.7*
Market cap	3.8E-07	5.6E-07	2.9E-07	-1.2E-08	6.6E-07
	3.29	*2.48*	*2.15*	*-0.01*	*1.41*
fee	0.1072	0.1166	0.0885	0.1225	0.1697
	18.17	*10.29*	*11.92*	*4.68*	*7.67*
Years of history	-0.0340	-0.0373	-0.0288	-0.0440	-0.0425
	-14.74	*-8.47*	*-9.91*	*-4.38*	*-4.88*
new	-0.1190	-0.1358	-0.0972	-0.1614	-0.1290
	-13.51	*-7.93*	*-8.83*	*-4.34*	*-3.79*
subgroup	-0.1399	-0.1618	-0.1219	-0.0366	-0.2042
	-20.51	*-12.95*	*-13.74*	*-1.12*	*-8.41*
alphabet	-0.0676	-0.0407	-0.0739	-0.0714	-0.0908
	-10.91	*-3.46*	*-9.48*	*-2.28*	*-3.98*
_cons	0.9158	0.8560	0.8921	0.8136	0.8755
	64.62	*28.42*	*51.62*	*8.88*	*14.03*
choice					
experience	0.3693	0.3961	0.3871	0.3691	0.0840
	17.29	*13.69*	*15.63*	*7.07*	*2.04*
amount	0.0000	0.0000	0.0000	0.0000	-0.0001
	18.04	*10.87*	*25.09*	*-4.29*	*-14.27*
_cons	-0.1608	-0.8669	-0.7425	-1.4887	-0.6464
	-6.46	*-24.96*	*-25.32*	*-26.59*	*-15.16*
mills					
lambda	-0.0615	-0.0372	-0.0489	0.0347	-0.0088
	-6.18	*-2.55*	*-6.26*	*0.87*	*-0.33*
rho	-0.4691	-0.2970	-0.3912	0.2802	-0.0625
sigma	0.1311	0.1252	0.1249	0.1239	0.1401

RISKCAT	all	occ1	occ2	occ3	occ4
Horizon	0.1750	0.1799	0.1606	0.1711	0.1165
	22.7	*11.63*	*14*	*2.94*	*2.41*
Compounded return	0.5826	0.7017	0.5652	0.4368	0.4889
	78.34	*46.46*	*59.66*	*12.34*	*20.55*
Human capital	2.5E-08	5.5E-08	3.3E-08	1.2E-07	1.6E-06
	12.18	*11.39*	*13.36*	*2.62*	*9.77*
Ln(net wealth)	1.4E-03	2.4E-04	1.5E-03	4.7E-03	6.8E-03
	2.27	*0.21*	*1.86*	*1.3*	*2.85*
Employed private sector	0.0159				
	2.07				
Self employed	0.0001				
	0				
Employment sector unknown	-0.0306				
	-2.11				
Proportion domestic funds	0.0217	-0.0325	0.0370	0.2684	0.0773
	1.22	*-1.03*	*1.6*	*2.91*	*1.21*
correct	-0.2003	-0.1868	-0.1879	-0.4195	-0.2473
	-21.75	*-11.75*	*-15.97*	*-7.32*	*-6.48*
shorter	-0.4173	-0.3753	-0.3940	-0.7629	-0.3879
	-25.33	*-13.97*	*-18.11*	*-9.8*	*-4.61*
longer	-0.1134	-0.1088	-0.1141	-0.4133	-0.0389
	-5.32	*-2.77*	*-4.3*	*-3.41*	*-0.44*
Market cap	8.9E-07	1.2E-06	8.8E-07	-2.0E-06	7.7E-07
	2.85	*2.04*	*2.36*	*-0.78*	*0.59*
fee	0.5889	0.6211	0.5446	0.6154	0.6862
	36.76	*20.64*	*27.02*	*7.57*	*11.14*
Years of history	-0.1510	-0.1632	-0.1368	-0.1402	-0.1809
	-24.1	*-13.96*	*-17.35*	*-4.5*	*-7.47*
new	-0.4844	-0.4874	-0.4317	-0.4961	-0.6317
	-20.25	*-10.73*	*-14.46*	*-4.29*	*-6.67*
subgroup	-0.0561	-0.1169	-0.0334	0.4359	-0.1455
	-3.03	*-3.53*	*-1.39*	*4.29*	*-2.15*
alphabet	-0.0754	0.0153	-0.0756	-0.1687	-0.2521
	-4.48	*0.49*	*-3.58*	*-1.73*	*-3.97*
_cons	3.2848	3.0584	3.2363	3.1070	3.2444
	85.28	*38.23*	*69.02*	*10.96*	*18.69*
choice					
experience	0.3693	0.3961	0.3871	0.3691	0.0840
	17.29	*13.69*	*15.63*	*7.07*	*2.04*
amount	0.0000	0.0000	0.0000	0.0000	-0.0001
	18.04	*10.87*	*25.09*	*-4.29*	*-14.27*
_cons	-0.1608	-0.8669	-0.7425	-1.4887	-0.6464
	-6.46	*-24.96*	*-25.32*	*-26.59*	*-15.16*
mills					
lambda	-0.1834	-0.1118	-0.1445	-0.0186	0.0248
	-6.75	*-2.88*	*-6.8*	*-0.15*	*0.34*
rho	-0.5090	-0.3334	-0.4230	-0.0500	0.0637
sigma	0.3602	0.3353	0.3416	0.3723	0.3899

STDEV_NC	all	occ1	occ2	occ3	occ4
Horizon	0.0161	0.0166	0.0149	0.0163	0.0100
	23.14	*12.15*	*14.29*	*2.84*	*2.34*
Compounded return	0.0417	0.0517	0.0408	0.0206	0.0356
	62.04	*38.62*	*47.43*	*5.91*	*16.91*
Human capital	3.2E-09	5.8E-09	4.0E-09	1.5E-08	1.4E-07
	16.81	*13.74*	*17.6*	*3.34*	*9.52*
Ln(net wealth)	9.4E-05	-5.4E-06	6.8E-05	2.9E-04	6.4E-04
	1.65	*-0.05*	*0.95*	*0.83*	*3.04*
Employed private sector	0.0016				
	2.34				
Self employed	0.0035				
	2.17				
Employment sector unknown	0.0004				
	0.33				
Proportion domestic funds	-0.0028	-0.0037	-0.0016	0.0106	-0.0028
	-1.72	*-1.32*	*-0.76*	*1.17*	*-0.5*
correct	-0.0209	-0.0199	-0.0193	-0.0489	-0.0235
	-25.08	*-14.14*	*-18.09*	*-8.65*	*-6.95*
shorter	-0.0483	-0.0450	-0.0447	-0.0835	-0.0589
	-32.47	*-18.93*	*-22.67*	*-10.87*	*-7.91*
longer	-0.0175	-0.0133	-0.0178	-0.0457	-0.0191
	-9.08	*-3.82*	*-7.4*	*-3.81*	*-2.43*
Market cap	1.5E-07	2.0E-07	1.5E-07	-3.6E-08	4.2E-08
	5.24	*3.78*	*4.4*	*-0.14*	*0.36*
fee	0.0572	0.0551	0.0541	0.0691	0.0656
	39.51	*20.69*	*29.6*	*8.62*	*12.02*
Years of history	-0.0165	-0.0165	-0.0156	-0.0187	-0.0186
	-29.15	*-15.95*	*-21.84*	*-6.08*	*-8.69*
new	-0.0551	-0.0561	-0.0498	-0.0636	-0.0679
	-25.47	*-13.95*	*-18.38*	*-5.58*	*-8.1*
subgroup	-0.0038	-0.0079	-0.0027	0.0392	-0.0109
	-2.29	*-2.68*	*-1.25*	*3.91*	*-1.81*
alphabet	-0.0182	-0.0102	-0.0170	-0.0223	-0.0380
	-11.95	*-3.68*	*-8.87*	*-2.32*	*-6.76*
_cons	0.1698	0.1475	0.1646	0.1748	0.1759
	48.83	*20.86*	*38.71*	*6.25*	*11.44*
choice					
experience	0.3693	0.3961	0.3871	0.3691	0.0840
	17.29	*13.69*	*15.63*	*7.07*	*2.04*
amount	0.0000	0.0000	0.0000	0.0000	-0.0001
	18.04	*10.87*	*25.09*	*-4.29*	*-14.27*
_cons	-0.1608	-0.8669	-0.7425	-1.4887	-0.6464
	-6.46	*-24.96*	*-25.32*	*-26.59*	*-15.16*
mills					
lambda	-0.0130	-0.0057	-0.0115	-0.0011	-0.0006
	-5.37	*-1.67*	*-6*	*-0.09*	*-0.1*
rho	-0.4111	-0.1966	-0.3759	-0.0306	-0.0183
sigma	0.0317	0.0289	0.0306	0.0367	0.0345

STDEV_36M	all	occ1	occ2	occ3	occ4
Horizon	0.0074	0.0062	0.0057	0.0070	0.0062
	10.13	*5.3*	*5.31*	*0.83*	*1.36*
Compounded return	0.0370	0.0498	0.0368	-0.0064	0.0358
	52.76	*43.61*	*41.55*	*-1.25*	*15.95*
Human capital	1.9E-09	3.3E-09	2.1E-09	8.2E-09	7.5E-08
	9.44	*9.24*	*9.14*	*1.27*	*4.77*
Ln(net wealth)	1.0E-04	4.2E-05	4.0E-05	4.8E-04	4.0E-04
	1.74	*0.49*	*0.54*	*0.93*	*1.76*
Employed private sector	0.0023				
	3.13				
Self employed	0.0120				
	7.07				
Employment sector unknown	0.0038				
	2.76				
Proportion domestic funds	-0.0158	-0.0148	-0.0123	-0.0137	-0.0161
	-9.44	*-6.19*	*-5.66*	*-1.03*	*-2.67*
correct	-0.0046	-0.0038	-0.0032	-0.0327	-0.0036
	-5.35	*-3.17*	*-2.89*	*-3.93*	*-1.01*
shorter	-0.0127	-0.0100	-0.0114	-0.0418	-0.0099
	-8.21	*-4.94*	*-5.6*	*-3.7*	*-1.24*
longer	-0.0055	0.0002	-0.0061	-0.0440	-0.0042
	-2.76	*0.06*	*-2.44*	*-2.5*	*-0.5*
Market cap	-4.4E-08	-9.8E-09	-4.0E-08	1.9E-08	-1.2E-07
	-1.51	*-0.22*	*-1.15*	*0.05*	*-1*
fee	0.0385	0.0231	0.0345	0.1022	0.0512
	25.54	*10.18*	*18.33*	*8.66*	*8.8*
Years of history	-0.0046	-0.0049	-0.0040	-0.0028	-0.0064
	-7.81	*-5.6*	*-5.4*	*-0.61*	*-2.8*
new	-0.0060	-0.0093	-0.0035	0.0189	-0.0152
	-2.67	*-2.7*	*-1.27*	*1.13*	*-1.7*
subgroup	-0.0157	-0.0212	-0.0129	0.0309	-0.0265
	-8.97	*-8.45*	*-5.73*	*2.09*	*-4.16*
alphabet	-0.0277	-0.0143	-0.0264	-0.0660	-0.0467
	-17.5	*-6.05*	*-13.37*	*-4.67*	*-7.79*
_cons	0.1343	0.1179	0.1342	0.1470	0.1725
	37.2	*19.55*	*30.7*	*3.57*	*10.45*
choice					
experience	0.3693	0.3961	0.3871	0.3691	0.0840
	17.29	*13.69*	*15.63*	*7.07*	*2.04*
amount	0.0000	0.0000	0.0000	0.0000	-0.0001
	18.04	*10.87*	*25.09*	*-4.29*	*-14.27*
_cons	-0.1608	-0.8669	-0.7425	-1.4887	-0.6464
	-6.46	*-24.96*	*-25.32*	*-26.59*	*-15.16*
mills					
lambda	-0.0002	0.0035	-0.0038	-0.0055	-0.0177
	-0.08	*1.2*	*-1.96*	*-0.31*	*-2.5*
rho	-0.0060	0.1415	-0.1256	-0.1018	-0.4418
sigma	0.0315	0.0245	0.0303	0.0542	0.0400

Tables A1-A4 report all coefficients from the Heckman estimations. I run five sets of regressions for four different risk measures: proportion of equity (A1), average risk category (A2), average standard deviation without considering covariance (A3) and portfolio standard deviation (A4). The five columns in each table report coefficients when the Heckman estimations are done for the entire sample (all), individuals employed in the public sector (occ1), private sector (occ2), self employed (occ3) and unknown sector (occ4). The horizon variable is the portion of horizon which is not explained by human capital (hc): $horizon_i = LN(65 - age) - (\alpha + \beta_1(hc)_i)$. Compounded return is the historical average three year return of the portfolio. Human capital is the present value of expected income. Log of net wealth is the logarithm of net wealth (market value of financial assets + market value of real estate – debt). Employed in public sector is the default whereas employed in private sector, self employed and employment sector unknown are the other occupational dummies. Proportion of domestic funds is the percent amount of the portfolio which is invested in domestic assets. Correct, shorter and longer are dummies which refer to generation funds. For example, if an investor with 20 years to retirement invests in a generation fund labeled "ten years to retirement", longer = 1. Market cap is a chosen fund's market capitalization in SEK multiplied by the portfolio weight of that fund. Fee is the weighted average percentage fee of the funds in the portfolio. Years of history represents the weighted average amount of years of the funds in the chosen portfolio. New = 1 if the fund has no history. Subgroup and alphabet deal with the order in which the funds are presented. In each asset class (equity-, mixed-, generation- and bond funds) there are subgroups (e.g. medical funds, European funds etc). The order of the subgroups and the alphabetical order within the subgroups are quantified into a number between 0 and 1. The two variables in the first step probit equation, listed under *choice* are; previous experience with risky assets and amount invested in this particular investment. Under *mills* I report the three parameters lambda, rho and sigma which are associated with the correction of the selection bias.

Chapter 2. Home Sweet Home: Home Bias and International Diversification among Individual Investors[1]

Anders Karlsson and Lars Nordén

1. Introduction

A striking feature of international portfolio investment is the extent to which equity portfolios are concentrated in the domestic market of the investor. French and Poterba (1991) estimate the domestic ownership share of the world's five largest stock markets to be 92.2 percent for the U.S., 95.7 percent for Japan, 92 percent for the U.K., 79 percent for Germany, and 89.4 percent for France (figures for 1990). Many authors try to explain this home bias that runs counter to well known benefits from international diversification.[2] So far, the literature has yet to provide a generally accepted explanation for the observed home bias. Also, relatively little is known regarding the characteristics of an individual who shows a relatively large reluctance towards international diversification and investments. In particular, a complete picture of individuals' home bias is lacking, where it is possible to control for several potential explanations for the bias.

In this chapter we investigate differences in home bias for individuals by studying portfolios formed within the new defined contribution pension plan in Sweden, and consider both rational and behavioral explanations for individuals' relative home bias. Specifically, we analyze whether home bias is related to investors' desire to hedge against inflation, level of investor sophistication, and overconfidence. Our analysis contributes to previous research in several ways. First, we study home bias in individuals' portfolio choices, whereas most previous work studies home bias on

[1] Published in the Journal of Banking & Finance 31 (2007) 317-333
[2] See Lewis (1999) and Karolyi and Stulz (2002) for excellent overviews and summaries of the home bias literature.

an aggregate level. Second, the inclusion of assets in the pension portfolios is made entirely in the interest of future returns. Hence, the investments contain no corporate control dimension that can be of significant importance when investigating the home bias issue in a stock market.[3] Third, we analyze the choices of a large representative sample of a nation's entire working force with a well-defined investment set. All portfolios are formed at the same time, and investors are provided with the same information brochure regarding available mutual funds. Hence, the set up is close to a laboratory experiment in terms of choice under uncertainty. Fourth, after controlling for several reasons for home bias found in previous research, our detailed data enable a categorization of individuals based on demographic and socioeconomic features such as gender, age, education level, wealth and income. Thus, besides analyzing established theories for home bias, we study relationships between investor demographics and the likelihood of being home-biased.

We use a multinomial logit model to estimate the likelihood for an individual to be home-biased. From our results, we identify the type of individual with the highest likelihood of home bias as an older, unmarried, poorly educated man working for the government, who invests only a small amount of money, and who has no experience with risky investments outside the pension plan. Using our theoretical framework, we find less sophisticated individuals to be relatively more home-biased. Moreover, our results are consistent with government employees, having a relatively high job security, and caring more about hedging domestic inflation than about international diversification, thus, having a bias towards domestic assets.[4] Finally, as men are regarded as relatively more overconfident than women with respect to investments, they will have a relatively greater tendency for a perceived information

[3] Dahlquist et al. (2002) argue that in a country like Sweden, where controlling shareholders are economically important, a large home bias is expected on an aggregate level.
[4] We argue that a low job security should be accompanied by international diversification since an economic downturn in the domestic market, which could lead to unemployment, should only have a limited effect on international investments, thus smoothing the individual's consumption.

advantage of domestic assets than women, and thus be more likely to overweigh their portfolios with domestic assets.[5] Hence, we can describe our home-biased candidate as a not so sophisticated man, who has a high level of job security, and seems to be somewhat overconfident.

The rest of the study is organized into four sections. Section 2 briefly presents the Swedish pension system, with emphasis on individual choice in the defined contribution part, and the data used in the analysis. Section 3 outlines the multinomial logit model as a framework for analyzing the likelihood of home bias for individual investors. We also present the explanatory variables, with motivation from previous research. Section 4 contains an analysis of the empirical results. Finally, the study is ended in section 5 with some concluding remarks.

2. Portfolio choice within the defined contribution pensions system

2.1 A brief description of the new Swedish pensions system

The new Swedish pension system was introduced in 2000 and consists of three parts. The first part is the income pension, which is based on 16 percent of the annual income and is used to finance current retirees. The amount paid in also serves as a base for calculating future pension payments. The second part, the premium pension, is based on 2.5 percent of the annual income. In the first round in 2000, 2.5 percent of the previous four years of income was invested. This amount was allocated at each individual's discretion. Each individual was presented with an investment opportunity set of 464 funds and invited to choose between 1 and 5 funds.[6] If no choice was made, the allotted money was invested in the default alternative, the Seventh Swedish Pension Fund, which is an equity fund run by the

[5] Gherig (1993) finds that home bias might be explained by an information advantage in domestic assets. Moreover, according to Barber and Odean (2001), men are relatively more overconfident than women. Although disputed by Dorn and Huberman (2005), the higher male overconfidence will generate a greater tendency for men to have a perceived information advantage of domestic assets than women.

[6] The Swedish pension system is described in further detail at www.ppm.nu and www.pension.nu. See also Engström and Westerberg (2003) and Karlsson (2005).

government. The accrued amount from the premium pension part will be paid out on a monthly basis to the individual during retirement. The third part of the system is a guaranteed pension level designed to ensure that no retiree will be completely without pension payments at the time of retirement, regardless of the previous income. In total, 18.5 percent of each individual's annual income is invested to finance the system. However, people earning more than 7.5 income base amounts per year will only be accredited an upper limit of 7.5 income base amounts, although they will still pay 18.5 percent of their income to the system.[7]

2.2 Individual choice of domestic and foreign assets

In 2000 all participants in the Swedish pension system were provided with a brochure containing the 464 mutual funds with accompanying information on risk, historical returns, fees, and a few words briefly describing each fund.[8] Given the brief fund description, we assess the proportion of domestic assets for each available fund, and calculate each individual's ratio of domestic holdings as the proportion of money allocated to domestic assets in the individual's portfolio of mutual funds. Table 1 presents the different fund categories that are used in the brochure, together with the associated average proportions of domestic and foreign assets. Unless the information in the brochure states otherwise, we use the following classification criteria: Swedish funds are classified as 100 percent domestic, Regional and Nordic funds are classified as 50 percent domestic and 50 percent foreign. European and Industry funds are classified as 100 percent foreign. Mixed funds and Generation funds are classified as either 100 percent, two thirds or 50 percent domestic. In the brochure, Fixed Income funds are sorted as 100 percent domestic or foreign or a

[7] For the year 2000, one income base amount equals SEK 38,800.

[8] The brochure contains fund-specific information on percentage returns for the last five years, which equals the compounded annual growth rate of returns for the years 1995 through 1999. The information on risk corresponds to an annualized percentage standard deviation of three-year monthly historical fund returns. The brochure also presents a categorized risk measure, where the risk is categorized into five different classes, and colors, with respect to standard deviation; Class 1: very low risk, dark green, percentage standard deviation in the range 0-2; Class 2: low risk, light green, 3-7; Class 3: average risk, yellow, 8-17; Class 4: high risk, orange, 18-24; Class 5: very high risk, red, 25-.

mixture thereof. In the latter case, we classify the fund as 50 percent domestic and 50 percent foreign.[9]

Table 1: Summary statistics for the mutual funds

Category 1	Category 2	Category 3	Number of funds	Mean domestic share	Mean foreign share	Mean percentage fee
Equity	Sweden	Sweden (normal)	28	1	0	0.92
		Sweden small cap	6	1	0	1.16
		Sweden index	7	1	0	0.41
	Regional	Swedish equity and foreign equity	11	0.5455	0.4545	0.74
		Nordic countries	12	0.5	0.5	1.27
		Europe	36	0	1	1.11
		Euroland	8	0	1	0.99
		Europe small cap	9	0	1	1.23
		Europe index	7	0	1	0.48
		North America and USA	26	0	1	1.05
		Asia and Far East	18	0	1	1.19
		Global	32	0	1	1.01
		New markets	21	0	1	1.56
	Countries	Japan	20	0	1	1.06
		UK	6	0	1	1.21
		Other countries	19	0	1	1.25
	Industry	IT and Communication	19	0	1	1.15
		Pharmacutical	7	0	1	1.36
		Other industries	16	0.125	0.875	1.17
Mixture	Mixture	Swedish equity and fixed income	3	1	0	1.08
		Swedish equity, Swedish and foreign fixed income	28	0.6577	0.3423	0.74
		Foreign equity and fixed income	22	0	1	0.93
Generation	Generation	Pension in less than 10 years	5	0.2	0.8	0.46
		Pension in less than 20 years	6	0	1	0.46
		Pension in more than 20 years	21	0.0714	0.9286	0.46
Fixed inc.	Fixed inc.	Sweden, short maturity	15	1	0	0.46
		Sweden, long maturity	15	1	0	0.45
		Europe and Euroland	18	0	1	0.70
		Others	15	0.0333	0.9667	0.79
All funds			456	0.2399	0.7601	0.85

[9] Note that the assessment of funds' domestic and foreign investment proportions is based entirely on information from the brochure, and is thus available to each individual investor at the time of the portfolio formation.

In Table 1, we see that the investment opportunity set comprises 24 percent domestic and 76 percent foreign assets. Hence, the opportunity set is skewed towards a high proportion of foreign investments.[10]

2.3 Data sources

Our data comes from the first round of the new Swedish pension system, involving the investment choices of 4.4 million individuals made in 2000. For a sample of 13,749 individuals, who are born in Sweden, we link the investment choices to very detailed demographic and socioeconomic data collected by Statistics Sweden for the year 2000.[11] For each individual, the data set contains information on how much money, which funds and how many funds the individual have invested in. Also, the age, gender, education, occupation, disposable income, and net wealth for the same individual are included.[12] From the 13,749 individuals in the data set, 9,415 individuals (68.5 percent) made an active investment decision. For these individuals we calculate the exact allocation of domestic and foreign assets in their portfolios. The remaining 4,334 individuals (31.5 percent) did not make an active investment decision, and are assigned to the default alternative.

3. Modeling the likelihood of home bias for individual investors

In order to investigate individual differences in home bias, we relate the share of domestic holdings to a set of explanatory variables, including individuals' demographic and socioeconomic characteristics. We divide the domestic share into five categories, 0-0.2, 0.2-0.4, …, 0.8-1.0, and use the multinomial logit model to

[10] Note that gains from domestic and foreign assets holdings, whether they accrue within or outside the Swedish pension system, are treated homogeneously in the Swedish tax system.

[11] Data sources from Statistics Sweden are, HEK 2000; a report on household economy, IoF 2000; income report, and SUN 2000; educational status. These three reports are for the total population in Sweden. They are linked to a survey on 15,000 households, conducted in the year 2000 by Statistics Sweden, reporting in-depth wealth statistics. In addition, information on individuals' country of birth is obtained from Statistics Sweden. Our final data set contains individuals born in Sweden only.

[12] Disposable income represents income after tax plus government subsidies. Net wealth is financial wealth plus market value of real estate minus debt. Financial wealth is reported in detail, how much is held in stocks, funds and risk-free assets. Note that disposable income and net wealth are reported on a household level.

estimate the likelihood of individuals choosing to allocate the portfolios according to the different categories. Thereby, we evaluate the individual relative likelihood of home bias given a set of explanatory variables. To avoid a selection bias, we jointly model the likelihood of home bias and the likelihood of making an active choice. Therefore, we add a sixth category, representing the default alternative, to the multinomial logit model. We assume that each individual simultaneously considers two investment choices: the choice of whether to be active or passive, and the choice of desired relative share of domestic holdings.[13, 14]

We include a set of core explanatory variables, based on previous research, and emphasize three different theories for individual home bias: inflation hedging, investor sophistication and overconfidence. First, relative home bias can be due to individuals' different desires to hedge inflation. Adler and Dumas (1983) and Cooper and Kaplanis (1994) argue that investing in domestic assets serves as a hedge against home country-specific risks, as e.g. inflation and changes in human capital, since investments in domestic assets are likely to follow the performance of the domestic market in general. The purchasing power of these investments will then be relatively constant and hedged against domestic market fluctuations. One drawback of being heavily invested in the domestic market is the risk of unemployment in case of a domestic market downturn, where unemployment is accompanied by poor performance in ones investment portfolio. For this reason,

[13] A natural alternative specification is a nested logit model, where an individual first chooses active participation or not, and second, given an active participation, the individual chooses the level (category) of domestic holdings. Given the information in the data set, we cannot distinguish between an individual who does not make an active choice, and ends up in the default fund, and one who actively chooses the default. Moreover, the two choices, of active participation and domestic share, both depend on the same set of core explanatory variables. In a nested model it would be difficult, if not impossible, to identify the effects of each core explanatory variable on each choice if treated separately. Therefore, we use the multinomial logit model rather than the nested alternative, where we do not distinguish between an active and a passive choice of the default alternative. Also, as a robustness check, we estimate a version of the logit model where we divide the domestic share into only three categories: 0-0.33, 0.34-0.66, and 0.67-1, adding a category for the default. The results from the two models are qualitatively similar. In the following analysis, we present the results only from the more detailed model with five plus one categories, as they are slightly easier to interpret than the ones from the model with three plus one categories. The results from the less detailed model are not presented in the text, but are available from the authors upon request.

[14] See Appendix A for a detailed description of the multinomial logit model.

investors concerned about the risk of unemployment are more inclined to diversify internationally. We argue that for people with a low risk of unemployment, domestic purchasing power is of greater concern than the need for international diversification. We use occupation as a proxy for inflation hedging. In Sweden, people working in the public sector have a high level of job security and the risk of unemployment is relatively small.[15] We expect a higher likelihood of home bias for government employees than for privately or self-employed people.[16]

Second, home bias might be related to a lack of investor sophistication. According to Grinblatt and Keloharju (2001), less sophisticated investors are more home-biased than sophisticated investors. We use five variables to represent investor sophistication: i) a dummy variable which equals one if the individual owns risky assets outside the pension system, i.e. who has previous experience with risky investments; ii) the education level, less than high school, high school or more than high school education; iii) the amount of money invested in the pension system; iv) disposable income; and v) net wealth. We hypothesize that these variables are positively correlated with investor sophistication and negatively related to the likelihood of home bias.

Third, we relate home bias to investor overconfidence. Gherig (1993) argues that an information advantage concerning domestic assets will lead to a home bias. Coval and Moscovitz (1999) find that fund managers have a strong preference for locally managed firms. This locality bias is accompanied with higher returns, suggesting that fund managers have an actual information advantage regarding firms that are located closely to them. This kind of over-weighting will, according to Gherig, occur whether the information advantage is real or only perceived, since no

[15] According to statistics from Statistics Sweden and The National Board of Labor Markets, in the year 2000, the percentage of employees loosing their jobs was 1 percent in the private sector and 0.1 percent in the public sector.
[16] The link between home bias and being employed in the public sector could be an issue of self-selection. A difference in home bias between individuals in the public and private sector might be confirming that home biased individuals are more likely to pursue a career in e.g. the public sector. This point of view is supported by Borjas (2003). The framework of DeMarzo et al. (2004) can also be interpreted to support the issue of endogeneity.

distinction is made between a perceived and an actual information advantage. If all investors are provided with the same information, we cannot assume that any particular group of investors has an actual information advantage. But we do argue that an overconfident investor will have a perceived information advantage concerning the investments that he or she is familiar with. Previous research shows that home bias is related to familiarity (e.g. Kang and Stulz, 1997 and Huberman, 2001). Huberman suggests that home bias is caused by a misperception of risk and that perceived confidence in local assets leads to an underestimation of local risks and an overestimation of foreign risks (Huberman, 2001, Kilka and Weber, 2000, and Heath and Tversky, 1991). In the literature, overconfidence is defined as overestimating the precision of ones estimates (see e.g. Fischhoff et al., 1977). Therefore, we link perceived information advantage with familiarity and ultimately overconfidence. The only assets that can be regarded as familiar in this context are domestic assets, since all investors are Swedish, or at least born in Sweden. Hence, since all individuals are familiar with Swedish assets, everyone is likely to be biased towards Swedish assets. Overconfident individuals are known to misjudge their ability in forecasting the performance of familiar assets, thus amplifying the home bias. Barber and Odean (2001) find that men in general are more overconfident than women.[17] If men are more likely to be overconfident, we expect men to be more home biased. Hence, we include a gender dummy variable in our model. Moreover, Barber and Odean (2001) also find that individual's marital status might weaken the gender effect of overconfidence. Therefore, we include a dummy variable for marriage and an interaction term between the dummy variables for marriage and gender.

In addition to the core explanatory variables, we include a set of control variables. We use historical returns to control for possible momentum effects (Chan et al.,

[17] However, Dorn and Huberman (2005) find little evidence to support that individuals' overconfidence is a factor for explaining differences in diversification.

1996). Returns are calculated based on the fund information from the brochure provided to individuals. We use the compounded annual return for the three years 1997 through 1999. For each individual, the return is calculated as the weighted average for all funds in the portfolio. We use two measures of portfolio risk. The first measure is simply each individual's weighted average category of risk, according to the classification based on standard deviation from the information brochure (see footnote 10). The second measure directly uses the numerical values of the annualized standard deviation of three-year monthly historical portfolio returns. The portfolio standard deviation is calculated by taking each portfolio's weighted average returns for the past 36 months and then calculating the standard deviation of this average return series, thus capturing covariance in returns.[18]

We use transactions costs for the individual portfolios to account for the claim of Black (1974), Stultz (1981) and Kang and Stultz (1997) that costs of international diversification exceeds the gains. Moreover, we add dummy variables for the number of mutual funds chosen. Benartzi and Thaler (2001) indicate that the complicated reality of portfolio diversification may cause inexperienced investors to diversify in a naïve manner, believing that many assets diversify better than few. Hence, a different number of funds might induce diversification effects that must be accounted for. Finally, we include individual age. Age is directly related to the investment horizon, which in turn might affect asset allocation decisions (Karlsson, 2005).

[18] Concerning the calculations of returns, risk categories and standard deviations, a number of new funds lack information. In the brochure, all funds are categorized as in Table 1. For the funds that lack information we assign the average value of the rest of the funds in that category. For example, in the first category (Sweden funds) there are 30 available funds to choose from, of which 14 lack risk measures and eight lack returns for 1999. In this category, the average return in 1999 is 63.7 percent, the average categorized risk measure is four, and the average standard deviation is 20%. Hence, these average values are assigned to those funds that lack information in this category.

4. Empirical results

4.1 Descriptive statistics

Table 2A displays average values for the relative share of domestic funds (w) held by each individual making an active choice, and the other numerical variables. Across all individuals who made an active choice, the average domestic share equals 0.34. This is slightly higher than 0.24, the average domestic share among all available mutual funds, and indicates an overall home bias on the aggregate level. Moreover, we see some stylized facts regarding individual home bias. For example, the average age (AGE) is higher in the groups with a larger share of domestic holdings. Moreover, individuals who choose a low domestic share tend to invest more money (MONEY) in the mutual funds. However, the variables INCOME and WEALTH do not give any clear-cut signals regarding the domestic share. The average transactions cost (FEE) equals 0.75 percent over all individual portfolios of mutual funds. The portfolios with the largest proportions of foreign assets are on

Table 2A: Summary statistics, average values

	All	$w < 0.2$	$0.2 \leq w < 0.4$	$0.4 \leq w < 0.6$	$0.6 \leq w < 0.8$	$w \geq 0.8$	Default
w	0.3399	0.0614	0.2954	0.4896	0.6713	0.8545	-
AGE	42.75	43.30	41.02	42.52	45.10	46.88	43.13
INCOME	364,131	376,219	366,444	352,806	352,787	319,572	316,261
WEALTH	452,631	538,209	399,456	418,728	402,295	350,034	379,443
MONEY	13,783	13,960	14,414	13,490	13,346	11,719	11,550
FEE	0.7579	0.8012	0.8189	0.6975	0.6496	0.6245	-
RETURN	0.5557	0.5558	0.5691	0.5551	0.5208	0.5028	-
RISKCAT	3.8131	3.7417	3.9675	3.8236	3.6293	3.4796	-
STDEV	0.1713	0.1654	0.1785	0.1732	0.1631	0.1552	-

Table 2A presents average sample values for the variables; w = the individual domestic share of invested pensions funds, AGE in years of the individual, INCOME = individual's disposable income in SEK, WEALTH = market value in SEK of financial assets and real estate holdings, net of debt for each individual, MONEY = amount in SEK of the total initial investments in pensions funds for each individual, FEE = percentage transactions cost (fee) for the individual investments, RETURN = the compounded annual growth rate of return for the three years 1997 through 1999, RISKCAT = average risk category of the individual portfolios, where each mutual fund is categorized on a scale in the range 1-5; 1 corresponds to "very low risk", 2 to "low risk", 3 to "average risk", 4 to "high risk" and 5 to "very high risk", STDEV = annualized standard deviation of three-year historical monthly returns. Average sample values are presented with respect to five groups based on the level of w. The column labeled "All" contains overall averages for all individuals, whereas "Default" represents individuals who ended up in the default alternative, rather than actively chose a level of w.

average more costly. Although the cost difference between domestic and foreign assets is unlikely to be a barrier towards international investments, following Black

(1974), it shows the importance of using costs as a control variable. For the return and risk variables, we see no pattern with respect to the chosen domestic share of the individuals.[19] In Table 2B, we display the categorical variables with respect to the domestic share categories, and the default alternative.

Table 2B: Summary statistics, proportion of individuals

	All	$w < 0.2$	$0.2 \leq w < 0.4$	$0.4 \leq w < 0.6$	$0.6 \leq w < 0.8$	$w \geq 0.8$	Default
All	13,749	0.2037	0.2090	0.1780	0.0631	0.0310	0.3152
MALE	6,973	0.2097	0.2042	0.1688	0.0647	0.0326	0.3201
FEMALE	6,776	0.1975	0.2140	0.1876	0.0614	0.0294	0.3102
OCC_1	4,052	0.1935	0.2162	0.1974	0.0671	0.0306	0.2952
OCC_2	7,736	0.2174	0.2208	0.1770	0.0655	0.0286	0.2907
OCC_3	637	0.1931	0.1554	0.1664	0.0361	0.0314	0.4176
OCC_4	1,324	0.1594	0.1443	0.1307	0.0491	0.0461	0.4705
EDU_1	2,415	0.1739	0.1449	0.1706	0.0741	0.0356	0.4008
EDU_2	7,148	0.2034	0.2128	0.1848	0.0639	0.0294	0.3057
EDU_3	4,186	0.2212	0.2396	0.1708	0.0552	0.0311	0.2821
EXP	8,692	0.2208	0.2361	0.1878	0.0587	0.0316	0.2651
UNEXP	5,057	0.1742	0.1625	0.1614	0.0706	0.0299	0.4014
FUND_1	5,675	0.1355	0.0000	0.0580	0.0153	0.0275	0.7637
FUND_2	1,236	0.2273	0.2265	0.3544	0.1165	0.0752	0.0000
FUND_3	1,985	0.2640	0.2972	0.2620	0.1335	0.0433	0.0000
FUND_4	1,897	0.2251	0.3869	0.2646	0.0986	0.0248	0.0000
FUND_5	2,956	0.2703	0.4296	0.2229	0.0622	0.0149	0.0000
MARRIED	10,717	0.2151	0.2160	0.1860	0.0650	0.0328	0.2852
UNMARRIED	3,032	0.1633	0.1844	0.1501	0.0561	0.0247	0.4215

The column labeled "All" contains the number of individuals in the sample, categorized with respect to MALE (dummy variable equal to 1 if the individual is male, and zero otherwise), FEMALE (dummy variable equal to 1 if the individual is female, and zero otherwise), OCC_1 (dummy variable equal to 1 if the individual is employed by the government, and zero otherwise), OCC_2 (dummy variable equal to 1 if the individual is employed in the private sector, and zero otherwise), OCC_3 (dummy variable equal to 1 if the individual is self-employed, and zero otherwise), OCC_4 (dummy variable equal to 1 if the individual is unemployed or not registered, and zero otherwise), EDU_1 (dummy variable equal to 1 if the education level of the individual is below high school), EDU_2 (dummy

[19] In the empirical analysis we will concentrate on the STDEV measure of risk, because the model fit is better using this measure rather than RISKCAT. However, the main results concerning individuals' relative likelihood of home bias are virtually the same irrespective of which measure of risk we use for control purposes. Furthermore, we find that the RETURN variable is highly correlated with STDEV. Hence, we do not use the RETURN variable directly in the analysis, but rather the residual (RETRES) from an ordinary linear regression of RETURN on STDEV.

variable equal to 1 if the education level is high school), EDU_3 (dummy variable equal to 1 if the education level is above high school), EXP (dummy variable equal to 1 if the individual has other risky holdings), UNEXP (dummy variable equal to 1 if the individual has no other risky holdings), FUND_1 through FUND_5 (dummy variables reflecting the number of mutual funds the individual has invested in, 1 through 5), MARRIED (dummy variable equal to 1 if the individual is married), and UNMARRIED (dummy variable equal to 1 if the individual is unmarried). Columns three through seven contain the proportion of the individuals in each row-wise category that fall in each column-wise group according to *w*. The last column contains the proportion of individuals in each row-wise category who ended up in the default alternative, rather than actively choosing a level of *w*.

We can see albeit a small tendency for men (MALE) choosing the extreme quintiles, and the default fund, more often than women (FEMALE). Occupation is divided into four groups: employment in the public sector (OCC_1), private sector (OCC_2), self-employment (OCC_3) and unemployment (OCC_4). From the descriptive statistics, we see only small differences between the public and private sector employees with respect to the domestic share categories. If anything, we observe a larger fraction of private employees (21.7 percent) in the category with domestic share less than 0.2, than of government employees (19.3 percent).

The education level is categorized into three groups; less than, equal to and more than high school education (EDU_1, EDU_2, and EDU_3). The least educated individuals are relatively more represented in the categories with domestic share above 0.6, than in the ones with less than 0.4. Moreover, we represent individuals' experience of risky investments with the dummy variable EXP (UNEXP) that equals one if an individual owns (does not own) risky assets outside the pension system. Experienced people are relatively over-represented in the three categories with domestic share less than 0.6, whereas inexperienced people are somewhat over-represented in the category with domestic share between 0.6 and 0.8.

The dummy variables FUND_1 through FUND_5 represent the number of chosen funds. The option with only one fund is dominated by the default alternative (76.4 percent). For the portfolios with more than one fund, we see international diversification effects. For example, individual representation in the category with domestic share above 0.8 is a decreasing function of the number of funds chosen. In the category with lowest domestic share, the pattern is the reverse, with

representation of only 22.7 percent of the individuals choosing two funds, up to a 27.0 percent representation of the people choosing five funds. Finally, when comparing the summary statistics for married (MARRIED) with unmarried (UNMARRIED) individuals, it is difficult to see a relationship between marriage and international diversification, since the married individuals show higher representation in all categories of domestic share.

4.2 Multinomial logit estimation

Table 3 presents the results from the estimation of the multinomial logit model, with the results for the core variables in Panel A, and the control variables in Panel B. For each explanatory variable, we estimate five coefficients (m = 2, 3, 4, 5 and 6). Each coefficient represents the effect of the variable on the probability of obtaining the outcome $y = m$ relative to the probability of obtaining y = 1, i.e. a domestic share in the lowest range 0-0.2, where y = 2, 3, 4, and 5 are outcomes for an increasingly larger individual domestic share, and y = 6 represents the default. We report a Wald test statistic for each explanatory variable, which is χ^2-distributed under the null hypothesis that the variable does not affect the allocation between domestic and foreign assets. Each Wald test statistic is derived under the null hypothesis that the coefficients are jointly equal to zero for m = 2, 3, 4, and 5. Hence, the likelihood of choosing the default is not considered in the test. For the core results in Panel A, the dummy variables for employment add significantly to the model. Moreover, employment in the private sector (OCC_2) differs significantly from government employment (the base case).

Table 3: Results from the multinomial logit model

Panel A: Core variables

	Pr(y = 2)	Pr(y = 3)	Pr(y = 4)	Pr(y = 5)	Pr(y = 6)	χ^2-statistic
CONSTANT	2.6424 (0.0001)	3.2521 (0.0001)	1.9391 (0.0001)	-0.1313 (0.7624)	1.6765 (0.0001)	-
OCC_2	-0.2433 (0.0003)	-0.3443 (0.0001)	-0.2215 (0.0295)	-0.2155 (0.1119)	-0.1722 (0.0369)	25.83 (0.0001)
OCC_3	-0.0908 (0.5539)	0.0040 (0.9786)	-0.4194 (0.0989)	-0.2271 (0.4173)	0.0877 (0.6023)	3.534 (0.4727)
OCC_4	0.0314 (0.7923)	-0.1185 (0.3387)	0.0036 (0.9834)	0.3839 (0.0437)	0.1882 (0.1684)	6.832 (0.1455)
EXP	-0.0804 (0.1883)	-0.2099 (0.0009)	-0.5793 (0.0001)	-0.3838 (0.0013)	-0.5111 (0.0001)	51.16 (0.0001)
EDU_1	0.1810 (0.0343)	0.3136 (0.0002)	0.3198 (0.0043)	0.1332 (0.3588)	0.3283 (0.0007)	16.19 (0.0028)
EDU_3	0.0360 (0.5636)	-0.0671 (0.3201)	-0.0462 (0.6451)	0.2155 (0.0958)	-0.0116 (0.8791)	6.002 (0.1990)
MONEY	-1.5e-5 (0.0013)	-3.1e-5 (0.0001)	-3.2e-5 (0.0001)	-7.7e-5 (0.0001)	-4.0e-5 (0.0001)	87.65 (0.0001)
INCOME	1.1e-8 (0.8581)	-0.3e-8 (0.9742)	4.4e-8 (0.7157)	2.6e-8 (0.9025)	-5.7e-8 (0.2765)	0.149 (0.9974)
WEALTH	-7.8e-8 (0.0200)	-3.6e-8 (0.0563)	-3.8e-8 (0.3822)	-2.0e-8 (0.6326)	-2.6e-8 (0.1671)	10.29 (0.0359)
MALE	-0.3026 (0.0180)	-0.1262 (0.3570)	0.2850 (0.1362)	0.6277 (0.0194)	0.1084 (0.4870)	19.36 (0.0007)
MARRIED	-0.5598 (0.0001)	-0.4607 (0.0001)	-0.4987 (0.0016)	-0.1081 (0.6444)	-0.4967 (0.0001)	33.68 (0.0001)
MAR_MALE	0.3768 (0.0072)	0.2250 (0.1328)	0.0611 (0.7708)	-0.2861 (0.3250)	0.0021 (0.9903)	21.85 (0.0002)
Panel B: Control variables						
STDEV	-2.1538 (0.0225)	-1.0203 (0.2589)	-3.5896 (0.0038)	-2.0723 (0.0622)	-	11.08 (0.0257)
RETRES	0.4036 (0.0001)	0.1594 (0.0574)	-0.0200 (0.9013)	0.0751 (0.7540)	-	41.95 (0.0001)
FUND_1	-	-0.0839 (0.3625)	-0.5748 (0.0003)	1.4958 (0.0001)	-	88.39 (0.0001)
FUND_2	0.3586 (0.0002)	0.8165 (0.0001)	0.6958 (0.0001)	1.7211 (0.0001)	-	117.2 (0.0001)
FUND_3	0.3584 (0.0001)	0.3826 (0.0001)	0.7974 (0.0001)	1.1575 (0.0001)	-	74.58 (0.0001)
FUND_4	0.7327 (0.0001)	0.6388 (0.0001)	0.8484 (0.0001)	0.9331 (0.0001)	-	116.2 (0.0001)
FEE	0.1450 (0.2105)	-1.2101 (0.0001)	-1.7079 (0.0001)	-1.2776 (0.0001)	-	228.5 (0.0001)
AGE	-0.0205 (0.0001)	-0.0137 (0.0001)	0.0109 (0.0137)	0.0112 (0.0468)	-0.0010 (0.7700)	95.85 (0.0001)

Table 3 contains results from the estimation of the multinomial logit model in Equation (1). The dependent variable y, has six possible outcomes (m = 1, …, 6), where each of the first five corresponds to an "active" choice in a range of the individual domestic share of invested pensions funds w according to $\{w < 0.2, \quad 0.2 \le w < 0.4, \quad 0.4 \le w < 0.6, \quad 0.6 \le w < 0.8, w \ge 0.8\}$, and $y = 6$ represents the "passive" default alternative. $\Pr(y = m \mid \mathbf{x})$ is the probability of observing outcome m given $\mathbf{x}$, and is a function of $\mathbf{x}\boldsymbol{\beta}_m$, where $\boldsymbol{\beta}_m$ includes a constant $\beta_{0,m}$ and coefficients $\beta_{k,m}$ for the effect of variable x_k on outcome m. The variables in $\mathbf{x}$ are AGE (in years), INCOME (individual's disposable income in SEK), WEALTH (market value in SEK of financial assets and real estate holdings, net of debt for each individual), MALE (dummy variable equal to 1 if the individual is male), MONEY (total initial investments in pensions funds for each individual in SEK), OCC_2 through OCC_4 are dummy variables for individuals occupation (OCC_2 represents private sector employment, OCC_3 self-employment, OCC_4 unemployment, whereas the base case represents government employment), EDU_2 is a dummy variable equal to 1 if the education level of the individual is high school, EDU_3 is a dummy variable equal to 1 if the education level is above high school (the base case represents education level below high school), EXP is a dummy variable equal to 1 if the individual has other risky holdings, FEE = percentage transactions cost (fee) for the individual investments, STDEV = annualized standard deviation of three-year historical monthly returns for the individual portfolio, RETRES = the residual from the regression of RETURN on STDEV, where RETURN = the compounded annual growth rate of return for the three years 1997 through 1999, MARRIED is a dummy variable equal to 1 if the individual is married, and FUND_1 through FUND_4 are dummy variables reflecting the number of mutual funds the individual has invested in, 1 through 4 (the base case is an investment in 5 funds). The probabilities are modeled as:

$$\Pr(y_i = m \mid \mathbf{x}_i) = \frac{1}{1 + \sum_{j=2}^{J} \exp(\mathbf{x}_i \boldsymbol{\beta}_j)} \quad \text{for } m = 1, \qquad \Pr(y_i = m \mid \mathbf{x}_i) = \frac{\exp(\mathbf{x}_i \boldsymbol{\beta}_m)}{1 + \sum_{j=2}^{J} \exp(\mathbf{x}_i \boldsymbol{\beta}_j)} \quad \text{for } m > 1$$

where $\boldsymbol{\beta}_1 = 0$ for the first outcome. The model is estimated using the maximum likelihood technique outlined in Berndt et al. (1974), and the heteroskedasticity-consistent covariance matrix according to White (1980). The estimated coefficients are presented for each probability and explanatory variable, with p-values in parentheses. Each χ^2-statistic results from a Wald test for the hypothesis that each explanatory variable does not affect the likelihoods of outcomes $y = 2$ through $y = 5$, relative the first outcome $y = 1$, and is χ^2-distributed with four degrees of freedom.

The null hypothesis that the OCC_2 coefficients in the probability equations for outcomes 2 through 5 are jointly equal to zero is rejected at any reasonable significance level. Furthermore, privately employed individuals have a significantly lower probability of choosing the outcomes 2 and 3, relative to outcome 1; i.e. people in the private sector have a lower likelihood for choosing funds with a large share of domestic securities relative to people employed by the government. Hence, there is a relatively higher likelihood for home bias among individuals in government employment. This result is in line with the idea that relative home bias is an effect of individuals' different desires to hedge inflation using domestic assets, due to a difference in job security. The result is also consistent with Borjas (2003) argument of home-biased individuals' self-selection into government employment. Also, we find evidence that privately employed people have a relatively lower likelihood of choosing the default alternative than government employees, as the

coefficient for OCC_2 is significantly negative at the five percent level, in the default equation.

Regarding investor sophistication, experience of holding risky assets (EXP) is significantly important for the choice of domestic share of funds. People who hold risky assets outside the pension system have a relatively lower probability of choosing outcomes 3 through 5. The coefficients for the EXP variable are increasingly more negative when compared across outcomes with an increasing domestic share. Hence, individuals experienced with risky investments have a relatively lower likelihood of home bias. Also, experienced individuals have a lower likelihood of a default choice, as the coefficient for EXP is significantly negative in the default equation.

The education level contributes significantly to the model, as the Wald test for the dummy variable EDU_1 rejects the null hypothesis that the set of coefficients is jointly equal to zero. However, the corresponding Wald test for the EDU_3 variable leads to no rejection. The coefficients for the EDU_1 variable are significantly positive at the five percent level in the equations for Pr(y = 2), Pr(y = 3) and Pr(y = 4), and increasingly so when comparing outcomes 2 and 3, and 3 and 4 respectively. Relative to individuals with high school education, less educated (less sophisticated) individuals have a higher likelihood of home bias. Also, people with less than high school education are significantly more likely to choose the default alternative.

The amount of money invested in the pension funds (MONEY) contributes significantly to the model. The null hypothesis that the variable MONEY can be excluded without loss of information is rejected at any reasonable significance level, and all coefficients for the variable MONEY are significantly negative. Also, for the probabilities of the outcomes y = 2 through 5, the MONEY effect appears to be increasingly more negative the higher the domestic share. The results indicate that

individuals who allocate a small amount of money into the pension system show a relatively higher likelihood of home bias. Hence, if we accept MONEY as a proxy for investor sophistication, then increased investor sophistication leads to a lower likelihood of home bias. We also note a lower likelihood of choosing the default alternative for a higher amount of money. This effect is captured by the significant coefficient for MONEY in the Pr(y = 6) equation, and supports the notion that a large amount invested in the pension scheme spurs investor activity.

Total net wealth (WEALTH) is a significant explanatory variable for the probabilities of choosing the different categories, at the five percent significance level, whereas disposable income (INCOME) is not. We see that wealthier individuals have a slightly lower likelihood of home bias, as they are significantly less likely to choose outcomes y = 2 and 3, relative y = 1. Neither wealth nor income is significant in explaining the likelihood of the default choice.

The inclusion of the MALE dummy variable in the model results in a Wald test statistic significantly different from zero. In particular, the coefficient in the Pr(y = 5) equation is significantly positive, implying that men have a higher likelihood of choosing outcome 5, and thus having a higher likelihood of home bias than women. Given the idea that an overconfident male investor perceives that he has an information advantage regarding the domestic assets, he will overweight this asset class and have a higher likelihood of home bias. The coefficient for the MALE dummy in the $\Pr(y = 6)$ equation is not significantly different from zero, which means that men and women are equally likely to choose the default alternative. This result contradicts Engström and Westerberg (2003) who find that men are less likely to make an active choice in the Swedish pension system.[20] Since we use a more

[20] Studies by Sundén and Surette (1998) and Madrian and Shea (2001) find the exact opposite result, that men are more likely to participate in defined contribution plans.

general model for investors' choice, we believe that our results are more robust relative to Engström and Westerberg (2003).

The Wald test statistics associated with the dummy variable for marriage (MARRIED) and the interaction term between marriage and gender (MAR_MALE) imply that marriage is a significant source of explanation for the proportion of chosen domestic funds. For the MARRIED variable, the coefficients are negatively significant in all probability equations except for $\Pr(y = 5)$. Hence, a married individual is more likely to have a relatively low proportion of domestic assets. The coefficient for the interaction variable (MAR_MALE) is significantly positive in the $\Pr(y = 2)$ equation, indicating that marriage has a different effect on men than women when choosing the domestic share. However, the most prominent gender difference, the significantly higher likelihood for men to choose outcome $y = 5$, is not affected by marriage. In the $\Pr(y = 5)$ equation, the coefficients for the marriage variable and the interaction term between gender and marriage are not significantly different from zero. Thus, we find no evidence that marital status, as in Barber and Odean (2001), weakens the gender overconfidence effect that men are more home-biased than women. Also, married individuals are less likely to end up in the default alternative, as the corresponding coefficient in the probability equation for outcome $y = 6$ is significantly negative.

Panel B of Table 3 contains the results for the control variables. The risk level of the individual portfolios (STDEV) is significant in the model. All coefficients for the STDEV variable are negative, indicating that all probabilities, relative $\Pr(y = 1)$, are decreasing in the risk level. Using the Wald test statistic, we reject the null hypothesis that portfolio standard deviation is not an important control variable at the five percent significance level. Similarly, the null hypothesis of excluding the return residual variable (RETRES), i.e. the part of the historical return that is

orthogonal to level of risk, is rejected at all reasonable significance levels. Consequently, it is also important to control for the return aspect.

The number of funds chosen by the individuals is important for the model. From the base case, individuals who invest in five funds, the dummy variables representing fewer funds (FUND_1 through FUND_4) are all significant according to the Wald tests. Most individual coefficients are significantly positive. For instance, people with one fund have a higher likelihood of home bias than people who invest in five funds. The fund fees (FEE) add significantly to the model. All coefficients for the variable FEE are significantly negative, except for $\Pr(y = 2)$. Thus, portfolios containing a large share of international assets are more costly than portfolios with predominantly domestic assets. Hence, it is important to control for transactions costs.

Age (AGE) is a significant source of explanation for the probabilities of choosing a certain domestic share. Using the Wald test statistic we reject the null hypothesis that the variable AGE does not affect the allocation between domestic and foreign assets at all reasonable significance levels. The probabilities $\Pr(y = 2)$ and $\Pr(y = 3)$ are associated with significantly negative coefficients with respect to age, whereas the corresponding coefficients for $\Pr(y = 4)$ and $\Pr(y = 5)$ are significantly positive. Hence, an increase in age increases the likelihood of home bias. The likelihood of choosing the default alternative is unrelated to age, as the coefficient for AGE in the equation for $\Pr(y = 6)$ is not significantly different from zero.

5. Concluding remarks

In the year 2000, the entire work force in Sweden was obligated to allocate 2.5 percent of their income from the previous four years into a defined contribution pension scheme. Each individual was provided with a choice of 464 mutual funds, of which a maximum of five could be combined into a portfolio. Information

regarding risk level, transactions costs, management and investment allocation of each fund was presented in a brochure distributed to all participating individuals. Hence, each individual could allocate her/his portfolio of mutual funds with the desired proportions of domestic and foreign assets.

Our analysis of the home bias issue in this particular context is unique with respect to previous research. The portfolio formation within the Swedish pension system constitutes a setting where barriers to international investments and information asymmetries are virtually non-existing, and where the investments contain no corporate control dimension. In addition, given the detailed set of data, we can investigate what characterizes an individual who is relatively reluctant towards international investments. Moreover, in this setting we are able to control for reasons for home bias found in previous studies, as e.g. inflation hedging (Adler and Dumas, 1983), investor sophistication (Grinblatt and Keloharju, 2001) and overconfidence (Barber and Odean, 2001).

The empirical analysis uncovers several significant relationships between individual characteristics and the likelihood of home bias. First, a higher tendency of home bias is found for individuals employed in the public sector than for individuals who work in the private sector. We argue that a high level of job security (being employed in the public sector) causes less concern for a job loss, which in turn causes less need for international diversification and more concern for hedging domestic purchasing power. Second, previous experience with risky investments, a higher level of education and a large amount of money invested in the pension scheme all indicate a lower likelihood of home bias. These results are consistent with the idea that investor sophistication reduces the likelihood for home bias. Third, men have a tendency to be relatively more home-biased than women. Previous research states that an overconfident individual will have a perceived information advantage of the asset class he or she is familiar with, and therefore will overweight that asset class.

The only asset classes we can assume is familiar to all investors in our sample are the domestic funds, all investors being born in Sweden. Home bias is thus arguably an overconfidence issue. The fact that men have a higher likelihood of being home-biased than women is consistent with previous research reporting men to be more overconfident than women.

Several of our control variables have significant impacts on the likelihood of being home biased. Our results show that portfolio risk and return, the number of chosen funds, fund fees, and individuals' age and marital status all are important control variables. From the results, we can identify a typical home-biased candidate. It would be an older, unmarried, poorly educated man working for the government, who only invests a small amount of money, and who has no previous experience with investments in risky assets prior to his participation in the defined contribution pension plan. Using our theoretical framework we would describe him as a not so sophisticated investor, who has a high level of job security and, arguably, seems to be somewhat overconfident.

Appendix A: The multinomial logit model

Define *w* as the individual domestic share of invested funds. The domestic share *w* is divided into five categories, and a sixth category is added, representing the default alternative. Let *y* be a nominal variable with $J = 6$ categories or outcomes defined as $y = 1$ if $w < 0.2$, $y = 2$ if $0.2 \le w < 0.4$, $y = 3$ if $0.4 \le w < 0.6$, $y = 4$ if $0.6 \le w < 0.8$, $y = 5$ if $w \ge 0.8$, and $y = 6$ to represent the default. Let $\Pr(y = m \mid \mathbf{x})$, m = 1, …, 6, be the conditional probability of observing the outcome *m* given the explanatory variables **x**. Further, assume that $\Pr(y = m \mid \mathbf{x})$ is a function of $\mathbf{x}\boldsymbol{\beta}_m$, where $\boldsymbol{\beta}_m = (\beta_{0,m} \cdots \beta_{k,m} \cdots \beta_{K,m})'$ includes an intercept term $\beta_{0,m}$ and coefficients $\beta_{k,m}$ for the effect of variable x_k on outcome *m*. Following Theil (1969), and using the notation in Long (1997), the probabilities for the *i*th individual can be written as:

$$\Pr(y_i = m \mid \mathbf{x}_i) = \frac{1}{1 + \sum_{j=2}^{J} \exp(\mathbf{x}_i \boldsymbol{\beta}_j)} \qquad \text{for } m = 1 \tag{A.1}$$

$$\Pr(y_i = m \mid \mathbf{x}_i) = \frac{\exp(\mathbf{x}_i \boldsymbol{\beta}_m)}{1 + \sum_{j=2}^{J} \exp(\mathbf{x}_i \boldsymbol{\beta}_j)} \qquad \text{for } m > 1$$

The constraint $\boldsymbol{\beta}_1 = 0$ for $m = 1$ is made to ensure that the probabilities are identifiable.

To interpret the coefficients the multinomial logit model is expressed in terms of odds, or relative probabilities of different outcomes. The odds of outcome *m* versus outcome *n*, given **x**, is:

$$\Omega_{m|n}(\mathbf{x}_i) = \frac{\Pr(y_i = m \mid \mathbf{x}_i)}{\Pr(y_i = n \mid \mathbf{x}_i)} = \frac{\exp(\mathbf{x}_i \boldsymbol{\beta}_m)}{\exp(\mathbf{x}_i \boldsymbol{\beta}_n)} = \exp\left[\mathbf{x}_i (\boldsymbol{\beta}_m - \boldsymbol{\beta}_n)\right] \tag{A.2}$$

or by taking natural logs in Equation (A.2):

$$\ln \Omega_{m|n}(\mathbf{x}_i) = \mathbf{x}_i (\boldsymbol{\beta}_m - \boldsymbol{\beta}_n) \tag{A.3}$$

The partial derivative of the log-odds in Equation (A.3) with respect to variable x_k equals:

$$\frac{\Delta \ln \Omega_{m|n}(\mathbf{x}_i)}{\Delta x_k} = \beta_{k,m} - \beta_{k,n} \tag{A.4}$$

The difference $\beta_{k,m} - \beta_{k,n}$ can be interpreted as a contrast, as for a unit change in x_k; the logit of outcome *m* versus outcome *n* is expected to change with the magnitude of this difference. Also, since $\boldsymbol{\beta}_1 = 0$, each coefficient $\beta_{k,m}$ can be interpreted as the effect of x_k on the log-odds of outcome *m* relative to the first outcome.

Chapter 3. The Homeboy Bias: Evidence For and Determinants Of

Anders Karlsson and Grant McQueen

1. Introduction

We document and study a new investor preference that we call the homeboy bias. Homeboy is a slang term, from the Hip-Hop culture of North America, for a close friend from ones' own neighborhood, hometown, or region.[1] For our purposes, the homeboy bias refers to the tendency of investors to invest with financial institutions from their own country.

At least since Levy and Sarnat (1970) and clearly since French and Poterba (1991), finance researchers have known that investors have a preference for assets in their own country. Coval and Moskowitz (1999) and Huberman (2001) extend this preference for domestic to local assets. Massa and Simonov (2006) suggest that investors are biased toward familiar assets because of they have better information about them; in contrast, Grinblatt and Keloharju (2001) suggest that the bias is based on a familiar language or culture. Whereas the *home* bias is a preference for domestic or local assets, the *homeboy* bias is a preference for domestic institutions. We show that the homeboy bias is empirically distinct from the home bias. For example, when investors seek exposure to international assets, they prefer funds offered by local institutions to funds offered by foreign institutions.

In an article documenting the degree of and the mechanism by which advertising affects an individual's choice of mutual fund, Cronqvist (2006) finds

[1] Homeboy can also be a slang term for a fellow member of a youth gang. Although Homeboy can refer to males, we use the term androgynously.

that mutual funds run by domestic (Swedish) fund families received more money than their foreign-run counterparts. We explore, in greater detail, this homeboy bias by first, presenting its potential causes, second, documenting its size and significance, third, testing potential explanations, and fourth, measuring its determinants.

We posit five potential explanations for the homeboy bias; the first three are economic and the last two are behavioral. First, the preference for domestic institutions could be the consequence of an information advantage their managers have about domestic assets. Second, the selection of domestic managers could be driven by a preference for funds denominated in the domestic currency. Third, the choice of local institutions could be associated with benefits to the local economy in terms of employment and tax revenue. Fourth, the homeboy bias could be another manifestation of the general bias toward the familiar (see Zajonc (1968) and Huberman (2001)). Fifth, the bias toward domestic institutions could be evidence of behaviors such as national identity or pride and xenophobia (see Müller-Peters (1998), Rydgren (2004), and Smith and Kim (2006)). Our tests contradict the first three economic explanations and support the latter two behavioral explanations.

We document the homeboy bias using two compressive data sets from Sweden's introduction of a partially-privatized social security system where participants were required to invest in individually-directed retirement accounts. Approximately half of the mutual funds available in this new retirement system were offered by Swedish institutions. Yet, when we aggregate the choices of all participants, over 96 percent of the money went to funds managed by Swedish institutions. The average fund offered by a foreign institution received only 10.8 million Swedish Crowns (SEK); in contrast, the average Swedish-institution's

fund received 155.4 million SEKs—over 14 times more.[2] Foreign institutions funds were virtually shut out. Of course, the preference for funds from Swedish institutions could arise if they had lower fees, higher past returns, larger size, or a longer history than their foreign counterparts. Our *cross-fund* regressions (416 mutual funds) show that, after controlling for determinants of fund flows such as past return, fees, size, and age, the evidence of a homeboy bias strengthens; a typical Swedish institution's fund received 30 times more money than a similar foreign institution's fund.

We explore the determinants of the homeboy bias in individuals' portfolios using *cross-individual* regressions (17,951 participants in the Swedish retirement system). For this representative sample, we have individual portfolio choices along with demographic information including level of education, employment sector, nationality, and holdings of foreign assets. We perform cross-individual regressions where the dependent variable is a measure of the degree of homeboy biasedness. These regressions find that the homeboy bias is inversely correlated with proxies for financial sophistication and positively correlated with proxies for familiarity and nationalism, suggesting that behavioral explanations, rather than information, foreign exchange exposure, or economic stimulation, are at play.[3]

For five reasons, our Swedish pension data is uniquely suited for modeling the flow of money to mutual funds and testing for a homeboy bias. First, we have the portfolio choices of the complete working population of Sweden rather than a sample. We have investment data on all 4,413,831 participants in 2000 and participation was mandatory. Thus, our data is free from self-selection biases

[2] In late 2000, a typical exchange rate was 10 Swedish Crowns to 1 U.S. Dollar.

[3] Although our focus is on the homeboy bias, the signs and significance of our control variables extend the understanding of how individuals choose mutual funds including insights into the role of past returns, menu effects, and fund family characteristics.

common in data that comes from a financial institution or 401(k) program. Second, since our data corresponds to the initiation of a system, all participants made their choice at approximately the same time. In contrast, most financial data reflect individual choices over time as people enter, exit, rebalance, or make new contributions. Third, all participants chose from the same finite list of funds, were given the same catalog, and had access to the same information on each fund. Fourth, each fund was on equal information, simplicity, and convenience footing with other funds. Each fund had the same coverage in the catalog and online system, reported the same fund and fund-family information. The choices were free from channel bias as all funds were offered and selected in the same way as the other funds. Fifth, for a sample of investors, we have a rich set of demographic and economic variables that document why some individuals have a greater homeboy bias than other individuals. The demographic data were gathered during the same time period in which the initial investment choices were made.

Our article is organized as follows. In Section 2 we present various potential explanations for the homeboy bias. In Section 3 we document the size and significance of the homeboy bias in cross-fund regressions using the choices of all 4.4 million participants. In Section 4 we present tests of the various explanations of the bias. In Section 5, we test for the determinants of the homeboy bias using a sample of participants in cross-individual regressions. Section 6 concludes.

2. Potential Theories for Homeboy Bias

An obvious reason for a Swede to pick a local institution to manage their money is ease; it is more convenient to deal with an institution that has nearby offices and employees that speak your language. However, in the Swedish

system all funds—even those offered by foreign institutions—were distributed in the same manner with the same offices and employees. The funds were offered through the Premie Pensions Myndigheten (Premium Retirement Authority, PPM hereafter) a government agency that insured a level playing field for domestic (Swedish) and foreign institutions. Each fund received equal space in the PPM offering catalog and were literally one mouse click or one darkened bubble away. The catalog was a prospectus-like 97 page document giving detailed information about the participating mutual funds and their sponsoring institutions. Although the Swedish Government encouraged participants to make a conscious choice, it did not promote any specific funds nor did it promote domestic-managed funds over foreign-managed funds. Domestic funds did not enjoy any special government guarantees or backing. Foreign-managed funds did not have any hidden 12b1 fees or loads (neither were allowed for funds in the PPM system). Swedish-managed funds did not get an early start and funds were listed alphabetically, by style, in the catalog. Post-investment reporting on performance, service, support for individuals, and opportunities for reallocation, deposits, and withdrawals all went through the PPM system and followed identical procedures for funds sponsored by foreign and domestic institutions. Since the system itself rules out the convenience and ease explanation, we posit four alternatives.

First, the homeboy bias could be the consequence of asymmetric information. Karlsson and Nordén (2006) show that Swedes, like investors in other countries, have a home country bias. Consequently, Swedes may prefer domestic institutions whose managers presumably know more about the expected return and covariance (with other assets and non-financial income) of domestic assets than their foreign counterparts.

Second, the homeboy bias could be driven by a preference for funds denominated in Swedish Crowns (SEK). Since Swedes' consumption is primarily in SEK, investments denominated in foreign currencies expose consumers to foreign-exchange risk, a risk with no commensurate reward and a risk avoided by choosing funds denominated in SEKs. Since funds sponsored by Swedish institutions tend to be denominated in SEKs, a spurious correlation between the amounts of money a fund receives and the nationality of the fund could result from the underlying correlation between Swedish institutions and funds denominated in SEKs.

Third, participants could prefer domestic institutions because of the benefits to the local economy. If a locally-managed fund receives additional deposits, the fund may hire extra employees, increase purchases from local vendors, and pay more income, property, and sales tax. Such improvements to the local economy, presumably, would indirectly increase the participant's own utility.

Fourth, the homeboy bias could be another manifestation of the bias toward the familiar. The behavioral sciences, including Kahneman (1973), have shown that attention is a scarce recourse and that as cognitive misers, humans may rationally take shortcuts when making decisions.[4] Tversky and Kahneman (1974) show that heuristics simplify decisions and that even sensible heuristics can result in systematic errors. Most relevant to our study are the findings of Alba and Chattopadhyay (1985) and Zajonc (1968) dealing with consideration sets and exposure, respectively.

[4] See Pashler and Johnston (1998) for a review of psychological research examining the human limits to cognitive processing capacity. Simon (1955) coined the term "bounded rationality" to describe the limits and costs associated with the time, effort, and mental energy needed to optimize decisions. Miller (1956) documents the limits in humans' ability to process and store information. Wright (1975) shows that these limitations cause individuals to simplify their decision-making process.

Campbell (1969) documents that when making choices, consumers do not consider all alternatives; rather, they focus on a subset that Alba and Chattopadhyay (1985) call the consideration set. Nedungadi (1990) indicates that brands that are accessible to memory because of frequency and recentcy of exposure are most likely to be included in the consideration set. For our purposes, this string of research suggests that participants may not consider all 455 mutual funds; rather, they may limit their consideration to familiar, and consequently, Swedish funds.

Zajonc (1968) finds that an individual's attitude about an object is enhanced by mere repeated exposure to the object. Bornstein (1989, page 265) summarizes Zajonc's original finding as "familiarity leads to liking," then documents the follow-on studies. Huberman (2001) shows that this preference for the familiar carries over to investments. Massa and Simonov (2006) find a bias toward familiar assets that is based on investors having better information about the familiar. Cronqvist (2006) finds that exposure, through advertising and news articles, leads to an increase in the amount of money a mutual fund receives. A propensity for participants to favor familiar mutual funds could cause the homeboy bias if the Swedish funds are more familiar than their foreign counterparts.

Fifth, the homeboy bias could be an economic manifestation of Tajfel (1978 and 1979) and Tajfel and Turner's (1979) social identity theory. Tajfel and Turner suggest that individuals derive psychic utility by being part of a distinctive group such as a successful team, tribe, religion, race, or country and by inflating the abilities and character of one's own group while denigrate those of other groups. This need to belong often results in economic choices that are biased in favor of members of one's own group. Sharma, Shimp, and Shin (1995) list a plethora of publications documenting a consumer bias in favor of

domestically-produced products over their imported counterparts. Foreshadowing our cross-individual regression results, Sharma, Shimp, and Shin (1995) find that nationalistic consumer behavior grows weaker with income and education and stronger with age. Since Müller-Peters (1998) and Smith and Kim (2006) find that the degree of nationalism in Sweden is low relative to other countries, the homeboy bias we find in Sweden may be much stronger in other countries.

3. Macro Evidence of Homeboy Bias

3.1 Swedish Pension System

The Swedish government, like its U.S. counterpart, allowed those retired when the national pension system was instituted to receive benefits even though they did not pay into the system. In 2000, demographic changes toward more retirees and fewer workers forced the Swedish government to make changes to this pay-as-you-go pension program. In the first part of the new system, 16 percent of a worker's annual income funds a common pool that, like the Social Security System, is (a) used to pay current retirees, and (b) is used to determine the worker's future pension benefits. In the second part of the new system, called "Premiepensions" (Premier Pension), workers contribute 2.5 percent of their income into a self-directed 401(k)-like account. Whereas ongoing contributions are 2.5 percent of salary, the initial funding was 2.5 percent of the prior four years of an individual's income.

We obtained data on each participant's initial choice from PPM. 4,413,832 participants made their initial selection of funds in the last 4 months of 2000 with over 60 percent making their choice in November 2000. We aggregated the choices of all participants to calculate the amount of money each fund received and the number of individuals that put money into each fund. Summary statistics

for these two series are reported in Table 1, Panel A. The average fund received 91 million SEK and 22,872 investors. The relatively large standard deviations and the range between the maximum and minimum numbers both suggest substantial variations in a fund's ability to attract money and investors.

Table 1. Summary Statistics for Mutual Funds in Sweden's PPM Catalog

Variable	N	Mean	StdDev	Skew	Max	Min
	Panel A: Data on Fund Allocation from PPM					
Amount Invested in a Fund (Millions of SEK)	416	91.07	225.79	5.36	2,254.85	0.11
Number of Investors Choosing a Fund	416	22,872	55,767	6.0	665,303	36
	Panel B: Data on Funds from the PPM's Catalog					
Domestic Manager = 1, 0 otherwise	416	0.56	0.50	-0.22	1	0
Proportion of Assets Invested in Sweden	416	0.25	0.40	1.10	1.00	0.00
Management Fee (%)	416	0.93	0.42	0.94	3.97	0.15
Return 2000 (% through August)	416	9.51	10.98	2.13	105.23	-16.99
Return 1999 (%)	268	41.39	41.53	1.83	272	-14
Return 1998 (%)	202	12.74	21.46	-0.61	86	-81
Age of Fund (discrete years, censored at 5)	416	2.08	2.05	0.41	5	0
Market Capitalization (millions of SEK)	413	2,808	22,867	15.65	410,913	0
# of Employees in Fund Family	416	2,058	9,034	4.96	48,623	0
Currency SEK = 1, 0 otherwise	416	0.57	0.50	-0.27	1	0
Advertising in 1999 and 2000 (thousands of SEK)	416	18.1	33.1	1.98	159.9	0
Banking Industry Market Share in Sweden	416	3.90	8.40	1.98	30.3	0
Insurance Industry Market Share in Sweden	416	2.95	4.80	1.36	16.24	0

Panel A reports summary statistics measuring the amount of money given to and the number of investors attracted by a fund after all 4.4 million participants made their initial choice in 2000. A typical exchange rate in late 2000 when choices were made was 10 Swedish Crowns (SEK) per 1 U.S. dollar. Panel B reports summary statistics for the active (not dropped from the system) mutual funds in the 2000 catalog.

All participants were provided with a catalog called "Fondkatalog för dit premiepensionsval" (Fund Catalog for the Premium Pension Choice) containing 455 mutual funds. Participants could choose up to 5 funds.[5] These 455 funds were grouped into 4 major categories and 29 different styles as reported in

[5] If no choice was made, participants' money was put into a default fund managed by the government.

Table 2. After the catalog was printed, 39 funds withdrew from the system leaving only 416. Table 2 shows that 56 percent (231 out of 416) of the funds were offered by Swedish institutions. Of the 416 funds, 3 did not report their market capitalization; consequently, we use 413 funds in our analysis.[6]

The catalog reported information about each fund and the institution sponsoring the fund. Table 1, Panel B, reports summary statistics for the mutual funds. We use the name of the fund family to categorize a fund as being offered by a domestic (Swedish) institution. We use the objective of the fund and its style to assign each fund the proportion of assets invested in Sweden; the average fund's exposure to Swedish assets is 25 percent. Management fees averaged just less than 1 percent and ranged from a low of 15 basis points (a bond fund) to a high of 397 basis points (an emerging markets fund). Fund returns in the first 8 months of 2000 averaged 9.51 percent. The average annual return in 1999 was high, 41.39 percent, for two reasons: first, 1999 was a good year for global equity returns and second, fund families apparently selected their better performing funds to be included in the PPM system. Only 268 of the funds were in existence at the beginning of 1999 and only 202 at the beginning of 1998.

The age of the fund is based on the number of complete years of return history and is right-censored at 5 years in the catalog; thus, the reported average fund age was only 2 years. Some funds were new; specifically, created sometime after January 1999 and before September 2000. Fund market capitalizations ranged from 410,913 million SEK (Indocam's Mosais Japanese Equities fund) to brand new funds with no assets under management.

[6] After the catalog was printed, an additional 40 funds were added to the system so that participants could choose between 456 different funds. When the default fund is added, 457 different funds received money in the initial allocations in 2000. We focus our analysis on the 413 funds that were in the initial catalog, did not drop out, and were not missing market capitalization information.

Table 1. Summary Statistics for Mutual Funds in Sweden's PPM Catalog

Variable	N	Mean	StdDev	Skew	Max	Min
	Panel A: Data on Fund Allocation from PPM					
Amount Invested in a Fund (Millions of SEK)	416	91.07	225.79	5.36	2,254.85	0.11
Number of Investors Choosing a Fund	416	22,872	55,767	6.0	665,303	36
	Panel B: Data on Funds from the PPM's Catalog					
Domestic Manager = 1, 0 otherwise	416	0.56	0.50	-0.22	1	0
Proportion of Assets Invested in Sweden	416	0.25	0.40	1.10	1.00	0.00
Management Fee (%)	416	0.93	0.42	0.94	3.97	0.15
Return 2000 (% through August)	416	9.51	10.98	2.13	105.23	-16.99
Return 1999 (%)	268	41.39	41.53	1.83	272	-14
Return 1998 (%)	202	12.74	21.46	-0.61	86	-81
Age of Fund (discrete years, censored at 5)	416	2.08	2.05	0.41	5	0
Market Capitalization (millions of SEK)	413	2,808	22,867	15.65	410,913	0
# of Employees in Fund Family	416	2,058	9,034	4.96	48,623	0
Currency SEK = 1, 0 otherwise	416	0.57	0.50	-0.27	1	0
Advertising in 1999 and 2000 (thousands of SEK)	416	18.1	33.1	1.98	159.9	0
Banking Industry Market Share in Sweden	416	3.90	8.40	1.98	30.3	0
Insurance Industry Market Share in Sweden	416	2.95	4.80	1.36	16.24	0

Panel A reports summary statistics measuring the amount of money given to and the number of investors attracted by a fund after all 4.4 million participants made their initial choice in 2000. A typical exchange rate in late 2000 when choices were made was 10 Swedish Crowns (SEK) per 1 U.S. dollar. Panel B reports summary statistics for the active (not dropped from the system) mutual funds in the 2000 catalog.

Like fund capitalization, family size, as measured by the number of employees, had a wide range.[7] 57 percent of the funds were denominated in Swedish crowns; the others were denominated in Swiss franks, Euros, British pounds, Japanese yen, Norwegian crowns, or US dollars.

Advertising expenditures come from MarketWatch, a Swedish institution that tracks advertising in Sweden's print (newspapers, magazines, journals, billboards, and banners, but not on-line advertising) and broadcast media (television and radio). Advertising expenditures for fund families range from 159.9 million SEK to no advertising at all. For fund families owned or associated with a bank, the market share in the banking industry comes from the Swedish Banking Authority's publication Kreditmarknadsstatistick (Credit Market Statistics) 2000. Banking is highly concentrated with the largest bank, Handelsbanken, having a 30.3 percent market share and the top four banks controlling a collective 86 percent of the industry. For fund families owned or associated with an insurance company, the market share in the insurance industry comes from Sweden's Försäkringsförbundet (Insurance Association) Quarterly Statistics for the third quarter of 2000. The shares are based on premiums in the life and property and casualty segments for the first 9 months of 2000.

Whereas Tables 1 and 2 describe the population of potential *choices* (mutual funds), Table 3 describes the population of *choosers* (individuals). Table 3, Panel A, reports summary statistics on all 4.4 million individuals who made their first allocation in 2000; in Panel B we describe the 2.8 million individuals who did not end up in the default fund. We treat the default alternative as an

[7] Several fund families actually reported zero employees.

entirely passive choice even though an individual could have considered the default fund to be the optimal choice.[8]

Table 3 Panel A reports that the average amount invested by an individual was 12,651 SEK.

Table 3. Summary Statistics for Individuals Investing in Mutual Funds in 2000

Variable	Mean	StdDev	Skew	Max	Min
	Panel A: Data on all 4,413,831 Participants				
Amount Invested by Individual (SEK)	12,651	6,727	0.142	26,202	167
Age of Individual in 2000	42.1	11.3	-0.011	62	18
Proportion Allocated to Domestic Institutions	0.964	0.120	-4.353	1	0
Number of Funds Held by Individual	2.6	1.6	0.385	5	1
	Panel B: Data on the 2,863,711 Non-Defaulters				
Amount Invested by Individual (SEK)	13,506	6,580	0.055	26,202	167
Age of Individual in 2000	42.0	11.0	-0.011	62	18
Proportion Allocated to Domestic Institutions	0.945	0.145	-3.381	1	0
Proportion Allocated to Swedish Assets	0.343	0.248	0.467	1	0
Number of Funds Chosen by Individual	3.4	1.4	-0.390	5	1

Panel A reports summary statistics for all participants that invested money in Sweden's Premium Pension plan during 2000. Panel B summarizes the participants who did not end up in the default fund (made a conscious choice). The Proportion Allocated to Swedish Assets is not available in Panel A because the asset allocation of the default fund was not known in 2000.

Participants ranged in age from 18 to 62 years, with the average participant being 42 years old. The average individual had over 96 percent of their money in funds managed by Swedish institutions and ended up with 2.6 funds. Panel B of Table 3 indicates that participants who made a conscious choice tended to have slightly higher incomes (as measured by the amount invested) and were slightly younger than those in the default fund. Since the default fund was managed by a Swedish institution, the proportion allocated to domestic managers in Panel B is less than in Panel A by definition. However, even among choosers, the homeboy

[8] For a detailed analysis of default investors, see Engström and Westerberg (2003). The high proportion of individuals in the default fund is consistent with the findings of Choi, Laibson, Madrian, and Metrick (2001) who find that many 401(k) participants take "the path of least resistance."

bias is evident; over 94 percent of their money went to Swedish institutions. The home *country* bias is also evident in Table 3, with the typical Swede investing 34.3 percent of their money in Swedish assets. Consistent with Huberman and Jiang (2006), who found that U.S. 401(k) participants typically choose between 3 and 4 funds, Swedes chose 3.4 funds, on average.[9]

3.2 Variable Descriptions

To measure the size and significance of the homeboy bias, we report cross-fund regressions where the dependent variable is a measure of participants' preference for a fund and one of the independent variables measures whether the fund is offered by a domestic or foreign institution. We use two measures of preference: one based on how much money each fund received and the other based on the number of individuals that chose the fund. The following four dependent variables are transformed using natural logs to help correct for the skewness reported in Table 1.

Ln(AmtMM) = Natural log of the amount, in millions of SEK, the mutual fund received from investors in their initial allocation of money in 2000.

RelAmt = *Ln(AmtMM/MeanAmtMM)*, the relative amount the mutual fund received, where *MeanAmtMM* is the mean amount of money received by funds with the same style. *RelAmt* is approximately the percent above or below the average amount received by funds of the same style.

[9] 35.1 percent of the participants invested in the default fund. 9.5 percent choose only 1 fund. The proportion choosing 2, 3, 4, and 5 funds were, 8.8, 13.4, 13.2, and 20.3 percent, respectively. Unlike the finding of Huberman and Jiang (2006), an even distribution across funds (1/*n*) was less common (35 percent of the individuals choosing 2 or more funds) than those who chose an uneven allocation (65 percent).

Ln(Investors) = Natural log of the number of investors who chose the fund in their initial allocation of money in 2000.

RelInvestors = *Ln(Investors/MeanInvestors)*, the relative number of investors the mutual fund received, where *MeanInvestors* is the mean number of investors who chose the funds with the same style.

To test for the homeboy bias, we not only need a measure for domestic institutions, but also measures of other covariates believed to influence the investors' choice of mutual funds. We define our control variables for past performance, fund characteristics, fund family characteristics, and behavioral preferences as follows:

Measure of the homeboy bias:

Domestic = 1 if the fund is managed by a Swedish institution and zero otherwise. The catalog reported the name of each fund and the fund family; we use the name of the fund family to categorize funds as domestic or foreign.

Past performance measures:

ExRet00 = Excess return (nominal return less *MeanRet00*) in the first 8 months of 2000. Participants started investing in September of 2000. Monthly returns for 2000 were not in the catalog; however, they were available on-line.

MeanRet00 = The mean return for the first 8 months in 2000 for funds of the same style.

ExRet99 = Excess percent return for the fund in 1999. Newer funds that did not exist in 1999 received an excess return of zero. Data on a fund's total return in 1999 is reported, to the nearest percent, in the catalog.

MeanRet99 = The mean percent return in 1999 for funds of the same style.

ExRet98 = Same definition as *ExRet99* except for 1998.

MeanRet98 = Same definition as *MeanRet99* except for 1998.[10]

Fund characteristics:

RelFee = Natural log of the relative management fee, *Ln(Fee/MeanFee),* where *MeanFee* is the mean fee of funds with the same style. *RelFee* is approximately the percentage above or below the average fee for funds of the same style.

[10] The catalog included annual percent returns for 1995, 1996, and 1997. Our regressions generally found that the coefficients on these older returns were insignificant (but typically positive); consequently, we do not report regressions using the older return data in the paper. One explanation for the low significance of the older returns is the lack of recentcy (see Cromwell (1950) and Duncan and Murdock (2000)). An additional explanation is that the sample size decreases as one looks further into the past. The catalog also included two measures of risk: the standard deviation of monthly returns from 1997 to 1999 and ordinal risk categories (1 = lowest risk to 5 = highest risk). Although higher risk funds tended to get more money than their lower risk counterparts, the coefficients on risk were often not significant; consequently, measures of risk are not included in our regressions.

FundAge = Binary variables indicating a funds age. The catalog reports historical returns which indicate the age of the fund up to 5 years and rounded to the nearest year. Since *FundAge* is right-censored and discrete, we create the following binary variables: *FundAge0*, *FundAge1_2*, *FundAge3_4*, and *FundAge5*, where, for example, *FundAge1_2* = 1 if the catalog reports the fund as having 1 or 2 years of return history, and 0 otherwise. We drop *FundAge0* from our regressions so that coefficients on the remaining age-related binary variables are interpreted as the increase in money or investors relative to a new fund.

Ln(FundCap) = Natural log of 1 plus a fund's market capitalization measured in SEK as of December 31, 1999. Information on *FundCap* is reported in the catalog. The log transformation helps to correct for the skewness in the raw data reported in Table 1. We add 1 before taking logs because several funds were completely new and, consequently, had no money under management when the catalog was printed.

Fund family characteristics:[11]

Ln(NumEmp) = Natural log of 1 plus the number of employees for the fund family as reported in the catalog. We add 1 employee before taking

[11] When recording fund characteristics associated with the fund family (size, advertising, and market share), we combined the measures of related families. For example we combine the Robur with the Föreningssparbanken families (the companies had merged) and we combine the three divisions of the Folksam Company into one family. However, we did not combine related families if they reported different home countries. For example we did not combine Carlson Fondförvalting with Carlson Fund Management Company because the former reports Sweden and the latter reports Luxemburg as its country of registrations.

logs because several fund families actually reported zero employees.[12]

Behavioral variables:

Rank1 = 1 if the fund is listed first among competitors of the same style in the catalog and zero otherwise.

Swedish = Proportion of assets in the fund that are invested in Sweden as estimated from the fund's description.

Ln(Ads) = Natural log of the amount (in millions of SEK) spent by a fund's family and its parent company on advertisements in 1999 and 2000.

BShare = Market share in the banking industry based on an average of assets, loans, and deposits shares in 2000 for the mutual fund family's parent or associated bank.[13]

IShare = Market share in the insurance industry based on premiums in the first three quarters of 2000 for the mutual fund family's parent or associated insurance company.

[12] Following Cronqvist (2006) or Karlsson, Massa, and Simonov (2006), we initially included the number of funds that a family had in the PPM system (typically negative coefficient), the number of styles offered by a family (typically positive coefficient), and the number of funds within the style (typically negative coefficient) as independent variables. The coefficients on these additional variables were, at best, marginally significant and were not robust; consequently, we do not include them in our reported regressions.

[13] For the UBS fund family we use bank market share from 1999 and for Banque Invik we use the market share from 2001 because these two banks were not included in the 2000 statistics.

Style control variables:

Style = Binary variables indicating the funds style as reported in Table 2. We exclude style 1, Swedish normal equities, from regressions so the coefficients on style binary variables 2 to 29 represent styles' ability to attract money and investors relative to Swedish equities.

3.3 Homeboy Bias Test Results

Table 4 reports the estimated coefficients, p-values, and adjusted R^2 for six different specifications of our regression model. In the first three specifications the dependent variables are based on the amount of *money* a fund received and in the last three specifications the dependent variables are based on the number of *investors* a fund attracted. For each of these two dependent variables, money and investors, we estimate three specifications that differ in how they control for the 29 styles. Potentially, the domestically-managed funds could have received more money or investors than their foreign counterparts because they were in relatively popular styles—styles with high demand but few suppliers. In the first empirical specification, we control for styles by including each style's average return for 2000, 1999, and 1998.[14] In the second specification, we include 28 style dummy variables. In the third specification, we use the relative (to funds of the same style) amount of money or number of investors as the dependent variable.

Our particular interests are the *Domestic* coefficients in the first row of Table 4. The coefficients are positive and significant (p-values less than 0.0001) in each of the six regressions.

[14] None of the 7 funds in Style 8, European index, existed on January 1, 1998; consequently, *MeanRet98* is missing and the sample size falls to 406.

Table 4. Determinants of a Mutual Fund's Ability to Attract Money and Investors

	---------- Model Number and Dependent Variable ----------					
Independent Variables	*1 ln(AmtMM)*	*2 ln(AmtMM)*	*3 RelAmt*	*4 ln(Investors)*	*5 ln(Investors)*	*6 RelInvestors*
Domestic	3.46	2.87	2.55	3.21	2.80	2.50
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ExRet00	0.028	0.027	0.024	0.028	0.027	0.024
	(0.001)	(0.000)	(0.001)	(0.000)	(0.000)	(0.001)
MeanRet00	0.077	—	—	0.075	—	—
	(0.000)			(0.000)		
ExRet99	0.018	0.018	0.019	0.018	0.018	0.018
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
MeanRet99	0.014	—	—	0.015	—	—
	(0.000)			(0.000)		
ExRet98	0.009	0.006	0.006	0.009	0.006	0.006
	(0.087)	(0.183)	(0.299)	(0.089)	(0.179)	(0.281)
MeanRet98	0.000	—	—	0.000	—	—
	(0.955)			(0.865)		
RelFee	-1.18	–1.49	–1.55	-1.23	–1.45	–1.50
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
FundAge 1_2	0.28	0.19	0.28	0.28	0.18	0.28
	(0.101)	(0.216)	(0.106)	(0.070)	(0.211)	(0.088)
FundAge 3_4	0.78	0.70	1.08	0.84	0.74	1.10
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
FundAge 5	1.10	1.23	1.43	1.19	1.28	1.46
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ln(FundCap)	0.07	0.12	0.08	0.07	0.12	0.07
	(0.024)	(0.000)	(0.006)	(0.007)	(0.000)	(0.012)
ln(NumEmp)	0.23	0.21	0.17	0.21	0.20	0.17
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Rank1	0.41	0.37	0.48	0.30	0.29	0.40
	(0.074)	(0.059)	(0.034)	(0.148)	(0.125)	(0.064)
Swedish	–0.44	—	–0.90	–0.38	—	–0.89
	(.012)		(0.000)	(0.018)		(0.000)
Style Dummies	No	Yes	No	No	Yes	No
Constant	–2.46	–0.85	–4.29	3.40	4.94	–4.13
	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)
Adjusted R^2	0.737	0.802	0.573	0.757	0.800	0.592
N	406	413	413	406	413	413

Entries in each cell are regression coefficients and p-values are in parenthesis. The dependent variables are the natural log of the amount of money measured in millions of SEKs, *AmtMM*, received by the mutual fund for regressions 1 and 2, the relative (to funds with a similar style) amount of money, *RelAmt = ln(AmtMM/MeanAmtMM)*, in regression 3, the natural log of the number of investors, *Investors*, attracted by a fund in regressions 4 and 5, and the relative number of investors, *RelInvestors= ln(Investors/MeanInvestors)*, in Regression 6, where *MeanAmt* is the mean amount of money received by funds of the same style and *MeanInv* is the mean number of investors choosing funds of the same style. *Domestic* = 1 if the fund is managed by a Swedish institution and zero otherwise. *ExRet00* = excess return (relative to funds with the same style) for the fund in the first 8 months of 2000. *MeanRet00* = the mean return for the first 8 months in 2000 for funds of the same style. *ExRetYY* = excess percent return for the fund in year 19YY. Newer funds that did not exist in 19YY received an excess return of zero. *MeanRetYY* = the mean percent return in 19YY for funds of the same style. *RelFee* = natural log of the relative management fee, *Ln(Fee/MeanFee)*, where *MeanFee* is the mean fee for funds of the same style. *FundAge1_2*, *FundAge3_4*, and *FundAge5* are binary variables for the age of the fund; for

example, *FundAge1_2* = 1 if the catalog reports the fund as having 1 or 2 years of return history, and 0 otherwise. *FundAge0* is omitted from the regression. *Ln(FundCap)* = natural log of 1 plus a fund's market capitalization measured in SEK as of December 31, 1999. *Ln(NumEmp)* = natural log of 1 plus the number of employees for the fund family. *Rank1* indicates whether the fund was listed first in its style in the catalog. *Swedish* = proportion of assets in the fund that are invested in Sweden as estimated from the fund's description. *Style* indicates the inclusion of 28 binary variables for the funds style (see Table 1); style 1, Swedish normal equities, is omitted from the regressions.

For the first regression, the coefficient on the domestic variable is 3.46 with a t-statistic of 21.68 (not reported) that is higher than any other t-statistic in the regression. Thus, even after controlling for past returns, fund, fund family, and behavioral characteristics, participants are more likely to choose a fund managed by a domestic institution than a fund managed by its foreign counterpart. The interpretation of the coefficients indicates that a domestic manager is economically meaningful. Because the dependent variables are natural logs, the coefficients can be transformed using: $e^{3.08} = 31.8$. That is, the typical fund managed by a domestic institution received nearly 31 times more money than a similar fund managed by a foreign institution. Two clarifying comments are in order. First, when just comparing raw amounts of money, an average Swedish-managed fund received 14 times more money than the average foreign-managed funds. However, after controlling for returns, fees, age, and size, this multiple jumps to 31 times. Second, if we consider the fact that the default fund was also managed by a Swedish institution, the homeboy bias would be even stronger.

Before discussing the robustness of the *Domestic* coefficient, we comment on the overall significance of the regression, on the signs and significance of the control variables, and on various specifications of the model. Table 1 indicates that one fund received over 2,254 million SEK whereas another only 0.1 million; thus there is a lot of variation for the model to explain. The adjusted R^2 on Model 1 in Table 4 indicates that collectively our covariates explain nearly 74 percent of the variation in the natural log of the money received by each fund. If the *Domestic* binary variable is dropped from the regression, the adjusted R^2 falls

to 42 percent. No other variable, when omitted, causes such a large drop in explanatory power. For comparison purposes, the adjusted R^2 of a regression that omits all six return variables, but maintains *Domestic*, only drops to 59 percent. Formally, the p-value from an F-test restricting the six return coefficients to equal 0 is larger than the p-value from a t-test restricting the domestic manager coefficient to equal 0.

In Model 1, the six coefficients on the past returns indicate that, contrary to the efficient markets theory and evidence that alphas tend to be serially uncorrelated, participants "chased" both hot *styles* and hot *hands*.[15] With regard to hot styles, *MeanRet00* and *MeanRet99* both have positive and significant (p-values less than 0.0001) coefficients indicating that the higher the returns for a particular style leading up to the allocation decision, the more money funds of that style received. The coefficients indicate investors may remember back only two years since the coefficient on *Mean98* (and *Mean95* to *Mean97*, not shown) is not significant. With regards to hot hands, investors apparently also looked at fund alphas, as measured relative to a style benchmark. The coefficients on excess returns are all positive and for 2000 and 1999 they are highly significant (p-values of 0.001 and less than 0.0001, respectively). The *ExRet98* coefficient is also significant, but at the 90 percent confidence level (p-value = 0.087). This weakening may not only be due to lack of recency, but also to the data incompleteness. Since the catalog did not contain 2000 return data, the significant coefficients on *ExRet00* and *MeanRet00* indicate that some participants viewed and considered online data when making their decision.

[15] Hendricks, Patel, and Zeckhauser (1983) document persistence in poor performing mutual funds; Carhart (1997) finds that fees explain some of this persistence. Grinblatt and Titman (1992) and Ibbotson and Goetzmann (1994) find some persistence in positive performance although Brown and Goetzmann (1995) caution that survivorship biases may impede tests of persistence. Sirri and Tufano (1998) show that, consistent with search costs, the relationship between past performance and fund flows is enhanced by marketing.

The negative and significant coefficient on *RelFee* indicate that participants shied away from funds with fees greater than their competitors'. The coefficient on the three fund-age binary variables indicated a monotonic preference for older funds. The coefficient on *Age5* indicates that funds established before 1994 received $e^{1.10} = 3.0$ times more money than a brand new fund with similar characteristics. Due to economies of scale, or perhaps broader awareness and lower search costs, the larger the fund the more money it received. Like fund size, fund family size, as measure by employees, has a significant and positive effect on the amount of money a fund receives.

The positive and significant coefficient on *Rank1* indicates that being listed first in the catalog among funds of the same style results in $100(e^{0.41} - 1) = 51$ percent more money than otherwise. Evidence from psychology and marketing suggests that being first on the list may be an advantage. Lund (1925) first documents the primacy effect, Asch (1946) and Rosnow (1966) document the role of primacy in memory, and Glanzer and Cunitz (1966) find a related serial position effect.[16] Of course, choosing the first on a list is also evidence of what Choi, Laibson, Madrian, and Metrick (2001) call "the path of least resistance" and what Madrian and Shea (2001) call "inertia."

The negative and significant coefficient on *Swedish* does not constitute evidence against the home bias anomaly. Swedes invested 15 to 30 times more money in Swedish *assets* than would be called for by a value-weighted internationally diversified portfolio. International diversification would call for 1 or 2 percent exposure to Swedish assets, yet the average active participant had over 34 percent exposure. That is, Swedes exhibited home country bias by

[16] In addition to testing for a primacy effect, we also tested for two additional menu effects: a recentcy effect (last on the list, see Cromwell (1950)) and a page position effect (first on a page, see Hanssens and Weitz (1980) and Fankel and Solov (1963)). Neither menu effect's coefficient (not reported) was significant.

investing in funds restricted to Swedish assets (see styles 1 to 4, 20, 21, 26, and 27 in Table 2). At the same time, when choosing among, say Nordic or European funds, the Swedes preferred the funds with fewer Swedish assets.[17]

In Model 2 of Table 4, 28 style dummies are included; style 1, Swedish equities, is omitted. Although not reported, the style dummies indicate that funds in styles 17, 18, and 24 (IT/ Communications, Pharmaceuticals, and Retirement in less than 20 years, respectively) received the most money and funds in styles 26 and 28 (Swedish short-term fixed income and European fixed income) received the least. Since Model 2 includes the style binary variables, we exclude the independent variable, *Sweden,* because of multicollinearity—the majority of the styles are defined by the proportion of Swedish assets.

In Model 3, the style dummies are excluded and the dependent variable is the relative amount of money, *RelAmt*. Thus, regression 3 models the percent above or below the typical amount of money received by funds of the same style. In Models 2 and 3, the size of the *Domestic* coefficient decreases relative to Model 1, but the coefficient remains significant.

Models 4, 5, and 6 in Table 4 repeat Models 1, 2, and 3 using the number of investors in a fund in lieu of the amount of money a fund received. Redefining the dependent measure of a fund's success has little effect on the model's coefficients and the change in the dependent variables has no effect on the strong evidence in favor of the homeboy bias.

[17] Potentially, Swedes' preference for domestic institutions could be driven by a relatively low turnover of assets within the funds or by unusually consumer-friendly regulation of Swedish institutions. However, the investors did not have access to turnover data and the turnover and regulation of Swedish institutions was not atypical.

3.4 Robustness of the Homeboy Bias

In Model 1 of Table 4, we made several choices regarding the measurement and inclusion of the dependent variables. Our choices were guided by theory and prior empirical work. We now examine the robustness of the *Domestic* coefficient using Model 3 as the basis case since it had the lowest coefficient, 2.55. That is, we stack the deck against finding robustness by choosing the weakest model to perturb. We change the number of observations by (1) excluding the independent variable measuring market cap that was missing data for three funds, (2) excluding the premix or generational funds, (3) including only equity funds, and (4) including only non-equity funds. These changes cause the *Domestic* coefficient to vary around 2.55 between a low of 2.42 and a high of 3.00 with p-values always less than 0.0001. Switching measurements by (1) using a fund's reported age rather than the age-based binary variables, (2) using the family market cap to measure size rather than the number of employees, (3) using total fee rather than relative fee, (4) dropping the number of employees (because some families reported no employees), and (5) adding a measure of risk as an independent variable causes the *Domestic* coefficient to vary between 2.38 and 2.64 with p-values always below 0.0001. Finally, omitting the 2000 returns (because they were not in the catalog) leaves the coefficient on *Domestic* unchanged at 2.55.

We also modified our definition of a domestically-managed fund. Our definition in Table 4 is based on name recognition. However, the following five Swedish institutions established subsidiaries that were registered outside of Sweden: Carlson Fund Management (2 funds), Carnegie Fund Management (5 funds), Handelsbanken Fondbolag (1 fund), H&Q Fund Management (2 funds),

and Moderna Fonder (3 funds).[18] Defining a domestic fund based on the country of registration, rather than its name, causes the coefficient on *Domestic* to fall slightly from 2.55 to 2.20 with a p-value still less than 0.0001. Given that around 95 percent of the money went to Swedish institutions and that controlling for covariates actually increases the size of the homeboy bias, it is not surprising that the *Domestic* coefficient is robust.

The data in Table 4 documents the homeboy bias as of 2000 when the PPM system started; however, the homeboy bias is still strong. According to the most recent PPM statistics, all ten funds with the most money on December 31, 2006 were offered by Swedish institutions. Nine of the ten funds with the least amount of money were offered by non-Swedish institutions. The one Swedish fund with a relatively low funding was a brand new fund that entered the PPM system during 2006.

3.4 Causes of the Homeboy Bias

We now focus on what the cross-fund regressions tell us about the potential economic explanations (information asymmetry, foreign exchange preference, and local economy stimulation) and the behavioral explanations (familiarity and nationalism).

One potential explanation for the homeboy bias is the home bias itself. If, for whatever reason, participants have a strong preference for domestic assets, they would also have a preference for domestic institutions whose managers presumably would know more than foreigners about Swedish assets. We can reject this explanation by re-estimating the models of Table 4 after conditioning

[18] For example, Carlson Fondförvaltning AB is a Swedish institution that established a subsidiary, Carlson Fund Management Co. that is registered in Luxemburg. The Carlson family offered 11 funds in the PPM system that were registered in Sweden and 2 funds that were registered in Luxemburg.

on the proportion of Swedish assets in the fund, *Swedish*. We use Model 3 as a basis for comparison. Re-estimating Model 3 with just the 285 funds that could hold only foreign assets, the domestic coefficient remains at 2.55. This significant coefficient (p-value less than 0.0001) clearly indicates that the homeboy bias is distinct from the home bias because in a group of funds that were prohibited from holding domestic *assets*--making a home bias impossible--participants still had a strong preference for domestic *institutions*.

We can also rule out the information asymmetry explanation for the homeboy bias. Estimating Model 3 using the 128 funds that had some exposure to Swedish assets results in the coefficient on *Domestic* dropping to 1.43. The drop in coefficients from 2.55, for funds with no domestic assets, to 1.43, for funds with domestic assets, is not consistent with the information asymmetry explanation. An information explanation would predict just the opposite, a bigger bias for homeboys in arenas where the homeboys know more than their foreign competition. In a regression that allows the coefficient on *Domestic* to vary conditional on whether the fund has exposure to any Swedish assets, this interactive coefficient is significantly (p-value less than 0.0001) negative, meaning the homeboy bias is actually statistically weaker when domestic institutions have a potential information advantage.

In Table 5 we report information about the *Domestic* coefficient from regressions of *ln(FundAmtMM)* on fees, past excess returns, and fund age, by style. Although there are 29 different styles of mutual funds, only 16 regressions are of full rank.[19] The domestic coefficient is positive in all 16 regressions and is significant at the 90 percent confidence level in 14 of the regressions.

[19] For example, the first style, Swedish Normal, has 28 funds, but all of them were offered by domestic institutions (*Domestic* = 1) and of the 5 funds in the seventh style, United Kingdom, all were offered by a foreign institution (*Domestic* = 0).

Table 5. Results, by Style, for the Domestic Fund Manager Independent Variable from Regressions of *Ln(RelAmt)* on Fees, Past Excess Returns, Fund Age, and the Domestic Binary Variable.

		----------*Domestic* Coefficient----------			
Fund Style		N	Coefficient	t-value	p-value
Equities:	Nordic countries	9	3.47	4.73	0.009
	Europe	31	2.85	5.99	0.000
	Euroland (EMU members)	8	4.09	3.03	0.056
	Europe small cap	7	4.66	7.13	0.019
	Europe index	7	0.73	0.32	0.779
	North America and USA	23	2.22	4.80	0.001
	Asia and Far East	15	2.71	5.84	0.001
	Global	31	2.91	6.10	0.000
	Emerging markets	17	2.18	4.06	0.002
	Japan	18	2.25	3.84	0.002
	Other countries	13	2.78	2.17	0.062
	IT and communication	15	1.47	3.61	0.005
	Pharmaceuticals	6	2.00	4.58	0.137
	Other industries	14	2.32	4.38	0.002
Mixed:	Foreign equity and fixed income	22	4.93	4.93	0.001
Fixed Income:	Others	15	5.04	9.65	0.000

The independent variable, *Ln(RelAmt)*, is the natural log of the amount of money (in millions of SEK) received by the mutual fund in the first allocation of participants in 2000 divided by the mean amount for funds of the same style. The independent variables are the binary variable indicating a domestic manager, *Domestic*, the relative (to funds of the same style) fee, *RelFee*, the age of the fund (*FundAge*, not the fund age binary variables), and the excess returns in 2000, *ExRet00*, and in 1999, *ExRet99*. The table reports information on the *Domestic* coefficient only.

The two non-significant coefficients are from regressions with few observations: The Europe index style has only 2 funds offered by foreign institutions and the Pharmaceutical style only has 3 foreign-institution funds. Again, counter to an information explanation, the three largest *Domestic* coefficients are for styles where Swedish institutions may be at an information disadvantage: Europe small cap, Foreign equity and fixed income, and Other (non-Swedish) fixed income. Table 5 indicates that the homeboy bias is evident when the home bias is not a

factor, and that participants are most biased when the local manager may be at an information disadvantage.[20]

The bias toward Swedish institutions could be explained by an information asymmetry regarding risk rather than expected return. Investors, with their employment tied to the Swedish economy, may have preferred retirement portfolios with low correlation to Swedish assets. When choosing, for example, international stocks, Swedish managers may have an advantage over foreign managers in creating a portfolio of international stocks with low correlation with the Swedish stock market. However, such was not the case. We calculated the monthly return correlation coefficients relative to the Swedish All Shares Index between January 1997 and August 2000 after sorting by style and then by domestic/foreign institution. For the 16 equity fund styles that had funds managed by both Swedish and foreign institutions, the Swedish-institution funds had higher return correlation coefficients with the Swedish index than their foreign-institution counterparts.

The apparent preference for a domestic manager could actually be driven by a preference for funds denominated in Swedish crowns (SEK). That is, to avoid exposure to foreign-exchange risks, participants chose funds denominated in SEK resulting in a spurious correlation between the amount of money a fund receives and the nationality of the fund family. However, when we estimate Model 3 for the 178 funds denominated in foreign currencies, the domestic coefficient is still positive and significant (p-value leas than 0.0001). Thus, in a

[20] Hindsight suggests that information asymmetry was not a good reason for participants to prefer domestic institutions. We compared the returns on two equally-weighted portfolios created on December 31, 2000. The portfolio of funds from domestic institutions underperformed the portfolio of funds from international institutions from January 2001 to August 2004. The average compound return on the domestic-managed funds was -24.4 percent, whereas, the average for the foreign-managed funds was -19.1 percent. The 5.3 percent difference is statistically significant at the 5 percent level using either pooled variances or unequal variances. That is, if the bias towards domestic institutions was motivated by a belief in their superior skills, the participants have been sorely disappointed.

sample where a local currency preference is eliminated, the homeboy bias still exists.

For four reasons, the economic stimulation argument (jobs for locals and tax revenue from institutions) seems unlikely. First, the 416 funds in the catalog received less than 38 billion SEK, (less than 4 billion U.S. dollars), an amount that would have little material effect on Swedish employment rates or income taxes. Second, although fund managers make investment decisions, government employees at PPM handle processing and reporting to individuals; that is, a large portion of the economic benefits from additional investments into the system are captured locally regardless of where the managing institution resides. Third, the stimulation argument cannot explain why the bias is larger in styles dealing with foreign assets than domestic assets. Presumably, the boost in local employment or tax receipts associated with a fund manager getting a job is independent of whether the manager analyzes Swedish or non-Swedish assets. Fourth, evidence in the following section suggests that the homeboy bias is stronger in rural areas than in urban areas (the three largest cities) where the domestic institutions' employees tend to work and live.

The three homeboy bias explanations based on economic reasoning (information asymmetry, foreign exchange risk, and economic stimulation) are also indirectly contradicted by evidence presented in the next section using cross-individual regressions. Specifically, in the next section we find that the homeboy bias is strongest in participants characterized as financially unsophisticated and inexperienced; the type of individuals more likely to be effected by behavioral, rather than economic, motivation.

The familiarity explanation receives support from two potential proxies for familiarity: size and age. Presumably, older and larger funds and larger fund families are more likely to be household names and to have established

relationships with participants. Thus, the positive coefficients on age and size in Table 4 indicate support for the familiarity explanation.

To further test the familiarity explanation, we follow Cronqvist (2006) and included advertising expenditures, *Ln(Ads)*, and the market share of a parent or sister bank, *BShare*, and insurance company, *IShare*, as additional independent variables.[21] The coefficients (not reported) on the three familiarity proxies are all positive and statistically significant. We illustrate the affects of familiarity using Model 2 from Table 4 since this model is closest to the empirical model used by Cronqvist (2006). The coefficients on *Ln(Ads)*, *BShare*, and *IShare* have p-values of 0.014, 0.002, and 0.045, respectively. After adding the three familiarity proxies, the coefficient on domestic institution drops from 2.55 to 2.21.[22]

Ideally, one would test whether the preference for domestic institutions is driven by more than familiarity by creating a measure of *Domestic* that is orthogonal to the three measures of familiarity. However the data does not allow analysis of this kind because so few of the foreign institutions advertised in Sweden or owned banks and insurance companies in Sweden. Fiduciary International Ireland Limited was the only foreign institution that both advertised in Sweden and placed funds into the PPM system; and they only advertised in 2000 (not 1999) and only spent 65,000 SEK. Only one non-Swedish bank, UBS, both had banking market share in Sweden and offered funds in the PPM system. Thus, the evidence that ads and market share in related industries increased the

[21] In 1999, 26 financial services institutions spent a total of 278 million SEK on advertising and in 2000, 36 institutions spent 461 million. In contrast, from 2001 to 2005, only 117 million SEK was spent on advertising in a typical year. The unusual amount of advertising by financial services in 1999 and 2000 indicates that fund mangers believed that familiarity increased the flow of money to a fund.

[22] A t-test restricting the *Domestic* coefficient to equal zero has a higher p-value that an F-test collectively restricting the coefficients on *FundCap*, *Num Emp*, *FundAge1_2*, *FundAge3_4*, *FundAge5*, *Ads*, *BShare*, and *IShare*.

flow of money into a fund is primarily limited to funds offered by Swedish institutions. More work needs to be done to tell whether *Domestic* is capturing aspects of familiarity other than those proxied for by age, size, advertising, and market share, or whether it is driven by other preferences such as nationalism or xenophobia. For now, the significant familiarity coefficients and the slight drop in the *Domestic* coefficient imply two possible insights. First, investors chose funds they were familiar with and familiarity can be gained by advertising and prior experience (banking or insurance) with an institution. Second, some, but not all, of the reason that Swedes overwhelming invested with Swedish institutions is that they were familiar with them.[23]

The main point of the cross-fund regressions are: (1) the homeboy bias is large, significant, and robust, (2) the economic explanations of information, foreign exchange, and economic stimulation seem unlikely, and (3) behavioral explanations based on familiarity are at least possible.

4. Micro Determinants of Home Bias

4.1 Data

Whereas in Section 3 we presented evidence of the homeboy bias in cross-fund regressions modeling the flow of money to 416 mutual funds, in Section 4 we switch our focus to the determinants of the homeboy bias in regressions modeling the degree of bias in 15,497 individuals. The intention is to relate the share of domestically-managed money (*Homeboy*) to a set of explanatory variables including demographic characteristics related to investor sophistication, familiarity, and nationalism.

[23] Massa and Simonov (2006) would call the significant coefficients on the advertising and market share variables "informational" familiarity, the kind of familiarity that might limit search costs. The significant coefficient on *Domestic* may be an example of what Massa and Simonov call "pure" familiarity; familiarity based on nationality alone.

For these 15,497 individuals, investment choices are linked with individual demographics for the year 2000. Data on individuals come from Statistics Sweden, the Swedish version of the U.S. Census Bureau. Data sources include HEK 2000 (a report on household economics), IoF 2000 (a report on individual and household measures of income), and SUN 2000 (a report on educational status). Data from these three population reports are linked to an in-depth survey of 15,000 households, also made by Statistics Sweden, which represents a cross section of the Swedish population. This survey reports more detailed information, including the amount of foreign assets held by each individual in the households.

From the survey of 15,000 *households*, 17,591 *individuals* map into the PPM data. 2,454 of these individuals were either too young or unemployed and did not make a selection in the PPM system during 2000 leaving 15,497 individuals who invested money. Of these, 5,124 individuals did not make an active investment decision and were assigned to the default fund. The magnitude of the homeboy bias is evident in our sample--83.1 percent of the individuals in the sample invested all their money with domestic institutions (*Homeboy* = 1). In contrast, only 0.4 percent (42 individuals) invested all their money with foreign institutions--even though almost half of the funds were sponsored by foreign institutions and investing with a foreign institution was just as simple and convenient as investing with a domestic institution. The remaining 16.5 percent split their money between domestic and foreign institutions ($0 < Homeboy > 1$).

To test which demographic and economic characteristics lead to a homeboy bias, we define our dependent variable, *Homeboy*, as the weighted (by the amount of money an individual invested in each fund) average of each fund in the portfolio's measure of *Domestic*. For example, if an individual put 3,000 SEK in a domestically-managed fund and 1,000 SEK in a foreign-managed fund,

the level of *Homeboy* would be: ¾(1) + ¼(0) = 0.75; meaning that 75 percent of the money is managed domestically.

Summary statistics for the dependent variable, *Homeboy*, are reported in Table 6. Consistent with prior population evidence, the average individual in our sample allocated 95 percent of their investment to funds run by Swedish institutions.

Table 6. Summary Statistics for the Sample of Swedish Investors That Invested in the PPM System in 2000

Variable	Mean	StdDev	Skew	Max	Min
Homeboy (independent variable)	0.949	0.137	-3.49	1	0
Investor sophistication measures:					
Income (SEK)	202,269	188,121	10.27	6,442,832	-989,783
Wealth (SEK)	444,108	3,174,320	90.08	355,000,000	-30,000,000
Urban = 1 if living in metropolitan area	0.354	0.478	0.629	1	0
Town = 1 if living in small or midsized city	0.356	0.479	0.601	1	0
Rural = 1 if living in the country	0.290	0.454	0.928	1	0
EDL1 = 1 if less than HS education	0.181	0.385	1.655	1	0
EDL2 = 1 if HS diploma	0.513	0.500	-0.053	1	0
EDL3 = 1 if some post-HS education	0.306	0.461	0.844	1	0
PPMTrades, Average number of trades	0.284	0.327	19.526		
Familiarity and nationalism measures:					
ForeignA = 1 if owned foreign assets	0.148	0.355	1.981	1	0
Age in 2000	43.028	11.111	-0.151		
Native = 1 if born in Sweden	0.887	0.316	-2.449	1	0
ImmigrantW = 1 if Married to Immigrant Wife	0.017	0.131	7.386	1	0
ImmigrantH = 1 if Married to Immigrant Husband	0.014	0.116	8.376	1	0
Demographic measures:					
Male = 1 if male	0.509	0.500	-0.036	1	0
Married = 1 if married	0.779	0.415	-1.342	1	0
OCC1 = 1 if employed in public sector	0.287	0.452	0.944	1	0
OCC2 = 1 if employed in private sector	0.558	0.497	-0.234	1	0
OCC3 = 1 if self-employed	0.048	0.213	4.249	1	0
OCC4 = 1 if in unknown employment	0.108	0.310	2.533	1	0
Control variables for portfolio characteristics:					
PastReturn, return for portfolio from 1997 to 1999	1.429	0.519	1.525	6.70	-0.241
StdDev, Weighted average standard deviation	0.179	0.043	2.100	0.77	0.002
AveFee, Weighted average fee	0.734	0.300	0.442	1.98	0
AveCap, Weighted average market cap	3,419	11,178	7.983	208,666	0

Table 6 includes summary statistics for all variables used in the multinomial logit regression. *Homeboy*, the dependent variable, represents the portion of an investors assets managed by a domestic manager. *Income* represents gross income in 2000 including capital gains/losses. *Wealth* represents net wealth in 2000, which is comprised of the market value of all risky + non risk assets (cash) + an assessed value of real estate – market

value of debt. *Urban, Town* and *Rural* are dummy variables indicating whether the investor lives in an urban, town, or rural setting. *EDL 1-3* represents level of education, where 1 = less than High School, 2 = High School and 3 = more than High School. *PPM Trades* represents the average number of trades per year within the PPM system from 2000 to 2004. *ForeignA* is a dummy variable indicating whether an investor holds foreign assets outside the PPM system in 2000. Age is the investor's age in 2000. *Immigrant* = 1 if the investor is not born in Sweden. *ImmigrantW* = 1 if the investor is a Swedish man married with an immigrant woman and *ImmigrantH* = 1 if the investor is a Swedish woman married with an immigrant man, zero otherwise. *Male* = 1 if the investor is male and *Married* = 1 if the investor is married, zero otherwise. *OCC* contains four occupation dummies. *OCC1* = employed in the public sector, *OCC2* = employed in the private sector, *OCC3* = self employed and *OCC4* = unemployed or employment unknown and zero otherwise. *PastReturn* = weighted average past return for each investors initial portfolio. *StdDev* = the annualized standard deviation on the monthly returns for 1997 – 1999 on the initial portfolio of the investor. *AveFee* is the weighted average fee of the funds chosen in the initial portfolio of the investor. *AveCap* = Weighted average market cap of the funds chosen in the initial portfolio of the investor.

To model the homeboy bias, we need measures of other covariates believed to influence investors' choice of domestic manager. Our list of covariates are informed by the work of Karlsson and Nordén (2006), Dhar and Zhu, (2002), and Engström and Westerberg (2003) who test whether sophistication, experience, and other demographic characteristics influence the home bias, the disposition bias, and the degree of active management, respectively.

Our data includes several measures associated with the degree of financial sophistication: wealth, education, big city domicile, and frequent trading). If the homeboy bias is due to economic reasons, one would expect that more financially sophisticated investors would exhibit a stronger preference for domestic institutions than the less financially sophisticated. In contrast, unsophisticated individuals are more likely to choose local institutions for behavioral reasons such as familiarity or nationalism. Presumably, unsophisticated investors have smaller consideration sets (see Alba and Chattopadhyay, (1985)) or have less exposure (see Zajonc (1968)) to foreign institutions. Furthermore, unsophisticated investors may be more likely to derive psychic utility from the social identity associated with nationalism. We have several indirect proxies for an individual's familiarity with international fund families and their degree of national pride. If the homeboy bias is driven by familiarity, then an individual

who already owned foreign investments at the time of the PPM choice should be less homeboy biased than someone who did not own foreign assets. Likewise, a worker in Sweden who is an immigrant or married to an immigrant may be more familiar with international institutions and/or less concerned about national (Swedish) identity. Finally, evidence in Sharma, Shimp, and Shin (1995), Hjerm (2005), and Smith and Kim (2006), among others, suggests that younger individuals (perhaps do to advances in communications, travel, economic unions, common currencies, and the archaic nature of world wars) tend to be more comfortable with international people and institutions and less nationalistic. Following are the definitions of the dependent variables in our cross-individual regressions:

Measures associated with investor sophistication:

Income = Gross income in 2000. *Income* may be negative since our measure includes capital losses.

Wealth = Net wealth in 2000. *Wealth* may be negative since our measure nets out debt.

Urban = 1 if an individual lives in one of Sweden's three largest metropolitan areas (Stockholm, Göteborg, and Malmö) and zero otherwise.

Town = 1 if an individual lives in a city other than Stockholm, Göteborg, or Malmö and zero otherwise. We drop *Town* from our regressions

so that coefficients on the *Urban* and *Rural* are interpreted relative to living in a city.

Rural = 1 if an individual lives in the countryside and zero otherwise.

EDL = Binary variables indicating an individual's educational attainment. *EDL1*, *EDL2*, and *EDL3* indicate whether an individual has less than a high school education, a high school diploma, or more than high school education, respectively. We drop *EDL2* from our regressions so that coefficients on the *EDL1* and *EDL3* are interpreted relative to a high school education.

PPMTrades = Average number of trades per year within the PPM system from 2000 to 2004. Unlike the above independent variables, *PPMTrades* is not measured in 2000 at the time of the portfolio decision, rather is measured in the four years since the decision.

Measures associated with familiarity and nationalism:

ForeignA = 1 if an individual held foreign assets (stocks, fixed income, derivatives, mutual funds, or real estate) outside of the PMM system in 2000 and zero otherwise.

Age = The individual's age in 2000.

Immigrant = 1 if the individual was not born in Sweden and zero otherwise.

ImmigrantW = 1 if a Swedish man is married with an immigrant wife and zero otherwise.

ImmigrantH = 1 if a Swedish woman is married with an immigrant husband and zero otherwise.

Demographic variables:

Male = 1 if the individual is male and zero otherwise.

Married = 1 if the individual is married and zero otherwise.

OCC = Binary variables indicating an individual's occupation. *OCC1 = 1* if the individual is employed in the public sector, *OCC2* =1 if in the private sector, *OCC3* = 1 if self employed, and *OCC4* = 1 if the individual has an unknown occupation. We drop *OCC1* from our regressions so that coefficients on the *OCC2* to *OCC4* are interpreted relative to employment in the public sector.

Control variables for portfolio characteristics:

PastReturn = Historical return on the individual's portfolio of chosen mutual funds from 1997 to 1999. If a fund was not in existence during this period, the fund's return was the average of extant funds of the same style.

StdDev = Annualised standard deviation of an individual's monthly returns on the portfolio of chosen mutual funds from 1997 to 1999. Funds without returns received the average return of funds with the same style.

AveFee = Weighed average fee of the funds chosen by the individual.

AveCap = Weighed average market cap of the funds chosen by the individual.

4.2 Determinants of Homeboy Bias

The distribution of the dependent variable, *Homeboy*, makes standard OLS or Probit/Logit models problematic. On one hand, only 16.5 percent of the individuals chose a degree of *Homeboy* on the continuum between 0 and 1 (some, but not all, domestic institutions), with the remaining 84.5 percent looking like a binary choice model best analysed using Probit or Logit model. On the other hand, a Probit or Logit model would ignore 16.5 percent of the data and have little power due to the fact that only 42 individuals exclusive chose funds offered by foreign institutions. That is, the degree of the homeboy bias is so strong that we have relatively few observations of *Domestic* below 1 and almost no observations at 0.

We test for the determinants of the homeboy bias using a multinomial logit model by dividing the portfolio choice, y, into the following four categories:[24]

y=1 if $0.0 \leq Homeboy < 0.5$

y =2 if $0.5 \leq Homeboy < 1.0$

y =3 if $Homeboy = 1.0$

y =4 if no choice (default fund)

In order to avoid a selection bias, we jointly model the likelihood of homeboy bias and the likelihood of making an active choice. By including the default outcome, we presume that each individual simultaneously considers two investment choices: the choice of whether to be active or passive and the choice of how much to allocate to domestic institutions. We split the sample at homeboy equals 0.5 and 1.0 (rather than, say, at 0.33 and 0.66) because the bias is so strong that at split at 0.5 is need to get more observations in the first bin, y =1.

The results of the estimation of the multinomial logit model are presented in Table 7. For each explanatory variable, three coefficients are estimated. Each coefficient represents the effect of the variable on the probability of obtaining the outcome y = 2, 3, or 4 relative to the probability of obtaining y = 1, i.e., a low level of homeboy bias, $Homeboy < 0.5$. y = 4 is added to control for the selection bias in our sample. We therefore make no attempt to interpret the meaning of its coefficients. Table 7 also reports a Wald test statistic for each explanatory variable, which is $\chi 2$-distributed under the null hypothesis that the variable does not affect the allocation between domestic and foreign institutions.

[24] See Long (1997), Theil (1969), and Greene (2000) for information about the multinomial logit model.

Table 7. Multinomial Logit Model of and Individual's Degree of Homeboy Bias

	Pr(y = 2)	Pr(y = 3)	Pr(y=4)	Wald
Variables associated with investor sophistication:				
Income	-0.0973	-0.1208	-0.0845	8.9265
	0.030	*0.004*	*0.402*	*0.012*
Wealth	-0.0290	-0.0246	-0.0285	3.5023
	0.066	*0.113*	*0.450*	*0.174*
Urban	-0.3673	-0.7299	-0.2218	35.4200
	0.063	*0.000*	*0.646*	*0.000*
Rural	0.0891	0.2456	0.1235	4.4775
	0.707	*0.294*	*0.849*	*0.107*
EDL1, Education less than High School	0.0459	0.4045	0.3257	12.4951
	0.881	*0.175*	*0.658*	*0.002*
EDL3, Education more than High School	-0.9580	-1.2587	-0.7143	58.2465
	0.000	*0.000*	*0.142*	*0.000*
PPMTrades	-0.0373	-0.6497	-	97.9223
	0.780	*0.000*	-	*0.000*
Variables associated with familiarity and nationalism:				
ForeignA	-0.6141	-0.7338	-0.6451	13.5724
	0.004	*0.000*	*0.187*	*0.001*
Age, in 2000	0.0175	0.0339	0.0218	36.0506
	0.039	*0.000*	*0.324*	*0.000*
Immigrant	-0.2036	-0.3721	-0.0104	4.0311
	0.427	*0.138*	*0.986*	*0.133*
ImmigrantW, married to immigrant wife	-1.6077	-2.0603	-1.0161	24.9736
	0.000	*0.000*	*0.219*	*0.000*
ImmigrantH, married to immigrant husband	1.4670	1.1200	0.9292	3.3160
	0.168	*0.288*	*0.856*	*0.191*
Control variables for investor characteristics:				
Male	-0.0199	-0.1824	-0.0653	6.3026
	0.915	*0.322*	*0.891*	*0.043*
Married	0.7271	0.7309	0.0833	13.7511
	0.000	*0.000*	*0.863*	*0.001*
OCC2, Employed in private sector	0.0006	-0.0676	-0.1542	0.8272
	0.998	*0.751*	*0.793*	*0.661*
OCC3, Self employed	-1.1228	-1.2680	-0.6403	13.0626
	0.002	*0.000*	*0.393*	*0.001*
OCC4, Employment unknown	-0.8739	-0.7086	-0.3999	3.7632
	0.054	*0.104*	*0.672*	*0.152*
Control variables for portfolio characteristics:				
PastReturn , return of portfolio from 1997 to 1999	0.1391	0.2812	-	10.4205
	0.234	*0.007*	-	*0.005*
StdDev, Weighted average standard deviation	-3.7862	-6.2865	-	47.7122
	0.005	*0.000*	-	*0.000*
AveFee, Weighted average fee	-1.2372	-2.7382	-	261.7570
	0.000	*0.000*	-	*0.000*
AveCap, Weighted average market cap	-0.0122	-0.0996	-	150.0624
	0.000	*0.000*	-	*0.000*
Constant	8.3681	11.8863	4.1645	-
	0.000	*0.000*	*0.013*	

Table 7 contains results from the estimation of the multinomial logit model. The dependent variable, *y*, has four possible outcomes (*m* = 1, 2, 3, 4), where each of the first three corresponds to an "active" choice in the range of the individual's domestically-managed share of invested pension funds, *Homeboy*, according to {*Homeboy* < 0.5, 0.5 ≤ *Homeboy* <1.0, *Homeboy* = 1.0}, and *y* = 4 represents the "passive" default alternative. The estimated coefficients are presented for each probability and explanatory variable, with p-values below the coefficient. The Wald tests are distributed CHI-square with four degrees of freedom for the null hypothesis that each explanatory variable does not affect the likelihoods of outcomes *y* = 2 through *y* = 4, relative to the first outcome, *y* = 1. *Income* represents gross income in 2000 including capital gains/losses. *Wealth* represents net wealth in 2000, which is comprised of the market value of all risky + non risk assets (cash) + an assessed value of real estate – market value of debt. *Urban, Town* and *Rural* are dummy variables indicating whether the investor lives in an urban, town or rural setting. We drop *Town* in our regressions so the coefficients for *Urban* and *Rural* are interpreted relative to *Town*. EDL *1-3* represents level of education, where 1 = less than High School, 2 = High School and 3 = more than High School. We drop *EDL2* in our regressions so the coefficients for *EDL1* and *EDL3* are interpreted relative to *EDL2. PPM Trades* represents the average number of trades per year within the PPM system from 2000 to 2004. *ForeignA* is a dummy variable indicating whether an investor holds foreign assets outside the PPM system in 2000. Age is the investor's age in 2000. *Immigrant* = 1 if the investor is not born in Sweden. *ImmigrantW* = 1 if the investor is a Swedish man married with an immigrant woman and *ImmigrantH* = 1 if the investor is a Swedish woman married with an immigrant man, zero otherwise. *Male* = 1 if the investor is male and *Married* = 1 if the investor is married, zero otherwise. *OCC* contains four occupation dummies. *OCC1* = employed in the public sector, *OCC2* = employed in the private sector, *OCC3* = self employed and *OCC4* = unemployed or employment unknown and zero otherwise. We drop *OCC1* in our regressions so the coefficients for *OCC 2-4* are interpreted relative to *OCC1*. *PastReturn* = weighted average past return for each investors initial portfolio. *StdDev* = the annualized standard deviation on the monthly returns for 1997 – 1999 on the initial portfolio of the investor. *AveFee* is the weighted average fee of the funds chosen in the initial portfolio of the investor. *AveCap* = Weighted average market cap of the funds chosen in the initial portfolio of the investor.

The multinomial results in Table 7 indicate that investor sophistication measures help explain an individual's choice regarding the proportion of funds allocated to domestic institutions. The Wald tests indicate significance for income levels, urban dwellers, education levels, and trading frequency. Furthermore the signs of the coefficients are in the direction predicted by behavioral explanations—the unsophisticated are more biased. For example, as income increases, individuals become less likely to end up in outcomes 2 and 3 and more likely to have consciously chosen some foreign institutions. Increased wealth also reduces the likelihood of ending up in outcome 2. City dwellers and individuals with post-high school education have relatively lower probability of ending up in outcomes 2 and 3 and higher probability of allocating money to foreign-managed funds. The more an individual trades, the less likely they are to

be completely homeboy biased, $y = 3$.[25] Overall, our proxies for sophistication indicate that financially-sophisticated individuals are more likely to make a conscious choice and less like to choose domestically-managed funds. This negative correlation between the homeboy bias and financial sophistication suggests that behavioural reasons, such as familiarity and nationalism, rather than economic reasons drive the homeboy bias.[26]

The variables associated with familiarity and nationalism generally add to the explanatory power of the model. Specifically, the Wald tests for prior exposure to foreign assets, age, and a Swedish man married with an immigrant, are all statistically significant in the model. If the homeboy bias is caused by an unobserved convenience or simplicity in choosing a domestic- institution fund or even by an information asymmetry or desire to boost the local economy, then the coefficients on *Immigrant*, *ImmigrantW*, *ImmigrantH*, *ForeignA* and *Age* should all be insignificant. In contrast, we find that individuals that may be more familiar with or open to international influences are less likely to be homeboy biased.

The coefficients on *Married*, and *OCC3*, indicate that non-married individuals who are self employed are more likely to be in the least homeboy-biased outcome, $y = 1$, than are investors who are married and employed in the public sector.

The last four variables in Table 7 control for the characteristics of the funds chosen other than the domicile of the sponsoring institutions. Possibly, individuals in, say, outcome $y = 3$ were really expressing a preference for high

[25] Data on the number of trades, as well as the other control variables at the bottom of Table 7, are only available for individuals who made a conscious choice (non-defaulters).

[26] In general, our evidence on sophistication and experience confirm the findings of Karlsson and Nordé (2006), Dhar and Zhu, (2002), and Engström and Westerberg (2003), who find that sophisticated and experienced investors are less biased.

past return, low risk, low fee, small-cap funds, rather than a preference for domestically-managed funds. The inclusion of the last four variables control for this possibility and indicated that individuals with a relatively strong homeboy bias tend to choose funds with high past returns, low risk and fees, and smaller market caps.

Our Table 7 results are consistent with behavioural preferences based on pure familiarity and nationalism—the type of preferences found by Grinblatt and Keloharju (2001), who relate language and culture to the home bias. A caveat on semantics is warranted. Although we call the tendency for Swedes to choose Swedish institutions a bias, we are careful not to use the word irrational. A behavioural preference is not irrational. Psychic utility, although not derived from consumption, is still utility. If investors gain psychic utility when they deal with the familiar then a behavioural preference for domestic-managers is rational.

5. Conclusions

We find a strong bias for mutual funds offered by domestic institutions relative to those offered by foreign institutions. We call this preference for domestic institutions, as opposed to assets, the home*boy* bias. Cross-*fund* regressions using Sweden's privatized retirement system data show that funds offered by domestic institutions received 30 times more money than similar (same fee, age, return history, size, and style) funds from international institutions. This bias would be even stronger if we considered the fact that the default fund was managed by a Swedish institution. The homeboy bias is empirically distinct from the home bias. Even when shopping for exposure to, say, Asian or North American assets, the Swedes still prefer a Swedish institution over an international institution.

The homeboy bias is not driven by economic explanations such as information asymmetries or foreign exchange risk. The bias is actually strongest when domestic managers are at an information disadvantage and the bias is evident even when foreign exchange risk is held constant. In contrast, behavioral or preference explanations receive some support; more familiar funds, as measured by advertisement and market share in the banking or insurance industry, received more money and more investors.

In support of behavioral explanations such as familiarity and nationalism, our cross-*individual* regressions find that financially unsophisticated and provincial (not familiar with or comfortable with international people and institutions) individuals are the most homeboy biased. We note that the bias is not necessarily irrational--investors may gain utility from dealing with domestic institutions as theorized by Tajfel and Turner's (1979) social identity theory or by lowering decisions cost by limiting their consideration set as theorized by Alba and Chattopadhyay's (1985).

Although we introduce the homeboy bias, many questions remain. Is the homeboy bias universal or limited to Sweden? Is the homeboy bias limited to mutual funds, or does it extend to other financial products or services? Is the preference for the familiar driven by consideration sets (limiting search costs) or some form of social identity utility based on shared culture, language, nationalism, or xenophobia? Much work remains to be done.

Chapter 4. Portfolio choice and menu exposure[1]

Anders Karlsson, Massimo Massa, and Andrei Simonov

1. Introduction

Are investors more likely to invest in a growth fund if fund companies offer more growth funds? The goal of this chapter is to address this question by investigating how investment is affected by the available alternatives. We study whether investors choose to invest in assets that are proportionally more represented in the menu of choices available to them and how an exogenous change in the menu representation affects investor decision.

We consider the unique experiment in portfolio choice provided by the new Swedish retirement scheme based on private accounts. Since 2000, Swedes have had to allocate a fraction of their yearly income to a fully funded retirement account. They can choose up to five investment vehicles ("the funds") out of more than 464 possibilities ("the menu"). No other choice is available and all the capital has to be invested. Each household receives a booklet containing a standardized description of all the available funds. Each fund belongs to a different category, such as growth funds. All the funds charge uniform and low management fees. No other fees are charged. The choice can be altered every day with no search or switching costs. The cost structure characteristics of the financial firms determine the funds that are offered in the menu.

We show that investors choose funds according to the funds' representation in the menu. Categories that offer more funds are chosen proportionally more. Moreover, if, for some exogenous reason, the

[1] Winner of the Best Paper Award in Behavioural Finance, European Finance Association 2006

number of funds of a specific category rises, the investors rebalance their portfolios increasing their investment in that category. This is the case for both the new funds being offered and the existing funds belonging to the category (i.e. growth) in which the number of funds has increased. For example, assume that investors already hold shares in the UBS growth fund and the Allianz income fund, and a new fund, HSBC growth fund, is offered. The mere addition of the HSBC growth fund, increasing the percentage of growth funds available in the menu, is enough to stimulate investment in the already existing UBS growth fund and to reduce the investment in the Allianz income fund.

We define this phenomenon – the tendency to invest in the stocks that are more heavily represented in the menu – as "menu exposure". It posits that a bigger representation in the menu conveys the idea that the particular category is better, regardless of its intrinsic merit. We show that an increase in the over-representation (with respect to the market representation) in the menu of a category of one standard deviation leads investors to over-invest in such a category by 4.4%.

We propose two alternative explanations of this phenomenon. The first is based on a form of "representativeness bias"; the second relies on an "information-induced motivation". We use the degree of informativeness of the investors to distinguish between these two hypotheses. Using the fund-picking ability of the investors and their degree of portfolio concentration as proxy for informativeness, we find a positive correlation between the sensitivity of the investors to the menu and their degree of informativeness. Investors who base their choice on the way funds are represented in the menu consistently choose lower performing funds and display a lower degree of portfolio concentration.

In particular, investors who display an exposure to the over-represented categories one standard deviation greater than the average, invest in funds that deliver a yearly performance 0.54% lower than average. In the context of the pension system, this can result in a 16% lower than expected pension income. Also, investors who display an exposure to the over-represented categories one standard deviation greater than the average, display a degree of portfolio concentration 18% lower than average (unconditional). We argue that this suggests that less-informed investors are affected by the menu representation, as this provides a source of information from which they infer the quality of the available choices.

Moreover, exposure to the menu bias is negatively related to investor trading. The more the investors have invested in categories that are over-represented in the menu, the less they trade. An increase in the exposure to the over-represented categories by one standard deviation lowers turnover by 7.7%. This also would be consistent with an information story, as less-informed investors – the ones more exposed to the menu representation – are less willing to act on the basis of their information (e.g. Hellwig, 1980, 1982).

Our findings contribute to two main literatures. The first is the literature on "style investing" and the second is the literature on the "home bias". In the literature on style investing, Mullanaithan (2002) and Mullanaithan and Shleifer (2006) argue that investors have a tendency to think through categories. Barberis and Shleifer (2003) and Barberis *et al*. (2005) show that investors classify risky assets into different styles and that news about one style can affect the prices of other apparently unrelated styles. Stocks returns are therefore affected because they belong to a particular style (Teo and Woo, 2003).

However, one question that has been left unanswered is where these styles come from. One possibility is that investors who classify stocks according to characteristics such as book-to-market originate the style characteristics. Mutual fund families cater to investor demand by offering funds specialized in these categories. This implies that the sprawling of different funds in various categories may be explicable in terms of the desire of the fund families to cater to investor needs. An alternative possibility is that styles are generated by the fund families themselves in order to segment the market. In this case, it is the investors who "adjust" to the funds being offered. That is, the mere fact that funds are established changes the representation of the different styles in the menu from which investors choose and this affects investor demand. Demand is not catered to, but created: *fund families create investor demand* by offering new products and changing the representation of the categories in the supply.

We show that an exogenous change in the way a category is represented in the menu does indeed affect portfolio choice. An increase in the number of funds of a category raises the investment in the funds already existing in the category. In particular, an increase of the weight of a category in the menu by 1% induces the investors to increase their investment in this category by 0.38%. This supports the interpretation that style-based demand is created and not just catered to.

Let us now consider the implications for the home bias, the stylized fact that investors tend to allocate most of their portfolios to domestic assets (Baxter and Jermann, 1997). Standard explanations include some form of limited information, transaction and search cost or limited investor awareness (Cooper and Kaplanis, 1994; Kang and Stultz, 1997; Lewis, 1999; Strong and Xu, 2003). For example, the "investor recognition hypothesis" posits that investors have a limited knowledge of

the menu of available assets and explains the home bias in terms of the fact that information is available only for domestic assets (Merton, 1987). Alternatively, it has been argued that limited capacity in processing information induces investors to focus on the few assets on which they have a lower information cost and that these are more likely to be the domestic ones (Sims, 2000; Peng and Ziong, 2002; Van Nieuwerburgh and Veldkamp, 2004). Our contribution is to show that, *even in the presence of complete awareness about the menu and its components and no transaction or switching costs,* the home bias still persists. We provide an explanation based on the menu and argue that the *bias is embedded in the menu*. As the menu representation of domestic and international categories changes, so does investor demand. The mere fact that local investment companies offer more investment packages containing domestic stocks induces the investors to over-represent domestic stocks in their portfolios. This would also be consistent with the fact that the reduction in the home bias experienced in most countries during the last two decades has occurred in concomitance with the burgeoning of the offer of international funds.

Our findings also relate to the literature on familiarity (Huberman, 2001; Huberman and Sengmuller, 2002) and proximity bias (Coval and Moskowitz, 1999, 2001; Grinblatt and Keloharju, 2001). However, in our case, proximity is defined not so much in terms of geographical distance, as in terms of menu representation. Finally, our results also contribute to the burgeoning literature on how investors choose investment in their retirement saving accounts (Agnew, 2002; Agnew *et al.*, 2003; Choi *et al.*, 2001 and 2004; Huberman and Sengmueller, 2004).

Some features make our contribution unique. First, in our set-up, the offering of the funds and their representation in the menu is largely *exogenous*, being a function of the cost structure of the funds and of their

managing companies. That is, we have an experiment in which we can estimate investor demand in the presence of an exogenous identification on the supply side.

Second, we know the menu each investor can choose from and the set of standardized information on the funds in the menu that is available to each investor. In the previous studies on portfolio choice based on field data, the menu available to the investors is not known in its entirety and data on the information available to the investors on the different choices are severely restricted. Allocation constraints, as well as some unwritten rules not easily detectable by the econometrician, may limit investment ability. Moreover, even if the investor's full awareness of all the options is available, their completeness is in general doubtful. In our case, however, each investor is provided with a menu of choices in an easily accessible and standardized form. This, by itself, helps to test for alternative explanations based on limited awareness of the investors.

Our third unique feature is related to the fee structure. In general, information on the effective overall costs faced by the investors is not available. For example, most of the studies on mutual fund demand use the information on the fees that is contained in the prospectus as a proxy for the price charged. However, this does not take into account the documented practice of fee-waiving (Christofersen, 2002). Moreover, load fees are not homogenous across investors, as their size depends on the investment horizon of the investors. In our case, not only do we know the overall *effective* fees charged by each fund, but also the range of variation of the management fees is very limited, while load and switch fees are prohibited.

The remainder of the paper is structured as follows. In the next section, we describe the experiment, provide institutional details and lay

out the hypotheses. In Section 3 we describe the data. In Sections 4 and 5, we provide evidence on the impact of menu exposure on investor choice and study its dynamic implications. In Section 6, we investigate the nature of menu exposure, testing whether there is a relation between sensitivity to menu exposure and investor trading, investment skills and degree of portfolio concentration respectively. A brief conclusion follows.

2. The experiment

2.1 Institutional details

The experiment is provided by the adoption of a new pension system in Sweden. The looming demographic crisis induced the Swedish authorities to restructure the national pay-as-you-go pension system, moving towards a capitalization-based system. Starting in 2000, part of the pension system was replaced by a self-financing system. This part has been designed, in particular, for individuals born after 1954 and not yet retired.

The resulting pension system consists of two parts.[2] The first part is a defined benefit component called the "income pension". This entitles the receiver to receive a fixed stream of income that is a function of its yearly payments (approximately 16% of the annual pension-based income), years of contribution, ... The second part, the premium pension (PPM), is a defined contribution component. Each year, individuals invest 2.5% of their annual pension-based income. The accrued amount from the premium pension part will be paid out on a monthly basis to the individual at the time of retirement. The money is allocated at each individual's discretion, from a choice of 464 funds. The individual can choose between one and five funds. If no choice is made, the allotted

[2] There is also a minimum guaranteed pension level designed to ensure that no retiree will be completely without pension payments at the time of retirement, regardless of previous income.

money is invested in the default alternative, the Seventh Swedish Pension Fund, which is an equity fund run by the government.

The funds are presented in a booklet ("the menu") mailed to each individual investor. Fees are capped and homogenous. Load and switch fees are prohibited and management fees are generally much lower than for the same funds offered outside the PPM. Figure 1 reports the information provided for each fund.

Figure 1: Extract from the PPM Investor Information Folder, fund example

Fund number	Fund name, Management company	Information regarding the funds	Fund fee (%)	Percentage return 99-12-31 (after fees)						Total risk
				In the year					Last 5	(last 3
				95	96	97	98	99	years	years)
191080	Baring Global Emerging Markets Baring International Fund Managers (Ireland) Ltd	Emerging markets' equity and equity related assets	1.59	-32	10	25	-25	77	25.3	32 (Red)

The percentage return for the last five years equals the compounded annual growth rate of return for the years 1995 through 1999. The total risk corresponds to an annualised percentage standard deviation of three-year monthly historical fund returns. The total risk is also categorised into five different classes, and colours, with respect to standard deviation; Class 1: very low risk, dark green, percentage standard deviation in the range 0-2; Class 2: low risk, light green, 3-7; Class 3: average risk, yellow, 8-17; Class 4: high risk, orange, 18-24; Class 5: very high risk, red, 25-.

Source: Premiepensionsmyndigheten (2000).

Apart from the fund's identification number, name and fund family, information is provided on fees, past returns and risk. Risk is represented by two measures. The first is a simplified graph displaying a jagged red line for very high risk or a flat green line for very low risk. There are five risk categories of this kind. The second risk measure is a

number, which is the fund's annualized standard deviation using the past 36 monthly returns.

The amount of income to be invested is defined in terms of the annual income. All the annual income from the age of 16 is included. However, individuals earning more than 7.5 times the "income base amount per year"[3] will only be credited an upper limit of 18.5% of the 7.5 income base amount, although they will still pay 18.5% of their pension-based income to finance the pension system. For example, for individuals with a pension-based income in the year 2006 of 360,000 SEK[4], only 18.5% of 333,750 SEK (7.5 * 44,500) will count towards their pension. In other words, their defined contribution to the PPM in 2006 is 8,344 SEK (0.025*333,750), which is the maximum contribution per individual for that year.

In the first PPM choice, the pension-based income from the previous four years served as a base for the amount invested. During these four years, 2% for 1997 and 1998 and 2.5% for 1999 and 2000 was paid into the system. In total, the maximum amount invested in 2000 was 26,202 SEK. The average amount invested was 12,651 SEK for the entire population and 13,506 SEK for those who made an active choice. The slightly higher average for active investors is consistent with Engström and Westerberg (2003) who show that active investors tend to have slightly higher income than the "default" investors.

The introduction of the PPM system was preceded by a massive and unprecedented advertising campaign by both the government and the mutual fund industry (Cronqvist and Thaler, 2004; Cronqvist, 2006). While low on information content (level of fees, risks, etc.), the ads

[3] For the year 2006, the base amount equals 44,500 SEK.
[4] The typical exchange rate in 2000 was 10 SEK for 1 USD (8.5 SEK/Euro) and for 2006 around 7 SEK for 1 USD (9.3 SEK/Euro).

helped to create a positive image of investing in financial markets. More than 86% of the investors were exposed to TV ads; 75% were exposed to some advertising in print media; 59% saw some kind of outdoor ad; and 36% listened to radio ads (Cronqvist, 2006). Thus, virtually the whole population of the country was exposed to information about investing in financial markets.

The representation of the funds in the menu deviates quite significantly from the weights in the universe of mutual funds offered for sale in Sweden (as reported by Morningstar Sweden) as well as from the weights they would have in the world market portfolio. As confirmed from conversations had with the industry, the reason of the deviation from the market lies in the different costs experienced by different categories of funds and the fact that fees are effectively capped. This induces a higher offering of funds proportionally less expensive to run.

The biggest deviations from the market representation are reported for the funds related to Sweden and to the Nordic region in general. These are grossly overweighed in the menu with respect to what their weight in market portfolio would be. For example, the weight of North America and US in the World market portfolio[5] exceeds 50%, while in the universe of the funds offered for sale in Sweden it is only 12.5%, and in the PPM menu this weight drops to just 8.4%. On the other hand, the weight of the Swedish Small Cap Funds in the World market portfolio is 0.17%, in the universe of mutual funds offerings in Sweden it is 1.9%, and in the PPM menu it rises to 2.3%.

[5] The bulk of the data on market sources comes from MSCI Global Capital Markets Indices. We complemented it with the data from Dimson, Marsh, and Staunton (2002).

2.2 The spirit of the test

We start with a simple example. Suppose investors can choose from among four bond funds and six equity funds. How will they compose their portfolio? Our working hypothesis is that the investors are affected by the representation of the fund categories in the menu. If there are more equity funds than bond funds, they will select more equity funds and fewer bond funds. If the number of equity funds is very high when compared with the other fund categories, investors will select only equity funds. This suggests a first testable restriction.

H1. There is a positive correlation between investment and representation in the menu.

That is, the higher the representation of a category of funds in the menu, the more the investors will hold the funds that belong to such a category. This is similar to the "*1/n heuristic*". This heuristic posits that investors offered choices from *n* assets, invest 1/n in each asset (Benartzi and Thaler, 2001; Huberman and Jiang, 2004). In our setting, the *1/n heuristic* would induce investors to put 1/10 in each fund and therefore have 40% in bond funds and 60% in equity funds, if the choice is based on funds. If, instead, investors just choose the category, the *1/n heuristic* would induce investors to put 50% in equity and 50% in bonds.

The intuition behind menu exposure is different. Investors do not invest in all the choices equally, but use the menu representation of the categories to select the funds that belong to the *most represented categories*. Investors see a higher representation of the equity funds and therefore invest more in equity funds. They do not think of allocating 1/n of their investment to each alternative. They prefer to *allocate most of their money to the "better" alternative* – the category with the highest

representation. That is, they would allocate less than 40% to bond funds and more than 60% to equity funds, if the choice is based on funds. If, instead, investors just choose the category, they would allocate more than 50% to equity funds.

A test that can distinguish our hypothesis from the *1/n heuristic* is based on the dynamic of the choice in the presence of a change in the menu. Let us continue with the previous example and assume that four new equity funds are added to the menu, so the investor can now choose between ten stock equity funds and four bond funds. The test would focus on whether a *change in category representation affects investor behavior in such a category,* that is, whether an increase in the number of funds belonging to a category stimulates investment in the *already existing* funds belonging to that category. This suggests a second testable restriction.

H2. There is a positive correlation between an increase in representation of a category and an increase in investment in the already existing funds of the same category.

Existing theories would have a hard time explaining such an increase. Let us start with the *1/n heuristic*. How would investors react according to the *1/n heuristic*? They would reallocate their portfolio so as to invest 71% (i.e. 10/14) in equity funds and 29% (i.e. 4/14) in bond funds. However, they would also reduce their holdings of each of their existing equity funds in order to be able to invest in as many as possible of the new equity funds. In the case of a constraint limiting the investors to hold no more than five funds, they would either hold the same portfolio as before or, if possible, invest in one of the new equity funds. However, investors would not increase their investment *in the equity funds in which they were already investing*. Indeed, any further reallocation towards

equity funds they were already holding would tilt their portfolio away from the desired 1/n allocation.

In the case of rational portfolio theory, if a new fund is added to the menu, investors may rebalance their portfolio by investing in the new fund. In order to do this they would reduce their holdings of some other existing funds. However, only in very special cases – depending on the correlation of fund returns – would investors increase their investment in funds in which they were already investing. This would not generate a systematic cross-sectional relation.

The "investor recognition hypothesis" (Merton 1987) posits that the choice is affected only by the options of which the investor is actually aware. Given that in our case the entire menu is known, it is unlikely that this theory would help explain the choice of the funds and the rebalancing after a change in the menu. Indeed, one of the nice features of our experiment is the fact that investors receive the same type of *standardized and easily accessible* information that makes them aware of all the funds in the menu. It may be argued that the ordering in the booklet may play a role. For example, a cursory reading will overlook funds reported last in the booklet containing the menu. We will explicitly control for this in the tests.

These considerations suggest that neither the *1/n heuristic*, nor rational theories, nor the investor recognition theory, nor theories based on limited awareness of the available choices would be consistent with an increase in the investment in *existing funds* as their category representation increases as a result of new funds belonging to the same category being added to the menu.

How can we rationalize this in finance? One consistent theory is that investors think through categories (Mullainathan, 2002; Barberis and

Shleifer, 2003; Mullanaithan and Shleifer, 2006) and the representation of a category plays a major role. A bigger representation in the menu conveys the idea that the particular category is better, regardless of its intrinsic merit. Representation in the menu plays the same role that "space in the shelf" has in marketing of consumer goods: bigger space suggests a better product. A bigger representation in the menu therefore conveys the idea that the category is better, regardless of its intrinsic merit. This is similar to "partition dependence". "If people are biased to allocate investment funds evenly over the options that have been identified, then the particular way in which the investment space is partitioned should influence the resulting distribution of funds" (Fox *et al*. 2004).

Why do investors fall in the "trap" of basing their choice on the representation of a category in the menu? We entertain two alternative explanations. The first is based on a behavioral "representativeness bias". Investors tend to give more importance to and believe more in things that they see repeated more often. In this case, the number of funds in a category stresses the importance of the category, thus inducing investors who were already investing there to increase their holdings. The second explanation relies on a rational "information-induced" motivation. Investors compensate for their lack of private information[6] by using the size of the category the fund belongs to as a signal: the bigger the category, the higher the perceived quality of the funds. This would be related to the recent theories that explain investor choice in terms of cost of processing publicly available information (Sims, 2000; Peng and Ziong, 2002; Van Nieuwerburgh and Veldkamp, 2004). This provides a

[6] Indeed, while investors are aware of all the choices and are presented with a standardized set of information, some investors may have private information on the different funds.

further testable restriction, based on the correlation between menu exposure and investor informativeness.

H3. An information-induced story posits a negative correlation between menu exposure and investor informativeness, while a behavioral bias posits no correlation.

Before moving on to the tests, we describe the data and the main variables used in our analysis.

3. The data

The data on individual retirement accounts come from PPM. For each individual, we have semi-annual information on the value of their positions as well as the daily value of their transactions. This information contains all individual choices made from the introduction of the system until October 2004. Data come in the form of the percentage choice (made as a percentage of the portfolio) as well as the amount invested. Both the transaction dates and the clearing dates are known. We also know the universe of funds that was available to investors at any point of time. Moreover, for each individual, we have a set of other demographic variables, such as age, gender, place of residence (defined as church parish) and total amount invested in PPM system (highly correlated with income).

Some preliminary descriptive statistics are provided in Table 1. We can see that, conditional on making an active choice, individuals hold on average 3.47 funds. Males hold fewer funds than females (although the difference is statistically significant, the economic magnitude is tiny) and younger investors (less than 40 years old) hold more risky funds than the older investors (respectively 3.71 and 3.27).

Table 1: Descriptive statistics

This table reports the descriptive statistics of the variables used (Panel A) and the description of the funds' categories (Panel B). Static measures are defined on a subset of the individuals making the first PPM choice. To construct the measures of excessive weight of the category in the individual portfolio and in the menu, we proceed as follows. Let $w_{i,j}$ be the fraction of the PPM portfolio of the *ith* investor allocated to the *jth* category and $w_{j,mkt}$ be the ratio between the value of the assets managed by the funds offered in the market that belong to the *jth* category and the total value of the assets managed by all the funds in the market. We define the market as Morningstar Sweden. We define the investor choice as: $I_{i,j} = w_{i,j} - w_{mkt,j}$. Alternatively, we construct a measure of choice in which the market is the representation of the category in the world market portfolio as defined above ($I_{i,j}^{*}$). For both measures we also construct a measure based on the log-difference: $\hat{I}_{i,j} = \ln\left(w_{i,j} / w_{mkt,j}\right)$. The menu tilt is defined as $M_j = m_{j,menu} - m_{j,mkt}$, where $m_{j,menu}$ is the representation of the *jth* category in the menu and $m_{j,mkt}$ is the representation of such a category in the world market. We define $m_{j,menu}$ ($m_{j,mkt}$) as the ratio between the *number of funds* offered in the menu (market) that belong to the *jth* category and the overall number of funds in the menu (market), regardless of category of affiliation. We also consider a log menu tilt and denote it as: $\hat{M}_j = \ln\left(m_{j,menu} / m_{j,mkt}\right)$. We use as market the list of funds and categories in Morningstar Sweden. We also consider an alternative proxy for the menu bias based on the actual size of the category (M_j^{*}). This defines $m_{j,menu}$ as the ratio between the assets managed by the funds of the *jth* category in the menu and the assets managed by all the funds in the menu, and $m_{j,mkt}$ as the ratio between the assets managed by the funds of the *jth* category in the market portfolio and the assets managed by all the funds in the market. We consider the world-market portfolio. To define the world market portfolio, we use the data from MSCI Global Capital Markets Indices and we complement it with the data from Dimson, Marsh, and Staunton (2002) and Dahlquist, Pinkowitz, Stultz, and Williamson (2003). We ignore generational funds and bond funds.

Dummy Gender is defined as 1 if the PPM investor is male and 0 otherwise. *Age* is investors' age in years. *Risk category* is a discrete variable between 1 and 5 that is assigned to each fund by PPM administration and given in the choice booklet. *New fund dummy* is equal to one if the fund did not exist before PPM system starts (most of the funds were added exclusively for PPM system). *Fee* is the funds' annual fee (in percents). *Flows* are defined as the sum of all flows to funds that belong to subcategory over the previous 6 months, as a fraction of total money under management at t-1. *Log_amt* is the log (base 10) of the amount invested in PPM (in SEK) plus 1 SEK. *Alphabet* is the order of fund within the menu category (normalized to be between 0 and 1). *Number of choices* is the number of funds investors choose within the PPM system (no more than five). The *Number of times choices were changed* is calculated over 6 month intervals. *Dynamic variables* are defined as semi-annual changes. The *Changes in weights* are defined on category level as $\Delta w_{i,j,t} = w_{i,j}$ (t)- $w_{i,j}$ (t-1). The *Change in category menu weight* is $\Delta m_{j,t} = m_{j,menu}$ (t)- $m_{j,menu}$(t-1)

Table 1, Panel A: Main variables

Variable	*Mean*	*StdDev*	*Minimum*	*Maximum*
Excessive Weight in individual portfolio, categories I_{ij}	0.206	0.182	-0.240	1.000
Excessive Weight in individual portfolio, categories $\hat{I}_{ij}$	2.192	2.060	-4.001	8.071
Excessive Weight in individual portfolio, categories I^{*}_{ij}	0.027	0.348	-0.990	0.999
Excessive Weight in individual portfolio, categories $\hat{I}^{*}_{ij}$	0.717	1.036	-1.833	3.258
Gender (male=1)	0.516	0.500	0.000	1.000
Age	40.220	10.630	18.000	62.000
Risk category	4.222	0.502	3.000	5.000
New fund dummy	0.515	0.500	0.000	1.000
Fee (in pct.)	0.906	0.400	0.200	3.970
Flows	0.087	0.106	-0.647	0.444

Variable	*Mean*	*StdDev*	*Minimum*	*Maximum*
Log amount	4.081	0.291	2.224	4.720
Alphabet	0.513	0.269	0.028	1.000
Excessive Weight in menu M_j	0.012	0.025	-0.071	0.061
Excessive Weight in menu $\hat{M}_j$	0.942	0.431	0.353	4.016
Excessive Weight in menu M_j^*	0.002	0.070	-0.341	0.059
Excessive Weight in menu $\hat{M}_{1j}^*$	0.992	1.666	-5.613	6.908
Number of choices	3.473	1.409	1.000	5.000
Number of times choices were changed	1.645	1.866	1.000	155.000
Changes in category weights, Δw_{ijt}	0.030	0.277	-1.000	1.000
Changes in category menu weight, Δm_{jt}	0.001	0.008	-0.016	0.028

Table 1, Panel B: Description of mutual funds

Category 1	Category 2	Category 3	Number of Funds	Mean Percentage Fee
Equity	Sweden	Sweden (normal)	28	0.92
		Sweden Small Cap	6	1.16
		Sweden Index	7	0.41
	Regional	Swedish Equity and Foreign Equity	11	0.74
		Nordic	12	1.27
		Europe	36	1.11
		Euroland	8	0.99
		Europe Small Cap	9	1.23
		Europe Index	7	0.48
		North America and USA	26	1.05
		Asia and Far East	18	1.19
		Global	32	1.01
		New Markets	21	1.56
	Countries	Japan	20	1.06
		UK	6	1.21
		Other Countries	19	1.25
	Industry	IT & Telecommunications	19	1.15
		Pharmaceutical	7	1.36
		Other Industries	16	1.17
Balanced Funds	Balanced	Swedish Equity and Fixed Income	3	1.08
		Swedish Equity, Swedish and Foreign Fixed Income	28	0.74
		Swedish and Foreign Equity	22	0.93
Generation Funds	Generation	Pension in less than 10 years	5	0.46
		Pension in less than 20 years	6	0.46
		Pension in more than 20 years	21	0.46
Fixed Income	Fixed Income	Sweden, short maturity	15	0.46
		Sweden, long maturity	15	0.45
		Europe and Euroland	18	0.70
		Others	15	0.79

It is also interesting to note that the portfolios of older investors are less exposed to equity and less risky. Indeed, older (than the median) investors hold 80% of their portfolios in equity. This figure rises to 86% for younger investors. If we consider the transactions, we see that, on average, investors rebalance their portfolios 1.64 times over the time period of our sample (with the median of 1, a 99th percentile of 8 and a maximum of 155). This frequency increases in the case of males (the females rebalance 1.59 times, the males 1.70 times) and does not depend on the age of the investors in any significant way. In Panel B, we report information on the representation of the different categories and sub-categories in the initial menu.

4. Are investors affected by the representation in the menu?

We test whether there is a positive correlation between investment and representation in the menu (H1). The test will focus on the relation between the investment and the way the choices are laid-out in the menu.

4.1 A definition of menu tilt and the test layout

We start by defining the "menu tilt". The tilt captures whether an investor is offered a menu that has a different partition of choices from the one the investor would find in the market. In particular, we define menu tilt as the difference between the representation of the mutual fund categories in the menu and their representation in the market portfolio. It quantifies the degree by which the menu representation differs from the world market representation. To avoid burdensome notation, we will omit the time subscript, except for the cases in which it is strictly necessary. For the *jth* category (e.g. growth), the menu tilt is:

$$M_j = m_{j,menu} - m_{j,mkt},$$

where $m_{j,menu}$ is the representation of the *jth* category in the menu and $m_{j,mkt}$ is the representation of such a category in the world market. We define $m_{j,menu}$ ($m_{j,mkt}$) as the ratio between the *number of funds* offered in the menu (market) that belong to the *jth* category and the overall number of funds in the menu (market), regardless of category of affiliation. For example, if three growth funds are offered and a total of 100 funds are in the menu, $m_{j,menu}$ is equal to 3% for the growth category. For two reasons, we focus on the *number* of funds as opposed to actual size of the category defined in terms of the asset under management. First, we argue that what drives investors is the actual gamut of the menu as opposed to the market size of the funds. Second, by using the number of funds as opposed to the assets the funds manage, we purge our analysis of potentially endogenous price effects. The difference between the two ratios (M_j) represents how much the *jth* category is over(under)-represented in the menu with respect to the standard market offering. We also consider a normalized menu tilt and denote it as: $\hat{M}_j = ln\left(m_{j,menu} / m_{j,mkt}\right)$. We use as market the list of funds and categories in Morningstar Sweden.

M_j is reasonably exogenous. Indeed, the offer of the funds is directly affected by the severe limits on charging fees imposed on the fund. For example, there are proportionally fewer South East Asian funds in the menu and proportionally more Europe-focused funds, due to cost constraints. Managing a fund invested in South East Asian stocks is more expensive than a fund invested in European stocks; the fact that fees are capped prevents the fund providers from passing on to the investors these additional costs. This makes it less lucrative to offer more "expensive" South East Asian funds and more advantageous to propose "cheaper" Europe-focused funds. The net effect is the creation of an artificial

distortion, *not related to the investor demand*, but exogenously induced, that sets $M_j \neq 0$. [7]

There may, however, be some residual spurious correlation due to the fact that the fund management companies offer new funds to cater specifically to recent investor demand. To address this issue, we test whether M_j is related to recent fund flows. We collected quarterly data on the flows from Swedish investors into mutual funds (including the ones located abroad) from Money Mate. We then regressed M_j on these flows. The (unreported) results show that there is scarcely any correlation between the two variables (the *R-squared* is essentially 0). This would suggest that the offering of the new funds is not related to investor demand but is mostly related to some supply- and cost-related factors. In any case, we will also include these flows among the control variables in the main regressions.

We also consider an alternative proxy for the menu bias based on the actual – i.e., value-based – size of the category (M_j^*). This defines $m_{j,menu}$ as the ratio between the assets managed by the funds of the *jth* category in the menu and the assets managed by all the funds in the menu, and $m_{j,mkt}$ as the ratio between the assets managed by the funds of the *jth* category in the market portfolio and the assets managed by all the funds in the market. We consider the world-market portfolio. To define the world market portfolio, we use the data from MSCI Global Capital Markets Indices. We complement it with the data from Dimson, *et al.* (2002). [8]

To define the assets managed by the funds in the menu, we cannot use the actual representation in the menu, as this is not available to

[7] Another reason for exogenous shock to the menu was the reluctance of international fund families to participate in a yet unproven system. Many chose to sit on the sidelines during the initial introduction of the system.
[8] We use both definitions based on total market capitalization and free float. The results do not differ, and we report only ones based on market capitalization.

investors when they make their first choice and it is jointly determined with the choice itself. We therefore proceed as follows. First, we determine the average (median) size (i.e., asset value) of the funds in each category using Morningstar. Then, for each category, we multiply the average (median) fund size time the number of funds that are offered in each category in the menu.

We then define the portfolio choice of the investor. Let $w_{i,j}$ be the fraction of the PPM portfolio of the *ith* investor allocated to the *jth* category and $w_{j,mkt}$ be the ratio between the value of the assets managed by the funds offered in the market that belong to the *jth* category and the total value of the assets managed by all the funds in the market. As before, we define the market as Morningstar Sweden and $w_{j,mkt}$ as the ratio between the value of the assets managed by the funds belonging to the *jth* category and the total value of the assets managed by all the funds in Morningstar Sweden. We define the investor choice as: $I_{i,j} = w_{i,j} - w_{mkt,j}$. That is, the fraction of the retirement money that the *ith* investor allocates to the *jth* category in *excess of its weight in the "market*". We also construct a measure based on the normalized difference: $\hat{I}_{i,j} = ln\left(w_{i,j} / w_{mkt,j}\right)$. Alternatively, we construct a measure of choice in which the market is the representation of the category in the world market portfolio as defined above ($I_{i,j}^{*}$).

We will test whether $I_{i,j}$ ($I_{i,j}^{*}$) correlates with M_j (M_j^{*}). That is, we study whether the decision of the *ith* investor to invest in the *jth* category is related to the representation of the category in the menu. Our working hypothesis posits a positive correlation between $I_{i,j}$ ($I_{i,j}^{*}$) and M_j (M_j^{*}). Investors, influenced by the representation in the menu, replicate in their portfolios the tilt towards some funds contained in the menu. If instead, the investors are not affected by the menu bias and choose their portfolio

the same way as they allocate their investment in ordinary mutual funds, we expect no correlation.

4.2 The test and the findings

To implement the test, we regress the investor choice on the menu tilt and a set of category- and investor-specific control variables:

$$I_{ij} = \alpha + \beta M_j + \gamma controls_{ij} + \varepsilon_{ij}, \qquad 1)$$

where, for the *ith* investor, I_{ij} is the fraction of the retirement money invested in the *jth* category in excess of its weight in the world market portfolio. The control variables are both category- and investor-specific. The investor characteristics are the investor's age and gender; the overall amount the investor has invested in the PPM retirement account (an increasing function of their income); the number of funds chosen in the PPM system (between one and five), and their residential municipality. This latter variable accounts for potential social effects in investment (Hong, Kubik and Stein, 2004). The category characteristics are the average performance and the average risk category of the funds chosen within the category; an index that represents the fraction of the assets invested within the category with domestic (as opposed to foreign) fund management companies; an index that represents the fraction of the assets invested in the category with funds that did not exist before and has been created only as an investment vehicle for PPM; the average (weighed according to the fraction of the assets invested) level of fees charged; the flows into the fund category in the prior six months, and the printed menu representation of the category. This variable is constructed as follows. We first define a variable (Alphabet) that proxies for the position of the fund in the printed booklet and is normalized to be between zero and one. We then construct a variable (Alphabet*(1-Alphabet)). This variable reaches

the maximum (0.25) for funds in the middle of the alphabetic menu and has value 0 for the start and end of the menu. We concentrate on the sample of people who made their choice in the fall of 2000[9].

The results are reported in Table 2. We consider alternative specifications. In Panel A, the dependent variable is the fraction held in excess of the Morningstar portfolio. The menu bias is proxied by M_j. In Panel B, the dependent variable is the fraction held in excess of the world-market portfolio. Menu bias is proxied by M_j^*.

Table 2: Static regression

In this table we report the results of regression of the fraction of the retirement money invested in the j[th] category in excess of its weight in the Morningstar menu (Panels A and C) and the World Market Portfolio (Panels B and D) and on set of individual, fund-specific and menu-specific characteristics. Panels A–D report the results for I_{ij}, I^*_{ij}, I_{1ij}, I^*_{1ij}, respectively. The variables are defined in Table 1. All the estimates are done using a White heteroscedasticity-consistent estimator clustered over municipalities. The sample contains 6,915,063 observations. The *t-statistics* are reported in parentheses. The coefficients for Age and Age^2 are multiplied by 10,000. For specification (4) we also report the level of significance obtained via bootstrapped estimates with 10,000 replications. Significance on 10%, 5%, 1% and 0.1% is denoted as *, **, *** and ****, respectively.

Panel A: Regression for I_{ij}

	(1)		(2)		(3)		(4)	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
M_j	1.731	(33.76)	1.927	(38.46)	1.421	(42.99)	1.427	(42.87)****
Alphabet*(1-Alphabet)	-0.089	(-22.70)	-0.009	(-2.43)	-0.047	(-11.13)	-0.044	(-10.56)****
Flows	-0.326	(-)	-0.326	(-82.61)	-0.281	(-98.35)	-0.282	(-99.23)****
Risk category			0.012	(-19.03)	-0.010	(-18.39)	-0.010	(-19.51)****
New fund dummy			0.025	(35.96)	0.001	(1.32)	0.001	(1.65)
Fee			-0.026	(-9.59)	-0.015	(-6.09)	-0.015	(-6.22)****
Domestic Fund Management			0.003	(-3.68)	-0.026	(-21.76)	-0.025	(-21.67)****
Log_amt					-0.025	(-28.22)	-0.030	(-33.93)****
Number of choices					-0.101	(-215.17)	-0.100	(-218.25)****
Gender							0.012	(70.37)****
Age							1.670	(1.71)
Age^2							0.006	(0.56)
Fund subcategory fixed effects	Y		Y		Y		Y	
Municipality fixed effects	N		N		Y		Y	
Adj. R^2	0.130		0.148		0.478		0.479	

[9] We also use the full sample. The results are similar. However, the full sample that contains both original choices and choices of newcomers to the already existing pension system is open to criticism as it compares choices made under quite different circumstances.

The results show a positive correlation between the over-representation in the menu of a category and the investor over-investment in it. In particular, a one standard deviation higher representation of a category in the menu with respect to Morningstar Sweden increases the investment in such a category from 20.6% to 25%, or by 24% of dependent variable's standard deviation. A similar overrepresentation of category in the menu with respect to world market increases overinvestment from 2.7% to 21.4%, or by 53% of dependent variable's standard deviation. Over-investment is lower for high-income savers. A one standard deviation higher pension contribution is related a 0.86% (0.80%) lower deviation from the market (world market) portfolio, corresponding to 4.8% (0.25%) of the dependent variable's standard deviation.

Panel B: Regression for I^*_{ij}

	(1)		(2)		(3)		(4)	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
M^*_j	3.012	(117.55)	2.655	(81.18)	2.666	(82.04)	2.665	(81.95) ****
Alphabet*(1-Alphabet)	-0.087	(-24.55)	-0.066	(-17.49)	-0.119	(-35.54)	-0.118	(-34.98) ****
Flows	-0.656	(-56.01)	-0.100	(-32.70)	-0.427	(-34.24)	-0.427	(-34.21) ****
Risk category			0.100	(92.09)	0.095	(65.70)	0.095	(65.69) ****
New fund dummy			-0.013	(-15.88)	-0.034	(-40.99)	-0.034	(-41.17) ****
Fee			-0.044	(-37.26)	-0.032	(-32.97)	-0.032	(-33.29) ****
Domestic Fund Management			0.019	(8.13)	-0.006	(-3.42)	-0.005	(-3.17) **
Log_amt					-0.028	(-30.34)	-0.028	(-29.46) ****
Number of choices					-0.098	(-203.42)	-0.097	(-208.57) ****
Gender							0.008	(46.90) ****
Age							-10.572	(-12.17) ****
Age^2							0.129	(12.14) ****
Fund subcategory fixed effects	Y		Y		Y		Y	
Municipality fixed effects	N		N		Y		Y	
Adj. R^2	0.584		0.587		0.670		0.670	

Also, over-investment is stronger for males (by about 0.8-1.2%, or by 3.5-4.5% of standard deviation of dependent variable) and older people. Moreover, over-investment is more pronounced for more risky funds and is reduced for new funds and by the level of fees. A one standard deviation higher level of fees is related to a 0.16% (0.33%)

lower the deviation from the market (world market) portfolio, equivalent to about 1% of the dependent variable's standard deviation.

It is interesting to note that the order by which the categories are reported in the booklet section does play a role. Categories reported in the middle of the booklet are chosen 2.9% less frequently than the ones at the beginning or end of the menu. Alternative specifications (not reported) show that the effect is even stronger for the categories appearing at the beginning of the menu.

The results are robust to the inclusion of category-fixed effects and to the use of different sub-categories of funds. Thus, we remove regional funds, leaving Sweden, Industry and Country Funds. We repeat the procedure removing Industry and leaving in Regional, Country and Swedish Funds, and so on. Our results are unchanged, suggesting that they are not driven by any particular category.

Can our results be driven by local interaction between investors? Indeed, Duflo and Saez (2002) explore peer effects in retirement savings decisions, while Hong *et al.* (2004) do the same for investment in mutual funds. They find significant own-group peer effects on participation and fund choice, but no cross-group peer effects. To address this issue, we add 290 municipality fixed effects. The results are reported in columns (3) and (4). They coincide with the ones reported earlier.

As an additional robustness check, we also re-estimate equation (1) removing Swedish counties one-by-one (26 in total). For example, we remove Stockholm County and re-estimate (1) for the 25 counties left. We repeat this procedure 25 times, removing one county from the list each time. The (unreported) results are not affected by any particular locality. Finally, we also apply some non-parametric measures of the

standard errors, bootstrapping the main specifications with both fund subcategory- and municipality-fixed effects. Our results hold.

In Panels C and D, we consider the cases in which we use normalized differences of tilt ($\hat{I}_{i,j}$ and $\hat{I}^{*}_{i,j}$). The results confirm that the menu representation affects portfolio choice. The results are robust both statistically and economically. Increasing the tilt in the menu by one standard deviation raising the dependent variable by 52% of its standard deviation.

Finally, it is worth mentioning that the same results hold if we estimate equation (1) using the fraction invested in the fund, as opposed to the fraction invested in the category. Overall, these findings suggest that the way categories are presented in the menu affects investor choice.

Panel C: Regression for $\hat{I}_{ij}$

	(1)		(2)		(3)		(4)	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
$\hat{M}_j$	2.549	(136.33)	2.594	(159.29)	2.490	(179.69)	2.490	(178.96) ****
Alphabet*(1-Alphabet)	-4.653	(-70.78)	-2.435	(-36.51)	-2.609	(-36.59)	-2.605	(-36.54) ****
Flows	-8.445	(-65.60)	-0.431	(-81.30)	-7.509	(-85.20)	-7.511	(-85.18) ****
Risk category			0.466	(31.09)	0.447	(33.31)	0.446	(33.46) ****
New fund dummy			0.954	(62.07)	0.367	(56.90)	0.367	(56.73) ****
Fee			-0.123	(-35.44)	-0.926	(-34.31)	-0.927	(-34.36) ****
Domestic Fund Management				(-14.45)	-0.200	(-23.13)	-0.199	(-23.19) ****
Log_amt					-0.058	(-27.24)	-0.060	(-22.31) ****
Number of choices					-0.347	(-175.68)	-0.346	(-177.69) ****
Gender							0.016	(14.44) ****
Age							-14.180	(-2.71) **
Age^2							0.212	(3.43) ***
Fund subcategory fixed effects	Y		Y		Y		Y	
Municipality fixed effects	N		N		Y		Y	
Adj. R^2	0.552		0.612		0.643		0.643	

Panel D: Regression for $\hat{I}^*_{ij}$

	(1)		(2)		(3)		(4)	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
$\hat{M}_j$	0.444	(188.20)	0.342	(133.33)	0.372	(169.11)	0.372	(168.95) ****
Alphabet*(1-Alphabet)	-0.319	(-5.52)	-0.330	(-5.64)	-0.177	(-5.76)	-0.173	(-5.66) ****
Flows	1.716	(60.95)	1.588	(50.61)	1.607	(56.29)	1.605	(56.24) ****
Risk category			0.665	(81.45)	0.633	(96.04)	0.632	(95.53) ****
New fund dummy			-0.072	(-8.75)	-0.143	(-16.74)	-0.143	(-16.77) ****
Fee			0.527	(91.56)	0.562	(114.58)	0.561	(114.66) ****
Domestic Fund Management			0.271	(21.86)	0.184	(16.51)	0.185	(16.60) ****
Log_amt					-0.080	(-38.35)	-0.084	(-45.42) ****
Number of choices					-0.295	(-318.46)	-0.295	(-306.46) ****
Gender							0.024	(24.84) ****
Age							-11.841	(-2.15) *
Age^2							0.120	(2.00) *
Fund subcategory fixed effects	Y		Y		Y		Y	
Municipality fixed effects	N		N		Y		Y	
Adj. R^2	0.717		0.753		0.788		0.789	

5. Change in representation and investor choice

We now move on to see what happens in the case of an exogenous change in the representation of a category. We test whether there is a correlation between an increase in representation of a category and an increase in investment in the already existing funds belonging to the same category (H2). We consider a dynamic specification that exploits the exogenous changes in the menu following the addition of new funds.

Let us define $\Delta w_{i,f,j}$ as the change in the fraction of the retirement money that the *ith* investor allocates to the *fth* fund that belongs to the *jth* category and $\Delta m_{j,menu}$ as the change in the representation in the menu of the *jth* category to which the fund belongs. Our hypothesis posits a positive correlation between $\Delta w_{i,f,j}$ and $\Delta m_{j,menu}$ in the case the *fth* fund was already part of the menu before the change in representation. To test it, we estimate:

$$\Delta w_{i,f,j} = \alpha + \beta D * \Delta m_{j,menu} + \gamma(1-D) * \Delta m_{j,menu} + \delta controls_{i,f,j} + \varepsilon_{i,f,j} \qquad 2)$$

where *D* is a dummy taking the value one, if the *fth* fund was already part of the menu before the change in the representation of the *jth* category to which it belongs to took place, and zero otherwise. For example, if a new fund – HSBC Growth – is added to the category of growth funds, *D* would be equal to one for the case in which the *fth* fund is a growth fund that was already part of the menu (e.g. ABN AMRO Growth) and zero otherwise (e.g. HSBC Growth).

The control variables are both fund- and investor-specific. The investor characteristics are the investor's age and gender; the overall amount the investor has invested in the PPM retirement account (an increasing function of their income); the number of choices the investor has made in the PPM system, and their residential municipality. The fund characteristics are the risk category of the fund; a dummy taking the value of one if the fund is managed by a domestic fund management company and zero otherwise; a dummy taking the value of one if the fund did not exist before and has been created only as an investment vehicle for PPM, and zero otherwise; the fees charged by the fund; the flows into the fund category in the prior six months, and the printed menu representation of the fund. This latter is defined as the position of the fund in the printed booklet normalized to one. We also include momentum (returns of the fund in the previous six months).

Univariate statistics show that for old funds (*D=1*) the weight in an increasing category is about larger than for weight in decreasing categories (with differences significant for both mean and median test),

while for new funds there are no statistically significant differences between weights in increasing or decreasing categories.[10]

Our hypothesis posits that $\beta > 0$. The results are reported in Table 3. In Panels A and B, we report the basic specification, while in Panel C, we interact $\Delta m_{j,menu}$ with a *LargeCategory (SmallCategory)* dummy equal to one if the size of the category is above (below) the mean. Panel A is based on a fund-by-fund estimate, while Panel B is based on category-by-category estimates. In Panels A and B, we consider a specification with no interaction between D and $\Delta m_{j,menu}$ (columns 1–3), as well as a specification with interaction (columns 4–6). We use different fixed effects (municipality, year). The standard errors are clustered at the municipality level.

The results show that a change in the way a category is represented in the menu affects portfolio choice. An increase in the number of funds of a category raises the investment in the funds already existing in that category. In particular, an increase in the category weight in the menu by 1% raises the fraction invested in the category by 0.38%. Moreover, the effect is driven mostly by existing funds. An increase in the category weight in the menu by 1% induces the investors to increase their investment in the already existing funds in this category by 0.77%. In the case of the interaction with the *Large Category Dummy* (Table 3, Panel C), we see that the effect is driven by old funds in large categories. The coefficients for small categories and for new funds that belong to large categories are actually negative.

[10] Actually, the weight of new funds in a decreasing category is about 0.15% larger than the one in a decreasing category.

Table 3: Dynamic estimates

This table reports the panel regression of changes in category and fund weights as a reaction to the change in menu by individuals. We aggregate all the changes for a given individual over six month intervals (January 1st-June 30th, July 1st-December 31st) Panel A is based on a fund-by-fund estimate, while Panel B is based on category-by-category estimates. In Panels A and B, we consider a specification with no interaction between *D* and $\Delta m_{j,menu}$ (columns 1–3), as well as a specification with interaction (columns 4–6). We use different fixed effects (municipality, year). In Panel A, the dependent variable is the change in holdings of fund *f*, belonging to category *j*. In columns 1–3, we consider Δm_j, and in columns 4–6 we interact it with the dummy *Newfund.* This dummy takes the value 1 if the fund *j* did not exist in the menu when the previous choice was made. In Panel B, the dependent variable is the change of the investor's holdings in category *j*. In Panel C, we interact Δm_j also with a *LargeCategory (SmallCategory)* dummy that is equal to one if the size of the sub-category is above (below) the mean. *Return6* are the returns of fund *f* in 6 months prior to the date of the choice. *Slweight1* is the fraction of the portfolio invested in all the funds that belong to the same category as fund *f* (i.e., category *j*) with the exception of fund *f*. *Flows* are the flows in the preceding six months as the percentage of funds under management in the category. All other variables are defined in Table 1. In the category-based specifications all fund-related variables are value-weighted average of variables for individual funds. All the estimates are done using the White heteroscedastisity-consistent estimator with clustering over municipalities. The sample consists of 3,237,529 observations. The *t-statistics* are reported in parenthesis. In columns 1, 3, 4 and 6 we used four year fixed effects, and in column 2, 3, 5 and 6 we used also 290 municipalities fixed effects. All the coefficients with the exception of Age^2 (multiplied by 10000) and Δm_{jt}, are multiplied by 100. For the specifications 3 and 6 we also report the level of significance obtained via bootstrapped estimates. Significance on 10%, 5%, 1% and 0.1% is denoted as *, **, *** and ****, respectively.

Table 3, Panel A: Fund-based estimate

	(1)		(2)		(3)		(4)		(5)		(6)	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
Δm_{jt}	0.2231	(4.44)	0.2285	(4.74)	0.2113	(4.28)****						
Δm_{jt} *Newfund							-0.6069	(-7.02)	-0.6409	(-7.35)	-0.6242	(-7.24)****
Δm_{jt} *(1-Newfund)							0.8008	(13.94)	0.8340	(15.48)	0.7929	(14.09)****
Gender	-0.0229	(-0.47)	0.0730	(1.45)	-0.0203	(-0.43)	-0.0215	(-0.44)	0.0726	(1.43)	-0.0191	(-0.41)
Age	-0.1071	(-4.30)	-0.0377	(-1.51)	-0.1068	(-4.28)****	-0.1096	(-4.38)	-0.0423	(-1.69)	-0.1093	(-4.36) **
Age^2	0.0947	(3.24)	0.0067	(0.23)	0.0946	(3.23) **	0.0972	(3.32)	0.0116	(0.39)	0.0972	(3.30) **
Log_amt	0.2767	(2.80)	0.5244	(5.58)	0.2868	(2.86) **	0.2572	(2.60)	0.5295	(5.65)	0.2684	(2.68) ***
Number of choices	-1.2896	(-12.91)	-1.2700	(-12.87)	-1.2873	(-13.08)****	-1.2897	(-12.93)	-1.2701	(-12.89)	-1.2873	(-13.10)****
Slweight1	-0.3948	(-52.00)	-0.3981	(-53.98)	-0.3976	(-55.29)****	-0.3946	(-51.65)	-0.3979	(-53.56)	-0.3974	(-54.86)****
Risk category	1.7426	(3.97)	1.5240	(3.45)	1.7650	(4.02)****	1.7615	(4.00)	1.5478	(3.49)	1.7845	(4.05)****
New fund dummy	29.1112	(168.91)	29.2356	(165.18)	29.0860	(169.01)****	29.2959	(161.82)	29.4254	(158.26)	29.2721	(162.25)****
Fee	-0.8345	(-7.22)	-0.7091	(-6.50)	-0.8352	(-7.19)****	-0.8119	(-6.89)	-0.6956	(-6.29)	-0.8119	(-6.85)
Return6	11.8731	(33.23)	10.9884	(35.06)	11.8702	(34.53)****	11.8258	(33.78)	10.9950	(35.64)	11.8223	(35.15)****
Domestic fund management company	-3.1461	(-28.47)	-3.1176	(-28.25)	-3.1322	(-28.72)****	-3.1376	(-28.78)	-3.1103	(-28.57)	-3.1246	(-29.02)****
Flows	6.4623	(11.76)	6.1166	(12.61)	6.5214	(12.20)****	6.4198	(11.55)	6.1248	(12.56)	6.4775	(11.98)****
Alphabet*(1-Alphabet)	-28.3847	(-4.47)	-28.0261	(-4.40)	-28.5657	(-4.54)****	-28.5308	(-4.48)	-28.2142	(-4.42)	-28.7112	(-4.55)****
M^*_{jt}	44.1110	(3.43)	49.1176	(3.90)	46.2066	(3.54) ***	48.6510	(3.71)	53.4096	(4.17)	50.7176	(3.81)****
Year fixed effects	N		Y		Y		N		Y		Y	
Municipality fixed effects	Y		N		Y		Y		N		Y	
Adj. R^2	0.5849		0.5847		0.5859		0.5852		0.5850		0.5862	

Table 3, Panel B: Category-based estimate

	(1)		(2)		(3)		(4)		(5)		(6)	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
Δm_{jt}	0.38137	(5.81)	0.36390	(5.53)	0.36103	(5.50) ****						
Δm_{jt} *Newfund							0.0730	(0.81)	0.0597	(0.67)	0.0645	(0.72)
Δm_{jt} *(1-Newfund)							0.7716	(8.30)	0.7493	(7.88)	0.7364	(7.87) ****
Gender	0.24130	(3.15)	0.28249	(3.72)	0.25458	(3.44) ***	0.24212	(3.16)	0.28280	(3.72)	0.25525	(3.45) ***
Age	-0.30412	(-7.41)	-0.27507	(-6.88)	-0.30294	(-7.43) ****	-0.30508	(-7.45)	-0.27661	(-6.91)	-0.30392	(-7.44) ****
Age^2	0.35695	(7.40)	0.32089	(6.84)	0.35549	(7.38) ****	0.35789	(7.45)	0.32255	(6.85)	0.35647	(7.40) ****
Log_amt	-0.89996	(-7.19)	-1.22163	(-10.96)	-0.92464	(-7.56) ****	-0.91230	(-7.36)	-1.22926	(-11.08)	-0.93586	(-7.73) ****
Number of choices	-6.43071	(-83.98)	-6.44726	(-89.25)	-6.45847	(-89.42) ****	-6.43054	(-83.94)	-6.44676	(-89.14)	-6.45804	(-89.30) ****
Slweight1	-0.55488	(-55.76)	-0.55958	(-60.81)	-0.55966	(-61.00) ****	-0.55519	(-55.90)	-0.55987	(-60.92)	-0.55994	(-61.12) ****
Risk category	1.10743	(2.78)	1.01810	(2.55)	1.10808	(2.78) **	1.12182	(2.81)	1.03325	(2.58)	1.12200	(2.81) **
New fund dummy	13.96487	(51.89)	13.89835	(56.23)	13.87913	(55.84) ****	14.02146	(51.81)	13.95409	(56.17)	13.93417	(55.74) ****
Fee	0.03770	(0.25)	0.12982	(1.06)	0.01887	(0.13)	0.05032	(0.33)	0.13875	(1.13)	0.03130	(0.21)
Return6	10.60002	(42.51)	10.15528	(31.76)	10.66428	(41.89) ****	10.57708	(42.40)	10.14814	(31.63)	10.64116	(41.74) ****
Domestic fund management company	-3.16669	(-35.36)	-3.11356	(-35.19)	-3.13249	(-36.20) ****	-3.16674	(-35.38)	-3.11488	(-35.19)	-3.13328	(-36.21) ****
Flows	3.31049	(4.81)	2.74615	(4.84)	3.27519	(4.75) ****	3.33382	(4.83)	2.78901	(4.92)	3.29722	(4.77) ****
Alphabet*(1-Alphabet)	5.79329	(0.75)	6.47897	(0.84)	5.96927	(0.77)	-2.10466	(-0.86)	6.32463	(0.82)	5.82508	(0.75)
M_{jt}^{*}	85.05276	(6.26)	90.47987	(6.59)	87.21895	(6.40) ****	87.89920	(6.29)	93.10464	(6.61)	89.90777	(6.43) ****
Year fixed effects	N		Y		Y		N		Y		Y	
Municipality fixed effects	Y		N		Y		Y		N		Y	
Adj. R^2	0.5621		0.562		0.5859		0.5622		0.5621		0.5622	

Table 3, Panel C: Controlling for the size of the *category*

	Fund-based estimate						Category-based estimate					
	(1)		(2)		(3)		(4)		(5)		(6)	
Variable	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
Δm_{it} *Newfund*SmallCategory	-3.7065	(-10.63)	-3.6554	(-10.60)	-3.6514	(-10.58) ****	-2.4531	(-8.07)	-2.3538	(-7.91)	-2.4137	(-8.03) ****
Δm_{it} *(1-Newfund) *SmallCategory	-0.8758	(-5.32)	-0.9910	(-5.98)	-0.8822	(-5.37) ****	-2.4790	(-7.96)	-2.2805	(-7.37)	-2.5406	(-7.98) ****
Δm_{it} *Newfund*LargeCategory	-0.2283	(-2.66)	-0.2335	(-2.73)	-0.2361	(-2.76) **	-0.0769	(-0.93)	-0.1211	(-1.47)	-0.0793	(-0.96)
Δm_{it} *(1-Newfund) *LargeCategory	0.7292	(11.80)	0.7475	(12.09)	0.7125	(11.36) ****	0.5496	(6.55)	0.4959	(6.06)	0.5425	(6.50) ****
Newfund*SmallCategory	9.8925	(12.95)	10.0109	(13.18)	10.1575	(13.15) ****	38.4846	(53.61)	38.2894	(53.79)	38.1528	(54.55) ****
(1-Newfund) *SmallCategory	38.3450	(49.06)	38.4584	(49.01)	38.5173	(48.75) ****	59.3495	(82.63)	59.2028	(82.54)	58.9670	(84.36) ****
Newfund*LargeCategory	8.2877	(9.83)	8.4164	(10.12)	8.5735	(10.11) ****	36.7063	(50.43)	36.5390	(50.61)	36.3588	(51.25) ****
(1-Newfund) *LargeCategory	38.3977	(48.15)	38.6311	(48.14)	38.6989	(47.92) ****	59.8040	(82.32)	59.5865	(82.30)	59.4679	(83.79) ****
Gender	0.1066	(2.10)	0.1071	(2.21)	0.1094	(2.24) *	0.2503	(4.10)	0.2839	(4.71)	0.2540	(4.18) ***
Age	-0.0206	(-0.81)	-0.0210	(-0.83)	-0.0215	(-0.84)	-0.1871	(-5.19)	-0.1519	(-4.43)	-0.1844	(-5.12) ****
Age^2	-0.0144	(-0.48)	-0.0125	(-0.42)	-0.0127	(-0.43)	0.2271	(5.30)	0.1838	(4.51)	0.2237	(5.22) ****
Log_amt	-0.6764	(-6.83)	-0.6396	(-6.53)	-0.6514	(-6.72) ****	-0.3842	(-3.26)	-0.7151	(-7.60)	-0.3901	(-3.33) **
Number of choices	-1.2807	(-12.61)	-1.2730	(-12.80)	-1.2779	(-12.78) ****	-6.1482	(-91.25)	-6.1526	(-93.10)	-6.1569	(-91.91) ****
Slweight1	-0.3962	(-52.39)	-0.3993	(-54.88)	-0.3988	(-55.34) ****	-0.2330	(-38.11)	-0.2347	(-39.13)	-0.2345	(-39.16) ****
Risk category	1.2515	(2.80)	1.3017	(2.92)	1.3058	(2.93) **	1.0416	(2.74)	0.9111	(2.38)	1.0415	(2.74) **
Fee	-0.5313	(-5.06)	-0.5725	(-5.16)	-0.5401	(-5.06) ****	-0.5446	(-6.40)	-0.4228	(-5.31)	-0.5569	(-6.46) ****
Return6	11.2601	(35.10)	11.3285	(38.42)	11.2472	(37.71) ****	8.1067	(32.12)	7.4666	(29.22)	8.1330	(31.65) ****
Domestic fund management company	-3.2906	(-29.67)	-3.2492	(-29.09)	-3.2645	(-29.80) ****	-2.5318	(-20.36)	-2.5057	(-20.31)	-2.5124	(-20.55) ****
Flows	7.0479	(16.56)	7.1397	(16.29)	7.0314	(16.56) ****	0.7332	(1.52)	-0.0443	(-0.09)	0.7088	(1.45)
Alphabet*(1-Alphabet)	-28.1734	(-4.38)	-28.4347	(-4.46)	-28.4951	(-4.46) ****	-8.3553	(-2.89)	-9.1722	(-2.98)	-8.6070	(-2.92) **
Year fixed effects	N		Y		Y		N		Y		Y	
Municipality fixed effects	Y		N		Y		Y		N		Y	
Adj. R^2	0.5888		0.5898		0.5898		0.7018		0.7024		0.7026	

Among the control variables, past returns affect the choice of fund: a 1% increase in returns by a fund in the previous six months raises the investment in the fund by 10.6%. Participants consistently choose to invest in funds run by non-Swedish fund management companies. Income also negatively affects changes in fund holdings. Controlling for income, age has a negative effect. Moreover, investors seem to tilt their portfolios towards riskier and less expensive funds. The results are robust to controlling for local effect (municipality-fixed effect, see columns (2), (3), (5) and (6)), as well as to removing groups of funds or counties.

Finally, we focus on a variable ("*Slweight1*") that accounts for the fraction of the portfolio invested prior to rebalancing in all the funds that belong to the same category as the *fth* fund with the exception of fund f[11]. The coefficient is negative. This suggests that, *all else being equal*, having invested in a category decreases the unconditional probability of investing in another fund of the same category.

The results suggest that investors react to exogenous changes in the representation in the menu of the assets they are offered. The positive and normative implications of these findings are striking. If demand is induced by the offering of new funds, supply creates its own demand. That is, styles in which more funds are offered will command higher demand, independently of the specific characteristics of the stocks and this will affect the equilibrium conditions of the stocks in those categories. For example, if product market competition between mutual fund families induces these to offer more growth funds as opposed to value funds, this will tilt the demand for growth assets in the market, reduce their price and increase their required rate of return, effectively shaping the value premium.

[11] We also looked at the full fraction of portfolio invested in the same category. The results do not change.

6. The nature of menu exposure

In the previous sections we ascertained the existence of a menu effect on portfolio choice. Now, we study whether this is information-induced or is just a behavioral heuristics. We argued that, in the case of a bias, there should not be any correlation between the degree of informativeness of the investors and their menu exposure. We therefore consider two proxies for information and one proxy for behavioral bias. The two proxies for information are the ability to select well-performing funds and the degree of concentration of the investor portfolio. The proxy for behavioral bias is the investor trading pattern.

6.1 Menu exposure and investor trading

Trading behavior reveals information about the investor. Indeed, if trading is information motivated and menu exposure is a way for investors to compensate for lower private information, we would expect a negative relation between it and menu exposure. Indeed, in the case the exposure were information-induced, investors more exposed to it should also be less willing to trade on the basis of their information (e.g. Hellwig, 1980, 1982). If, instead, menu exposure were associated with some behavioral bias, there should not be any correlation between trading and menu exposure. In fact, in the case of overconfidence (e.g. Odean, 1998; Barber and Odean, 2000), we would expect a positive relation between it and trading. We can, therefore, use trading to help shed light on the nature of the menu exposure.

In the PPM system, there is no limitation on the number of transactions. Transactions can be placed electronically (via a website) or by filing a paper form. Investors can choose to keep their old allocation, change only one fund, or change all five funds. There are no transaction

fees, as well as no front and end-load fees. Normally it takes between one and three days to process the transaction. There is no automatic rebalancing. Thus, if the allocation is to creep out of "optimal", it is the responsibility of the investor to adjust it in due course[12].

For each investor, we define trading as either the number of trades (rebalancings) the investor has placed within a given half year, or as portfolio turnover. Rebalancing is defined as the number of times during a six-month period the changes were made. Portfolio turnover is defined as the sum of the absolute values of the weight changes. We do not count events that are later marked as erroneous. Then, for each investor, we regress our proxy for trade – i.e. either rebalancing or portfolio turnover – on the exposure of the investor to categories that are over-represented in the menu ("menu exposure", *E*) and the control variables:

$$T_{it} = \alpha + \beta E_{it} + \gamma P_{it-1} + \delta controls_{it} + \varepsilon_{it} \qquad 3)$$

where for the *ith* individual, T_{it} is the aggregate measure of trading of the *ith* investor at time *t* and E_{it} is the *ith* investor's menu exposure. Menu exposure is defined as the exposure to the menu tilt. It is constructed as the weighted-average of the over-representation (tilt) in the menu of the categories of the funds the investor holds. It is: $E_{it} = \sum_{j=1}^{J} w_{jt} \left| M_{jt} \right|$. That is, for the *ith* investor, at time *t*, we calculate the weighted-average of the absolute values of the over-representation in the menu of the categories in which the *ith* investor holds funds. The weights are given by the fraction of the portfolio invested in each *j* category. For example, if the investor invests in two funds – 30% in a growth fund and 70% in a balanced fund – *E* is given by $0.3*M_g+0.7*M_b$, where M_g and M_b are the degree of over-representation of, respectively, the growth and balanced category, as

[12] A paper statement of the account is mailed to all PPM participants once a year.

defined before. The absolute value allows us to control for any deviation from the market portfolio. Indeed, categories may be over-represented (positive deviation) as well as under-represented (negative deviation). Any deviation may trigger the investor's portfolio tilt. The absolute value captures both types of deviations. We consider two measures of menu exposure: one based on M_j and one based on M_j^*. The control variables include the lagged dependent variable; demographic characteristics (age, age-squared, gender), and the amount invested in the PPM system.

The fact that we can estimate equation (3) only for the individuals who trade, induces a selection bias. To address this issue, we estimate a Heckman two-stage specification. The first stage uses a probit model in which the decision to trade is explained in terms of the menu exposure as well as of a set control variables. The control variables are menu exposure, menu exposure squared, gender, age, and age squared. The identifying restriction in the Heckman model is provided by the square of menu exposure and 26 county dummies.

The results are reported in Table 4. In Panel A, we report the results for the tilt measure based on M_j, while in Panel B, we report the results based on M_j^*. In the first column, we display the results for the first stage of the Heckman estimator (the decision to trade). We will directly focus on the frequency of trading. The results show that individuals who are more exposed to the menu bias trade less. Indeed, there is a strong and significant negative correlation between trading and menu exposure. The more the investors are invested in categories that are over-represented in the menu, the less they trade. In particular, investors who display a one standard deviation higher exposure to the over-represented categories have a 11% lower trading and 8% lower turnover with respect to the investor's unconditional means.

Table 4: Trade

We report the results of regression measures of investor trade (number of trades, turnover) on the investor menu exposure (E) and a set of control variables. The measures of investor trade are defined over 6 month-periods. E_{it} (E^*_{it}) is the *ith* investor's exposure to the menu bias and is defined as the average of the menu (Morningstar menu) exposure of the categories of the funds the investor holds. It is constructed as: $E_{it} = \sum_{j=1}^{J} w_{jt} \left| M_{jt} \right|, E^*_{it} = \sum_{j=1}^{J} w_{jt} \left| M^*_{jt} \right|$, where the weights are the weights of the categories in the investor portfolio. We report the results for the first stage (the decision to trade, reported in column 1) of the Heckman estimation and the second stage estimates. These estimations use year fixed effects, 290 municipality fixed effects and various combination of them. Panel A (B) reports the results for E_{it} (E^*_{it}). All the other variables are defined in Table 1. The estimates are done using White heteroscedasticity-consistent estimator with clustering of standard errors over municipalities. T-statistics are reported in parentheses. Estimates for Age^2 are multiplied by 100. For specifications with both year and municipality fixed effects we also report the level of significance obtained via bootstrapped estimates with 10000 resamplings. Significance on 10%, 5%, 1% and 0.1% is denoted as *, **, *** and ****, respectively.

Table 4, Panel A: Results for E_{it}

	1st stage		*ntrades*		*Turnover*		*ntrades*		*Turnover*		*ntrades*		*Turnover*	
	Coeff.	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
E_{it}	-0.403	(-9.18)	-0.486	(-22.63)	-35.959	(-17.94)	-0.286	(-63.65)	-20.740	(-40.83)	-0.485	(-23.95) ****	-35.769	(-19.34) ****
E_{it}^2	-0.022	(-0.39)												
Lagged Dependent Var			0.400	(27.15)	0.407	(13.49)	0.713	(82.34)	0.704	(86.68)	0.403	(30.97) ****	0.410	(14.87) ****
Gender	0.062	(43.87)	0.074	(22.10)	4.572	(14.64)	0.037	(34.44)	2.147	(15.46)	0.074	(22.47) ****	4.603	(13.25) ****
Age	0.020	(7.70)	0.030	(40.95)	-0.187	(-2.12)	0.029	(55.86)	-0.138	(-1.20)	0.030	(36.08) ****	-0.132	(-0.89)
Age^2	-0.022	(-10.11)	-0.032	(-59.94)	0.317	(4.96)	-0.030	(-51.19)	0.330	(2.44)	-0.032	(-53.75) ****	0.250	(1.86)
Log_amt			0.016	(8.06)	0.952	(6.60)	0.031	(2.62)	-3.704	(-1.89)	0.001	(0.19)	-3.894	(-5.43) ***
Heckman lambda			1.390	(31.31)	97.026	(21.32)	-0.031	(-5.17)	-1.998	(-9.70)	1.387	(34.68) ****	96.696	(22.38) ****
Intercept	-0.074	(-0.62)	1.300	(28.66)	99.653	(25.88)	1.355	(100.30)	101.445	(43.00)	1.301	(28.66) ****	98.685	(19.78) ****
Year fixed effects			Y		Y		N		N		Y		Y	
Municipality fixed effects			N		N		Y		Y		Y		Y	
Adj R2	0.124		0.504		0.406		0.492		0.390		0.507		0.407	

Table 4, Panel B: Results for E^*_{it}

	1st stage		*ntrades*		*Turnover*		*ntrades*		*Turnover*		*ntrades*		*Turnover*	
	Coeff.	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
E^*_{it}	-6.926	(-5.20)	-3.118	(-35.90)	-248.133	(-55.45)	-3.268	(-25.46)	-252.304	(-38.07)	-3.116	(-35.05) ****	-246.829	(-44.81) ****
$E^*_{it}{}^2$	-82.343	(-6.94)												
Lagged Dependent Var.			0.715	(69.55)	0.697	(78.24)	0.711	(82.17)	0.702	(85.41)	0.715	(72.76) ****	0.697	(84.62) ****
Gender	0.004	(3.26)	0.040	(25.16)	2.303	(28.28)	0.037	(30.29)	2.114	(14.10)	0.040	(24.88) ****	2.363	(15.23) ****
Age	0.015	(19.03)	0.023	(27.34)	-0.536	(-6.55)	0.030	(59.81)	-0.113	(-1.01)	0.023	(31.36) ****	-0.467	(-19.04) ****
Age^2	-0.024	(-31.21)	-0.024	(-25.11)	0.740	(8.34)	-0.030	(-54.11)	0.299	(2.27)	-0.024	(-29.70) ****	0.656	(17.98) ****
Log_amt			0.033	(6.59)	2.179	(6.99)	0.034	(2.88)	-3.517	(-1.80)	0.014	(1.25)	-4.367	(-2.09)
Heckman lambda			-0.009	(-11.58)	-0.294	(-10.14)	-0.054	(-15.27)	-3.531	(-32.67)	-0.009	(-16.80) ****	-0.291	(-20.44) ****
Intercept	1.326	(10.19)	1.683	(236.48)	125.089	(112.03)	1.530	(75.09)	114.818	(40.94)	1.690	(214.16) ****	124.028	(149.81) ****
Year fixed effects			Y		Y		N		N		Y		Y	
Municipality fixed effects			N		N		Y		Y		Y		Y	
Adj R2	0.127		0.498		0.396		0.493		0.392		0.500		0.398	

This would be consistent with an information-based story. Less informed investors are less willing to trade. This would, however, also be consistent with a positive relationship between menu exposure and higher risk aversion. That is, it may be the case that the investors who are less risk averse – more conservative, more willing to herd with the pack and therefore more exposed to menu bias – are also those who are more exposed to the menu representation (e.g. Hellwig, 1980, 1982). To distinguish between these two interpretations, we focus on portfolio performance.

6.2 Menu exposure and fund-picking skills

As we mentioned before, the less informed investors are, the less able they will be to choose better-performing funds. In the case of a bias, there should not be any (and definitely not a negative) correlation between the ability to select funds and menu exposure. Indeed, the bias is likely to affect all the investors in a similar way. That is, if the sensitivity of the investment to the menu is just a behavioral heuristic, the ability of the investors to select over-performing funds should not be systematically related to the way their investment is affected by the representation in the menu.

If, instead, more uninformed investors use the representation in the menu as a cheap source of information, there should be a *negative* correlation between the ability of the investors to select over-performing funds and their sensitivity to the menu composition. Indeed, in the case the exposure were information induced, we would expect the investors more exposed to the menu bias to be the less informed investors – that is, less capable of selecting over-performing funds. This would suggest that these investors use the menu representation as a source of information.

We consider as a measure of ability the skill of picking over-performing funds, that is, funds whose return is higher than that of the other funds *within the same category*. Notice that we are not studying the ability to select funds that provide higher gross returns, as this would depend upon the type of category that is preferred and this choice may be related to the investor's long-term objectives. For example, a young investor, having many years before retirement, may prefer equity funds. If the equity market were over – performing, we would erroneously attribute a simple, long-term portfolio allocation to skill.

For each fund, we calculate the difference between its return and the average return of all the other funds belonging to the same category. This represents the performance of the fund. Then we calculate the weighted-average of the performance of all the funds the investor holds. This represents the performance of the investor's portfolio. We then regress this measure of portfolio performance on our proxy for menu exposure (E) and a set of control variables:

$$P_{it} = \alpha + \beta E_{it} + \gamma P_{it-1} + \delta controls_{it} + \varepsilon_{it}, \qquad 4)$$

where for the *ith* individual, P_{it} is the performance of the portfolio of funds of the *ith* investor at time *t.* We consider four alternative measures of portfolio performance: the value weighed (according to the investment in each fund) difference between fund return and the average return of all the other funds belonging to the same category; the value weighed (according to the investment in each fund) difference between fund return and the average return of all the other funds belonging to the same category standardized by the volatility of funds; the equally weighed difference between fund return and the average return of all the other funds of the same category, and the equally weighed difference between fund return and the average return of all the other funds belonging to the

same category standardized by the volatility of funds. We also consider 6- and 12-month performance measures. E_{it} as well as the control variables are defined as in the previous specifications.

The results are reported in Table 5. In Panel A, we use as measure of performance the difference between the performance of the chosen funds and the average performance of the funds in the same category. In Panel B, we normalize the performance measure by the standard deviation of the funds in the same category.

The findings support our intuition. The more individuals invest in categories that are over-represented in the menu, the more they tend to choose worse-performing funds. An increase in the exposure to the over-represented categories by one standard deviation reduces by 0.54% per annum the performance of the funds the investor has selected. This ranges between 32% and 37% of the dependent variable's unconditional mean. In the context of the pension system it can result in about 16% changes in expected pension income[13].

It is interesting to note that, in terms of age, young individuals are better pickers of funds. These results are robust across specifications as well as alternative definitions of our exposure variables (equally or value-weighted, raw or standardized measures of performance). The results for gender show that, while performance is roughly the same for males and females, standardized performance is higher for males.

In the previous section, we related menu exposure to trading and we showed that there is a negative correlation. It may be that part of the lower performance is not only due to the choice of worse-performing funds, but also to the fact that investors exposed to menu bias trade less.

[13] We assume a 45–year horizon and real risk premia of 4% per annum.

Table 5: Investment performance

We report the results of regression measures of investor performance on the investor menu exposure (E) and a set of control variables Performance is defined over 6 month-horizon. E_{it} (E^*_{it}) is the *ith* investor's exposure to the menu exposure and is defined as the average of the menu (Morningstar menu) exposure of the categories of the funds the investor holds. It is constructed as: $E_{it} = \sum_{j=1}^{J} w_{jt} \left| M_{jt} \right|, E^*_{it} = \sum_{j=1}^{J} w_{jt} \left| M^*_{jt} \right|$, where the weights are the weights of the category in the investor portfolio. Panel A looks at performance relative to category average, while Panel B looks at the same variable standardized by the standard deviation of funds' performance. In columns (2) and (3), 290 municipalities fixed effects are used. In columns (1) and (3), year fixed effects are also employed. Panel C looks at performance and turnover. Columns 1–4 report the results of the regressions for performance relative to category average, while columns 5–8 report the results for the standardized measures of performance. In columns (1), (2), (5) and (6) we used residuals of regression of turnover on menu exposure measures. In columns (3), (4), (7) and (8) we interact menu exposure with dummies of High and Low turnover (dummies that take value 1 if turnover is above or below the mean for given time period, and zero otherwise). The reported estimates are done with year fixed effects and 290 municipalities fixed effects. All the other variables are defined in Table 1. All the estimates are done using White heteroscedastisity-consistent estimator with clustering over municipalities. The *t-statistics* are reported in parentheses. All the estimates are multiplied by 100 with the exception of Age^2, which is multiplied by 10,000.

Table 5, Panel A: Performance relative to category average

	Coeff.	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
E_{it}	-1.2943	(-104.41)			-1.1336	(-46.06)			-1.2928	(-112.37)		
E^*_{it}			-1.0878	(-57.51)			-0.9431	(-29.15)			-1.0862	(-61.66)
Gender	-0.0046	(-1.02)	-0.0064	(-1.38)	-0.0243	(-15.10)	-0.0262	(-14.66)	-0.0087	(-6.75)	-0.0105	(-7.62)
Age	0.0115	(5.58)	0.0120	(6.07)	0.0432	(9.92)	0.0437	(9.84)	0.0083	(5.89)	0.0088	(5.99)
Age^2	-0.0200	(-12.25)	-0.0206	(-13.23)	-0.0536	(-9.58)	-0.0542	(-9.51)	-0.0162	(-6.84)	-0.0168	(-6.88)
Log_amt	-0.2309	(-22.87)	-0.2308	(-22.84)	-0.0101	(-0.78)	-0.0106	(-0.82)	0.0166	(1.26)	0.0165	(1.26)
Lagged dependent variable	-20.0828	(-160.06)	-20.1402	(-156.71)	-21.0389	(-187.10)	-21.0968	(-182.76)	-20.1065	(-163.21)	-20.1639	(-160.25)
Year fixed effects	Y		Y		N		N		Y		Y	
Municipality fixed effects	N		N		Y		Y		Y		Y	
Adj. R^2	0.0863		0.0857		0.0742		0.0737		0.0870		0.0864	

Table , Panel B: Standardized performance relative to category average

	Coeff.	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
E_{it}	-7.3421	(-9.31)			-5.9218	(-2.24)			-7.3069	(-9.70)		
E^*_{it}			-6.1079	(-7.51)			-4.7867	(-2.34)			-6.0730	(-7.80)
Gender	0.3428	(5.07)	0.3316	(4.77)	0.6196	(28.57)	0.6008	(24.82)	0.3630	(4.37)	0.3518	(4.14)
Age	0.3187	(34.08)	0.3225	(32.80)	0.3744	(38.68)	0.3808	(36.42)	0.3366	(27.86)	0.3404	(29.43)
Age^2	-0.3440	(-38.66)	-0.3480	(-38.87)	-0.3170	(-27.60)	-0.3250	(-30.00)	-0.3640	(-18.48)	-0.3680	(-19.31)
Log_amt	-0.9101	(-6.33)	-0.9090	(-6.32)	-0.2403	(-1.42)	-0.2445	(-1.45)	-2.0710	(-9.51)	-2.0707	(-9.52)
Lagged dependent variable	-33.8982	(-86.63)	-33.9484	(-86.24)	-32.2133	(-132.91)	-32.2616	(-131.24)	-33.8632	(-81.66)	-33.9134	(-81.29)
Year fixed effects	Y		Y		N		N		Y		Y	
Municipality fixed effects	N		N		Y		Y		Y		Y	
Adj. R^2	0.1860		0.1860		0.1681		0.1680		0.1869		0.1867	

Table 5, Panel C: Performance and turnover

	Coeff.	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
E_{it}	-1.261	(-95.56)							-6.722	(-8.52)						
E^*_{it}			-1.056	(-52.61)							-5.542	(-6.84)				
E_{it}*Low Turnover					-0.436	(-3.89)							1.989	(1.35)		
E_{it}*High Turnover					-1.460	(-21.74)							-10.727	(-7.93)		
E^*_{it}*Low Turnover							-0.799	(-4.06)							0.757	(0.56)
E^*_{it}*High Turnover							-1.214	(-15.63)							-9.019	(-6.42)
Lagged dep. variable	-20.607	(-140.03)	-20.666	(-137.16)	-20.679	(-155.82)	-20.733	(-159.54)	-34.535	(-79.55)	-34.585	(-79.36)	-34.611	(-80.59)	-34.680	(-82.64)
Gender	-0.017	(-11.87)	-0.019	(-12.20)	-0.017	(-14.19)	-0.019	(-14.73)	0.197	(2.48)	0.188	(2.32)	0.182	(2.45)	0.175	(2.30)
Age	0.009	(5.33)	0.010	(5.42)	0.011	(11.96)	0.011	(16.51)	0.354	(40.71)	0.357	(42.49)	0.394	(26.67)	0.402	(22.73)
Age^2	-0.017	(-6.35)	-0.018	(-6.39)	-0.020	(-10.97)	-0.020	(-13.75)	-0.391	(-27.05)	-0.394	(-28.27)	-0.437	(-19.13)	-0.446	(-16.96)
Log_amt	0.037	(2.94)	0.037	(2.95)	0.037	(2.99)	0.037	(2.98)	-1.679	(-7.34)	-1.678	(-7.34)	-1.679	(-7.27)	-1.682	(-7.31)
Turnover					0.001	(7.23)	0.001	(6.37)					0.025	(16.28)	0.025	(13.89)
Res_turnover	0.143	(11.02)	0.144	(11.16)					2.753	(26.35)	2.762	(26.91)				
Intercept	1.015	(79.16)	0.964	(71.23)	0.801	(42.87)	0.745	(26.19)	-9.014	(-40.09)	-9.273	(-48.72)	-13.179	(-77.67)	-13.678	(-77.09)
Adj R2	0.088		0.087		0.088		0.087		0.189		0.189		0.190		0.189	

To address this issue, in Panel C, we interact our measures exposure with dummies that proxy for whether the investor is trading a lot. While we lack good instruments to deal with the endogeneity of the turnover measures, we can try to separate the effects of willingness to trade from those of menu exposure. In columns (1), (2), (5) and (6), we control for turnover using the residuals of a regression of turnover on the corresponding menu bias measure. The resulting variable (*res_turnover*) is purged of the direct effect of menu bias on turnover. In columns (3), (4), (7) and (8), we interact menu bias with dummies for high and low turnover (the dummy for high turnover takes the value one if turnover for a given investor is higher than the mean, and zero otherwise)[14]. The results show that the negative relationship is robust to controlling for turnover measures. The results are consistent across specifications and suggest that the negative performance is mostly due to scarce ability to pick well-performing funds.

6.3 Menu exposure and portfolio concentration

An alternative proxy for information used in the literature is the degree of concentration of the investor portfolio. The literature has shown that there is a positive correlation between portfolio concentration and performance, arguing that this implies that better informed investors do hold a more concentrated portfolio (Kacperczyk *et al.*, 2004). We therefore expect that, if the better informed investors are less subject to menu exposure, there should be a negative relation between menu exposure and the degree of portfolio concentration. We therefore regress the degree of portfolio concentration of the investors (*H*) on their menu exposure and a set of control variables:

[14] The results for number of trades are qualitatively similar. We omit them in the interest of brevity.

$$H_{it} = \alpha + \beta E_{it} + \gamma controls_{it} + \varepsilon_{it} \qquad 5)$$

where H_{it} is the degree of portfolio concentration of the *ith* investor at time *t.* This is the Herfindhal index of concentration of the portfolio, where each fund is represented on the basis of its weight in the portfolio. The menu exposure (E_{it}) as well as the control variables are defined as before.

The results are reported in Table 6. They show that an increase in menu exposure is related to lower portfolio concentration. In particular, an increase in menu exposure equivalent to one standard deviation is related to a lower concentration in the portfolio of the investor equivalent to 18% of the unconditional mean. This explains about 35% of its standard deviation. That is, the more the investors have chosen categories that are over-represented in the menu, the less concentrated their portfolios are. This confirms the previous findings and suggests a negative correlation between investor informativeness and choosing funds as a function of their representation in the menu.

All these findings taken together suggest that investors tend to invest more in the funds that are more represented. The investors who are more subject to this behavior are the ones who trade less, select the worse funds and have a less concentrated portfolio. We argue that this suggests that the bias is a way of coping with limited information. Given that individuals are fully aware of all the choices available in the menu, the lower information does not arise from limited awareness of the options, but is most likely related to the availability of private information.

Table 6: Portfolio concentration

We report the results of regression measures of investor portfolio concentration on the investor menu exposure (E) and a set of control variables. Portfolio concentration is defined using the Herfindahl index of the fraction of the investor's portfolio invested in the different funds. All the variables are as defined in Tables 1 and 4. In columns (2) and (3), 290 municipalities fixed effects are used. In columns (1) and (3), year fixed effects are also employed. All the estimates are done using White heteroscedastisity-consistent estimator and clustered over municipalities. T-statistics is reported in parentheses.

	(1)				*(2)*				*(3)*			
	Coeff.	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*	*Coeff.*	*t-stat*
E_{it}	-0.1876	(-12.30)			-0.1267	(-6.98)			-0.1829	(-14.29)		
E^*_{it}			-0.2586	(-11.33)			-0.2141	(-11.68)			-0.2519	(-12.87)
Gender	0.0079	(17.90)	0.0078	(18.34)	0.0069	(25.44)	0.0067	(25.03)	0.0073	(49.41)	0.0072	(51.37)
Age	-0.0135	(-21.10)	-0.0134	(-20.81)	-0.0130	(-72.81)	-0.0129	(-70.17)	-0.0140	(-44.22)	-0.0139	(-43.27)
Age^2	0.0002	(25.81)	0.0002	(25.45)	0.0002	(78.57)	0.0002	(75.82)	0.0002	(54.21)	0.0002	(53.01)
Log_amt	-0.0052	(-5.56)	-0.0052	(-5.65)	0.0041	(4.63)	0.0039	(4.42)	0.0236	(4.44)	0.0235	(4.44)
Lagged dependent variable	0.2021	(3.89)	0.2010	(3.89)	0.1346	(17.33)	0.1330	(17.35)	0.2287	(5.81)	0.2275	(5.81)
Year fixed effects	Y		Y		N		N		Y		Y	
Parish fixed effects	N		N		Y		Y		Y		Y	
Adj. R^2	0.0826		0.0834		0.0640		0.0648		0.0880		0.0888	

If phenomena such as the home bias puzzle can be explained in terms of menu exposure, this would suggest that exposure is fully rational and explainable in terms of investors trying to cope with limited information. It is therefore arguable that better information would reduce the role of such exposure. The information disadvantage also affects the degree of portfolio concentration, reducing it.

7. Conclusion

We study how the choice of investing in a class of assets is affected by the alternatives available in the class. We used the information on the choice of mutual funds that are available to Swedes to invest their retirement accounts. We have shown that investors choose the category to invest (e.g. growth funds) according to number of funds available in the category. More numerous categories attract more investment, regardless of their weight in the optimal or world market portfolio. More importantly, we have shown that an exogenous increase in the number of funds of a specific category induces investors to rebalance their portfolios so as to increase their investment in the already existing funds belonging to that category.

We define this phenomenon as "menu exposure" and argue that it requires investors to consider the representation in the menu as a sign of the quality of the fund category. By using information on the performance of the funds that investors choose and on their degree of portfolio concentration, we have shown that there is a consistent positive correlation between investors' sensitivity to the menu exposure and their lack of information. This suggests that menu exposure represents a rational way of coping with limited (private) information, although this decreases as information improves.

Our findings shed light on both the home bias puzzle and the determinants of style investing. Moreover, they have normative implications in terms of the US debate on the reform of Social Security. They suggest that the impact of the reform on the stock market will crucially depend on the type of menu offered to the investors, as this will determine the choice of fund and therefore, indirectly, the types of stocks (e.g. growth, value, big, small) that will be more affected by the reform.

Chapter 5. Stock market participation and pension reform[1]

Anders Karlsson, Andrei Simonov and Massimo Massa

1. Introduction

This chapter examines how the introduction of a funded defined contribution retirement system affects participants' propensity to participate in the stock market. This is an interesting topic, in view of the transition to funded individual account plans in many nations around the world. Here we focus on what we call the "Swedish experiment:" in the year 2000, the traditional Swedish pay-as-you go retirement system was partially replaced by a national defined contribution plan commonly referred to as the PPM system (Sunden, 2000). Our main question is whether individuals perceive their investments in the pension scheme as a substitute for direct investments, and whether allocating more equities in their pension accounts induces participants to reduce or increase their directly-held equity investments.

Standard portfolio choice theory suggests that investors would choose an optimal overall market exposure taking into account both direct and indirect (e.g., through a pension scheme) investment in equity. That is, pension and nonpension equity holdings would be seen as substitutes, at least to a first order. By contrast, behavioral finance theorists instead hypothesize that investors tend to categorize their investments, along with associated gains and losses, according to narrow categories. This approach, labeled "mental accounting" or "narrow framing," predicts that investors apply such mental accounting to stock holdings and react separately to gains and losses for different stocks (Barberis and Huang, 2001).

[1] Published Oct 2007 by the Pension Research Council at the Wharton School as a chapter in: "Transitioning to Retirement: How Will Boomers Fare?"

The implications of the two theories are starkly different. Standard portfolio theory would posit that inducing investors to hold their retirement accounts in pension funds would crowd out direct equity investment. Behavioral theory, instead, suggests that if investors perceive their investment in pension funds and in equity to be different sorts of accounts, one investment would not necessarily crowd out the other. The best way to examine this issue is to see whether introducing a defined contribution individual account system changes investor incentives to directly participate in the stock market. To do a proper analysis, we would require information on investors' portfolio choices before the change, and to explore what happened after the introduction of the new pension system. Until now, this sort of information has been unavailable.

Our contribution is to exploit data on a large and representative sample of Swedish individuals tracked over time. Not only do we have information on their wealth, their income, their tax position, and their demographic characteristics, but also we have semiannual information on the value of each investor's stock positions as well as the daily value of his transactions in the PPM system. In particular, the dataset on the individual retirement accounts includes all individual choices made from the introduction of the system until October of 2004. We use the evidence provided on investors' patterns in the PPM system to test whether having a pension account containing mutual fund investments changes investor incentives to participate directly in the stock market.

The results show that investors do not perceive direct investment in the equity market as a close substitute for their retirement accounts, suggesting that an individual account system does not crowd out direct equity market investment. The new Swedish system may actually help educate investors of the benefits of stock market participation, increasing participation and therefore, indirectly, boosting saving. In what follows, we first describe the Swedish experiment and provide

institutional details. Next we describe the data and our main evidence. A short conclusion offers policy implications.

2. The Swedish "Experiment"

The Swedish government made changes to its old pay-as-you-go national retirement program in the Fall of 2000. The "basic" portion of the system is a guaranteed benefit designed to ensure that no retiree will be completely without benefits in retirement, regardless of her or his previous income. There are two earnings-related elements of the new pension system, one being the Notional Defined Contribution (NDC) plan, financed by a tax of 16 percent of annual pay; revenue from this system is used to finance current retirees, and the amount paid in also serves as a base in calculating future pension payments. The second earnings-related component, of most interest here, is the premium pension (PPM) portion. This is financed by a mandated contribution 2.5 percent of annual pay, which is invested according to each individual's discretion. At inception, each participant was mailed a prospectus describing the investment menu of 464 funds, from which the participant had to elect from 1 to 5 of these choices.

Figure 1 represents the information provided to investors in the PPM system. In addition to the fund identification number, name, and fund family, information is provided on fees, past return, and risk. Risk is represented by two measures. One is a simplified graph displaying a jagged red line for very risky or a flat green line for very low risk. There are five risk categories of this kind. In addition each fund has a number representing the fund's annualized standard deviation over the past 36 months. If, for some reason, no menu choice is actively elected, the participant's money is invested in the default which is the Seventh Swedish Pension Fund; this is an equity fund run by the government. The accrued amount

from the premium pension part will be paid out on a monthly basis to the individual at the time of retirement.

Figure 1: Extract from the PPM Investor Information Folder, fund example

Fund number	Fund name, Management company	Information regarding the funds	Fund fee (%)	Percentage return 99-12-31 (after fees)						Total risk
				In the year					Last 5 years	(last 3 years)
				95	96	97	98	99		
191080	Baring Global Emerging Markets Baring International Fund Managers (Ireland) Ltd	Emerging markets' equity and equity related assets	1.59	-32	10	25	-25	77	25.3	32 (Red)

The percentage return for the last five years equals the compounded annual growth rate of return for the years 1995 through 1999. The total risk corresponds to an annualised percentage standard deviation of three-year monthly historical fund returns. The total risk is also categorised into five different classes, and colours, with respect to standard deviation; Class 1: very low risk, dark green, percentage standard deviation in the range 0-2; Class 2: low risk, light green, 3-7; Class 3: average risk, yellow, 8-17; Class 4: high risk, orange, 18-24; Class 5: very high risk, red, 25-.

Source: Premiepensionsmyndigheten (2000).

In total, 18.5 percent of the annual pension based income for each individual is invested to finance this system, and all annual income from the age of 16 is included. Yet there is a contribution crediting limit: individuals who earn more than 7.5 income 'base amounts'[2] per year are only credited with 18.5 percent of 7.5 income base amounts, even though they still pay the contribution on their entire pension-based income.[3] The first pension investments in the PPM system, in

[2] For the year 2006, one income base amount equals 44,500 SEK

[3] For example, for a person with 360,000 SEK as her pension-based income in the year 2006, only 18.5% of 333,750 SEK (7.5 * 44,500) will count towards her pension. In other words, her defined contribution to PPM in 2006 is 8,344 SEK (0.025*333,750), which is the maximum

October of 2000, involved 4.4 million individuals. At that time, workers' initial contribution was set at two percent of pay for 1997-98 and 2.5 percent for 1999 and 2000. The average amount invested was 12,651 SEK for the entire population and 13,506 SEK for those who made an active investment election; [4]the maximum invested was 26,202 SEK.

When the PPM system was introduced, the government and the mutual fund industry launched a massive advertising campaign (Cronqvist and Thaler, 2004; Cronqvist, 2006). Though low on information content (level of fees, risks, etc), the ads did help create a positive image of investing in financial markets. More than 86% of all investors were exposed to TV ads, 75% were exposed to some advertising in print media, 59% saw some kind of outdoor ad, and 36% listened to radio ads (Cronqvist, 2006). Accordingly, virtually the entire population was exposed to information about investing in financial markets at the time of the PPM launch.

Figure 2 shows how the Swedish stock market peaked in March of 2000 showing a 15 month gain of 141% (January 1999-March 2000). In the second half of 2000 (when the PPM was introduced), the market fell 20%, and it then continued to fall for the first 6 months of 2001, dropping about 15% further. About 6 percent of our sample population was invested directly in the stock market in 1999-06-30. This figure increased to 8.5% by June 2000 and further to 9 percent by the end of 2000 and 9.35% by June 2001. Thus, our three dates after this extreme bull market show an increase in stock market participation. This increased interest for the stock market could be a delayed reaction of the '99 bull market. But it is also possible

---

contribution per individual for that year. The typical exchange rate in 2000 was 10 SEK for 1 USD and for 2006 around 7 SEK/USD

[4] The slightly higher average for active investors is consistent with Engström and Westerberg (2003) who show that active investors tend to have slightly higher income than the "default" investors.

that rising interest in the stock market despite the bear market of 2000-01 can be attributed to the substantial marketing campaigns in 2000 encouraging individuals to make an active choice in PPM.

Figure 2: MSCI-Sweden Total Return Index, in SEK (Ibbotson Associates data id IBOR002028), January 1999=100.

Source: Morgan Stanley Capital International (2006).

3. The Data

To analyze investment choices under the PPM model, we have collected data from three different sources: investment allocations in the PPM system, investor stock market holdings and investor demographics. Our sample is a representative cross-section of the Swedish work force, whose PPM investment choices have been linked to individual demographic data collected by Statistics Sweden for the year 2000. This linked dataset makes it possible to study investment behavior in great detail.

3.1 Asset Allocation Information

Data on individual retirement accounts are collected from PPM, and they indicate all individual choices made from the introduction of the system until October of 2004. Both the transaction date and clearing date are known. We also know the universe of funds that was available to investors at each point in time. Thus, for each individual, there is information on the amount invested in each fund, and in which funds the individual invested. Approximately two-thirds of the 4.4 million participants in PPM made an active investment decision. For these individuals, we can investigate the exact allocation of assets in their portfolios. The remaining third did not make an active investment decision, so their money was defaulted into the government-run equity fund.

3.2 Portfolio Data

For each investor, we have detailed information about individual stockholdings, mutual funds, bank accounts, real estate, and other types of wealth for the period 1995-2000. The Security Register Center (VPC or Vardepapperscentralen) have data on both stocks held directly and in a street name, including holdings of US-listed ADRs. In addition, SIS Ägarservice AB collects information on the ultimate owners of shares held via trusts, foreign holding companies, and the like (c.f. Sundin and Sundquist, 2002). Overall, the records provide information about the owners of 98% of the market capitalization of publicly-traded Swedish companies. For the median company, we have information about 97.9% of the equity, and in the worst case, 81.6% of market capitalization of the company. The data provided by SIS Ägarservice AB are linked by Statistics Sweden to the demographic data described below.

3.3 Additional Data

We also use data on demographic statistics reported by Statistics Sweden, the Swedish version of the U.S. Census Bureau. Data sources include HEK 2000 (a

report on household economics), IoF 2000 (a report on individual and household measures of income), and SUN 2000 (a report on educational status). These three reports represent a cross-section of the Swedish population and contain detailed information including the amount of foreign assets held per person by asset class. The data also describe individuals' demographic status in 2000, at the time the initial PPM choice was made. We also have information on the investor's age as of Dec 31 2000, whether the investor is an immigrant, his income, and net wealth,[5], if the investor lives in an urban or rural area or small town, the investor's education level and major, his occupation, sex, and marital status. Excluding individuals too young or too old to make a selection in the PPM system in 2000, the sample includes 15,651 individuals, of whom 10,373 made an active choice.

Preliminary descriptive statistics appear in Table 1, where we note that the average investor holds 2.60 funds (conditional on making an active choice, 3.47 funds). Men hold slightly fewer funds than women; though the difference is statistically significant, the economic magnitude is only about 0.02. Younger investors (under age 40) hold more funds than older persons (respectively 3.27 and 3.71). It is also interesting to note that, while portfolios of older investors are less exposed to equity and less risky, the difference in exposure is small and not economically significant. Older investors hold 80% of their portfolios in equity versus 86% for young investors. Turning to transactions, investors on average rebalance their portfolios 1.64 times over our time period (with the median of 1, a 99th percentile of 8 and a maximum of 155). This frequency is lower for men (women rebalance 1.59 times per year, while the men did so 1.70 times per year) and there is no important age effect.

[5] Net wealth is made up of the market value of domestic and international financial assets such as bank account, fixed income, mutual funds, stocks and options, and an estimated market value of real estate, minus debt.

The Table also shows that the average age of our sample is 43; 89 percent were born in Sweden; average gross income in 2000 was 202,269 SEK, and average net wealth was 444,108 SEK. Some 29 percent lived in the countryside, with the remaining 71 percent living in either one of the three major cities in Sweden (35 percent) or in a town (36 percent).

Table 1. Descriptive Statistics for PPM Investors

	Mean	Std. Dev.
ENTERSIM	0.010	0.010
ENTERAFTER	0.012	0.012
EXITSIM	0.053	0.057
EXITAFTER	0.086	0.079
Choice	0.669	0.221
Equity	0.901	0.170
Number of funds	2.602	1.409
Age	43.028	11.111
Swede	0.887	0.316
Income	202,269	188,121
Net Wealth	444,108	3,174,320
Urban	0.354	0.478
Town	0.356	0.479
Rural	0.289	0.454
Education: less then High School	0.181	0.385
Education: High School	0.513	0.500
Education: College	0.306	0.461
Major: Social Science	0.262	0.440
Major: Engineering	0.239	0.426
Major: Medical	0.134	0.341
Major: Other	0.365	0.482
Occ: Public Sector	0.287	0.452
Occ: Private Sector	0.558	0.500
Occ: Self-Employed	0.048	0.213
Occ: Other	0.108	0.310
Male	0.509	0.500
Married	0.778	0.415

Notes: N=15,497. Key outcomes are ENTERSIM if between July-December, 2000; ENTERAFTER takes the value 1 if a person entered the stock market between January-June, 2001; both are set to 0 otherwise. EXITSIM (EXITAFTER) equals 1 if person exited the stock market between July 1-December, 2000 (January-June, 2001) and 0 otherwise. CHOICE is set to 1 if investor made an active choice in the PPM system, 0 otherwise; Equity represents the amount invested in equity divided by the total amount held in the PPM. Number of funds represents the number of funds included in the participant's PPM portfolio. Age refers to the participant's age as of Dec 31 2000; Swede is set to 1 if the individual was born in Sweden, 0 else; INCOME is the individual's net income in 2000; NET WEALTH is the individual's net wealth which is the market value of financial wealth + real estate – debt; Urban, Town and Rural are set to 1 respectively if the individual lived in one of the three major cities in Sweden, any smaller town or on the county side, correspondingly; education level is split into three dummies set to 1 if the person's schooling was less than, equal to, or more than high school, correspondingly; Education major is split into four dummies set to 1 if person' higher education major was social science, engineering, medical, or

other; occupation is split into four dummies set to 1 if the person was employed in the public sector, private sector, self employed, or unknown employment/unemployed; Male is set to 1 if the person was male; and Married is set to 1 if the person was married.

Few, 18 percent, had not completed high school; 51 percent had completed high school; and 31 percent went to college. Around one quarter majored in social science (including economics, law, sociology

4. Investment Patterns in the PPM System

In what follows, we first analyze how investors choose the funds in their retirement account. Some investors choose specific combinations of equity and bond funds. We call them active investors. Others do not make a deliberate choice but instead opt for the default government-run equity fund. Next we focus on stock market participation, to test the link between the choice in the PPM system and stock market participation. We show that for the case of the active investors, being active in their retirement choice increases the probability of stock market participation. That is, an individual who did not participate in the stock market is more likely to enter it once he has been presented with the new pension scheme. This effect is not only statistically significant, but also economically relevant.

4.1 Choice of Funds

The key question of interest is whether PPM participation affects stock market participation. We first approach this by providing descriptive statistics about how investors select their funds in the PPM system. Some investors make a deliberate choice and choose specific combinations of equity and bond funds, while others just opt for the default. This allows us to define active and passive investors.

Table 2 describes entry and exit rates, defined as ratio of people who entered or exited the stock market between two different periods. Our focus is on two time periods, simultaneous with the rollout of the PPM system (7/1/00 to 12/31/00), and after the PPM system was in place (1/1/01 to 6/30/01). Panel A compares people who actively choose funds in the PPM system to those who did not, whereas More

than 66% of the households made a deliberate decision. These households were younger, in general.

Table 2: Entry and Exit Rates as Function of PPM Investment Choice

This table reports entry and exit rates for subgroups of people who actively selected funds in the PPM system, and the ones who did not (Panel A); and subgroups of people who choose different number of funds (Panel B).

Panel A: Entry and Exit Rates as Function of Active Choice

	ENTERSIM		ENTERAFTER		EXITSIM		EXITAFTER	
Choice	N	Mean	N	Mean	N	Mean	N	Mean
0 (No choice)	4792	0.0065	4783	0.0084	332	0.0663	341	0.1114
1 (Active Choice)	9378	0.0125	9314	0.0135	995	0.0533	1059	0.0784
Test of the difference								
t-statistic		-3.3285		-2.6921		0.8877		1.8907
p-value		0.0009		0.0071		0.3748		0.0589
Wilcoxon' Z		-3.3266		-2.6919		0.8881		1.8890
p-value		0.0009		0.0071		0.3747		0.0589

Panel B: Entry and Exit Rates as Function of Number of Funds Chosen

	ENTERSIM		ENTERAFTER		EXITSIM		EXITAFTER	
Number of funds	N	Mean	N	Mean	N	Mean	N	Mean
1	6140	0.0085	6119	0.0083	490	0.0633	511	0.0978
2	1272	0.0071	1266	0.0150	120	0.0250	126	0.0556
3	1997	0.0080	1993	0.0085	212	0.0566	216	0.1157
4	1872	0.0107	1865	0.0123	173	0.0751	180	0.0611
5	2889	0.0177	2854	0.0196	332	0.0482	367	0.0763
Test of the difference between groups 1 and 5								
t-statistic		-3.8360		-4.5921		0.9126		1.1067
p-value		0.0001		0.0000		0.3617		0.2687
Wilcoxon' Z		-3.8330		-4.5870		0.9130		1.1070
p-value		0.0001		0.0000		0.3614		0.2685
Test of the difference between groups 1 and combined 2,3,4 and 5								
t-statistic		-2.0229		-3.3178		-0.8139		-1.1525
p-value		0.0431		0.0009		0.4158		0.2493
Wilcoxon' Z		-2.0229		-3.3170		-0.8140		-1.1530
p-value		0.0431		0.0009		0.4156		0.2492

If we concentrate on ENTERSIM, one can see that entry rate is almost two times larger (1.25% vs. 0.65%) among the people who made active choice. This difference is also statistically significant (p-value of both t-test and Wilcoxon test is 0.0009). Similar results were obtained for period after introduction of PPM system (January-June, 2001). People who made active choice 61% more likely to

enter stock market during this period (1.35% vs. 0.84%, p-value of mean and median test 0.0071). At the same time, active participation in PPM system does not affect strongly exit rate. Panel B divides the sample into subgroups of investors who choose different numbers of funds. As in Panel A, people who choose multiple funds are more likely than people who choose only one fund to enter the stock market and same chances to exit.

4.2 Pension Choice and Stock Market Participation

We next focus on the link between PPM system choice and stock market participation. We focus both on investors who did participate in the stock market and look at their exit rates, and also those who did not participate initially and evaluate their later entry rates. These we model as a function of investors' demographic characteristics and choices made in the PPM system. The sample is affected by selection bias inasmuch as people with certain demographic characteristics are more likely to invest in the stock market. To deal with this issue, we estimate a two-stage Heckman-type Probit selection model for the decision to enter the stock market (c.f. Van de Ven and Van Praag, 1981). That is, we first estimate the probability that the investor holds equity already at time *t-1*, and then, conditional on that behavior, we relate it to the decision to enter the stock market. For example, for investor *j* with characteristics x_j, z_j we estimate the following equation: $Entry_j=(x_j\beta+u_{1j}>0)$. The selection equation is $z_j\gamma+u_{2j}>0$, where $u_1 \sim N(0,1)$, $u_2 \sim N(0,1)$ and corr(u_1, u_2)=□. Choice in the pension scheme is posited to depend on a number of investor demographic characteristics, income, wealth, residential location, and education. [6]

[6] Specifically we control on investors' Age in 2000; Swedish nationality; Net Income (in 2000 measured as the logarithm of income in SEK plus one); Net Wealth defined as the market value of financial wealth and real estate less the value of debt (measured as the logarithm of Net

The dependent variables of most interest are indicators of individuals exiting or entering the stock market during this time period, and a continuous variable for how much the individual investor has increased or decreased her investments. In particular, ENTERSIM takes the value of 1 if the investor entered the stock market simultaneous with the PPM launch (July 1-December 31, 2000), while ENTERAFTER takes the value 1 if the person entered the stock market after the launch (January 1-June 30, 2001). Similarly, EXITSIM (EXITAFTER) takes the value 1 if the person exited the stock market between July 1-December 31, 2000 (or January 1-June 30, 2001).

The behavioral variables on which we focus most are Choice, Equity, and Number of Funds. We refer to these as proxies for activity in the PPM system. Choice takes on a value of 1 if the investor made an active choice in the PPM system, and equals zero otherwise. Equity represents the fraction invested in equity out of the total amount invested in PPM. Number of funds refers to the number of funds included in the investor's PPM portfolio. Also included are other demographic factors defined as above.

Table 3 illustrates one specification for the period during the introduction of PPM system (July-December, 2000) and a second after the introduction of the PPM system (January-June 2001). We start by considering the decision to enter the

---

Wealth plus one); Urban, which takes the value of 1 if the investor lived in one of the three major cities in Sweden; Town takes the value of 1 if the investor lived in a smaller town; and Rural equals 1 if the investor lived in the countyside. We also include three variables which measure educational levels, respectively set to 1 if the person's schooling was less than, equal to, or exceeded high school. We control for participants' education major by using four variables which take on a value of 1 if the investor's higher education major was respectively, social science, engineering, medical school, or other. Analogously, we control for occupational differences using four variables set to 1, respectively, if the person was employed in the public sector, private sector, self employed, or had unknown employment or was unemployed. Finally, we include Male and Married controls, respectively set to 1 if the investor was a male and 1 if the investor was married. To achieve identification, In Tables 3 and 4, we used Age, Age Squared and set of dummies related to education major in selection equation and not in the second stage.

stock market, and we provide results for models with and without a Past Stock Market Participation Dummy (set equal to 1 if investor held equity in 1998:12, 1999:06 or 1999:12). The results confirm that those who made a deliberate choice under the PPM system have a higher probability of stock market participation. In particular, there is always a positive relationship between our three proxies for activity in the PPM system, namely Choice, Equity and Number of Funds, and the ensuing decision to enter the stock market. This result is robust, holding for all specifications and cases presented. Not only is it statistically significant, but it is also economically relevant: thus making a deliberate choice under the PPM increases the probability of stock market participation by 0.28% (see bold value in row 1, column 5). This is equivalent to a 24% increase with respect to the unconditional mean. Analogously, a one standard deviation rise in the number of funds that the investor selects for his retirement portfolio raises the probability of stock market entry by 0.12% (bold value in row 2, column 11). This represents a 10% increase with respect to the unconditional mean.

Finally, an increase in the fraction invested in equity funds by one standard deviation increases the probability of stock market entry by 0.4%. This represents a 34% increase with respect to the unconditional mean. These results hold for both after the introduction of PPM system (January-June 2001) and also during the PPM introduction (July-December, 2000; and point estimates in this earlier period are somewhat larger).

Table 3: Effect of PPM Choice on Entry Decision.
This table reports the results of two-stage Heckman-type Probit selection models for the decision to enter the stock market during the introduction of PPM system (July-December 2000, reported in Panel A) versus after the introduction of PPM system (January-June 2001, reported in Panel B). We report the result for two specifications, with and without Past Participation Dummy (which is equal to 1 if the person held equity in our sample in 1998:12, 1999:06 or 1999:12. Other control variables are described in Table 1. Marginal effects (ME) are given for the second stage; t-statistics are reported in parentheses.

Panel A: Decision to Enter the Stock Market: July-December 2000

	Decision to participate in the stock market (ENTERSIM)																	
	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME
Choice	0.212	(4.25)	0.0038	0.201	(3.84)	0.0039												
Number of funds							0.056	(5.60)	0.0011	0.053	(4.80)	0.0011						
Equity													0.286	(4.98)	0.0054	0.279	(4.67)	0.0055
Past Participation Dummy	1.083	(10.08)	0.0763				1.100	(9.96)	0.0773				1.094	(10.16)	0.0755			
Pseudo R2	0.094			0.072			0.094			0.072			0.097			0.076		

Panel B: Decision to Enter the Stock Market: January-July 2001

	Decision to participate in the stock market (ENTERAFTER)																	
	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME
Choice	0.127	(2.51)	0.0028	0.1253	(2.79)	0.0030												
Number of funds							0.0570	(3.81)	0.0012	0.0533	(3.99)	0.0012						
Equity													0.1810	(2.39)	0.0039	0.1732	(2.55)	0.0040
Past Participation Dummy	0.604	(4.36)	0.0327				0.6786	(4.63)	0.0339				0.6494	(4.70)	0.0332			
Pseudo R2	0.103			0.096			0.107			0.100			0.104			0.098		

Note: All controls from Table 1 and an intercept are also included in the model.

We now consider the decision to exit the market, with estimates appearing in Table 4. Here, results indicate that making a deliberate choice under the PPM system is either irrelevant to the investor's exit decision, or it reduces the probability of leaving the stock market (after the introduction of PPM system). The results for the time period of July-December, 2000 (Panel A) show that our three proxies for activity in the PPM system do not affect the exit decision at this point. In Panel B, after the introduction of PPM (January-June 2001), there is always a negative relationship between our three proxies for activity in the PPM system and the probability of leaving the stock market. This result holds across all the different specifications. The results are strongly statistically significant, but the economic significance is modest. In particular, making a deliberate choice in the retirement account reduces the probability of leaving the stock market by 0.22%, or 2.8% of unconditional mean. Analogously, an increase in the number of funds that the investor selects for his retirement portfolio by one standard deviation decreases the probability of the probability of leaving the stock market by 0.07% (about 0.8% of unconditional mean). Finally, an increase in the fraction invested in equity funds by one standard deviation reduces the probability of leaving the stock market by 0.31%. This represents a 3.9% decrease with respect to the unconditional mean. [7]

Overall, then, the findings suggest a sort of "learning effect." That is, investors afforded the choice to invest their retirement money in equity funds and then make a deliberate choice to do so, are more likely to enter the stock market.

[7] It is interesting to note that, for most of our estimates, the correlation coefficients of the Probit and selection equation error terms are not statistically significant (i.e., in our case, selection does not affect the Probit results).This allows us to ignore the issues of selectivity and to concentrate on the issues related to the endogeneity of the choice.

Table 4: Effect of PPM Choice on Exit Decision

This table reports the results of two-stage Heckman-type probit selection models for the decision to exit the stock market during the introduction of PPM system (July-December,2000, reported in Panel A) and after the introduction of PPM system (January-June 2001, reported in Panel B). We report the results for two specifications, with and without Past Participation Dummy (which is equal to 1 if the person held equity in our sample in 1998:12, 1999:06 and 1999:12. Other control variables are described in Table 1. We report marginal effects (ME); t-statistics are reported in parentheses.

Panel A: Decision to Exit the Stock Market: July-December 2000

	Decision to participate in the stock market (EXITSIM)																	
	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME
Choice	-0.030	(-0.40)	-0.0002	-0.052	(-0.55)	-0.0005												
Number of funds							-0.009	(-0.23)	0.0000	-0.004	(-0.09)	0.0000						
Equity													0.113	(1.21)	0.0007911	0.104	(1.43)	0.0011
Past Participation Dummy	0.427	(6.24)	0.0033				0.427	(5.97)	0.0033				0.400	(5.48)	0.0039884			
Pseudo R2	0.086			0.053			0.085			0.052			0.086			0.053		

Panel B: Decision to Exit the Stock Market: January-July 2001

	Decision to participate in the stock market (EXITAFTER)																	
	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME	Est.	t-stat	ME
Choice	-0.215	(-3.19)	-0.0022	-0.215	(-3.18)	-0.0021												
Number of funds							-0.068	(-2.05)	-0.0007	-0.069	(-2.09)	-0.0006						
Equity													-0.315	(-3.74)	-0.0031	-0.315	(-3.79)	-0.0030
Past Participation Dummy	-0.075	(-0.85)	-0.0007				-0.064	(-0.69)	-0.0006				-0.075	(-0.84)	-0.0007			
Pseudo R2	0.047			0.046			0.048			0.047			0.052			0.051		

Note: All controls from Table 1 and an intercept are also included in the model.

4.3 Pension Choice and Stock Market Participation: Controlling for Endogeneity.

Next we consider the issue of endogeneity which arises because the investor determines the value for his Choice, Equity and Number of Funds variables jointly with the decision to participate in the stock market. To correct for this, we adopt an Instrumental Variables (IV) methodology. Specifcially we estimate both a 2-Stage Least Squares (2SLS) linear probability model and an IV-Probit model of the decision to enter the stock market, as a function of our three proxies for activity in the PPM system as well as the other controls. We use as instruments Age and Age squared.

Our estimates appear in Tables 5 and 6. The explanatory variable in Panel A is the active choice decision under the PPM system; in Panel B, it is the number of funds chosen; and in Panel C, it is the fraction of equity in the chosen portfolio. As in the previous case, we estimate a specification for the period during the introduction of PPM system (July-December, 2000) and after the introduction of the PPM system (January-June 2001).

As before, there is a positive relationship between the three proxies for PPM activity and the ensuing decision to enter the stock market. Furthermore, this result holds across all specifications and is economically significant. For example, an increase in the probability of making a deliberate choice in the retirement account triples the unconditional probability of stock market entry. Marginal effect related to making active choice is about 3%, which is three times unconditional probability of entry. Similar numbers for the effect of choosing multiple funds and choosing larger fraction of equity are 35% and 170% increase with respect to unconditional probability of entry, correspondingly.

Table 5: Effect of PPM Choice on Entry Decision: IV Estimates.

This table reports the result of 2SLS linear probability model and IV-Probit estimates of a decision to enter the stock market as a function of active choice decision (Panel A), number of funds chosen (Panel B) and fraction of equity in the chosen portfolio (Panel C). We used as instruments Age and squared Age of the investor. We report the results of first stage estimate and joint F-test of significance of the instruments. For 2SLS linear probability model we report Hansen overidentifying restrictions test. For IV-probit estimates we report Wald test of exogeneity. The control variables are as in Tables 1 and 3. We report marginal effects (ME) for probit estimates. t-statistics are reported in parentheses.

Panel A: Decision to Enter Stock Market and Make Active Choice in PPM System

	Entry June-Dec 2001 (ENTERSIM)							Entry Jan-June 2001 (ENTERAFTER)						
	First Stage		2SLS		Probit			First Stage		2SLS		Probit		
	Est.	t-stat	Est.	t-stat	Est.	t-stat	ME	Est.	t-stat	Est.	t-stat	Est.	t-stat	ME
Choice			0.0337	(1.91)	1.1501	(4.09)	0.0316			0.0576	(1.96)	1.6570	(7.24)	0.1149
Past Participation Dummy	0.0333	(0.71)	0.1256	(11.66)	0.9555	(4.89)	0.0878	0.0413	(0.58)	0.0669	(5.32)	0.4576	(2.41)	0.0577
	F-test of instruments		Hansen OIR Test		Wald Exogeneity Test			F-test of instruments		Hansen OIR Test		Wald Exogeneity Test		
	F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val		F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val	
	31.52	0.000	1.528	0.2164	8.300	0.004		29.78	0.000	0.161	0.688	22.020	0.000	

Panel B: Decision to Enter the Stock Market and Number of Funds Chosen in PPM System

	Entry June-Dec 2001 (ENTERSIM)							Entry Jan-June 2001 (ENTERAFTER)						
	First Stage		2SLS		Probit			First Stage		2SLS		Probit		
	Est.	t-stat	Est.	t-stat	Est.	t-stat	ME	Est.	t-stat	Est.	t-stat	Est.	t-stat	ME
Number of funds			0.0048	(1.77)	0.1806	(3.49)	0.0038			0.0066	(1.94)	0.2590	(3.49)	0.0062
Past Participation Dummy	0.0363	(0.31)	0.1265	(11.89)	1.0777	(7.76)	0.0795	0.0893	(0.59)	0.0687	(5.62)	0.6906	(2.27)	0.0373
	F-test of instruments		Hansen OIR Test		Wald Exogeneity Test			F-test of instruments		Hansen OIR Test		Wald Exogeneity Test		
	F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val		F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val	
	213.78	0.000	0.711	0.399	6.410	0.011		211.75	0.000	0.015	0.9033	6.470	0.011	

Panel C: Decision to Enter the Stock Market and Fraction of Equity Chosen in PPM System

	Entry June-Dec 2001 (ENTERSIM)							Entry Jan-June 2001 (ENTERAFTER)						
	First Stage		2SLS		Probit			First Stage		2SLS		Probit		
	Est.	t-stat	Est.	t-stat	Est.	t-stat	ME	Est.	t-stat	Est.	t-stat	Est.	t-stat	ME
Equity			0.0211	(1.79)	0.8268	(4.11)	0.0179			0.0270	(1.99)	1.0972	(4.35)	0.0317
Past Participation Dummy	0.0393	(0.88)	0.1259	(11.80)	1.0442	(7.09)	0.0765	0.0429	(0.64)	0.0680	(2.10)	0.6407	(2.32)	0.0382
	F-test of instruments		Hansen OIR Test		Wald Exogeneity Test			F-test of instruments		Hansen OIR Test		Wald Exogeneity Test		
	F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val		F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val	
	155.78	0.000	0.503	0.478	8.230	0.004		152.56	0.000	0.122	0.7253	13.270	0.000	

Note: All controls from Table 1 and an intercept are also included in the model.

Table 6: Effect of PPM Choice on Exit Decision: IV estimates.

This table reports the result of 2SLS linear probability model and IV-probit estimates of a decision to exit the stock market as a function of active choice decision (Panel A), number of funds chosen (Panel B) and fraction of equity in the chosen portfolio (Panel C). We used as instruments Age and squared Age of the investor. We report the results of first stage estimate and joint F-test of significance of the instruments. For 2SLS linear probability model we report Hansen overidentifying restrictions test. For IV-probit estimates we report Wald test of exogeneity. The control variables are as in Tables 1 and 3. We report marginal effects (ME) for Probit estimates and t-statistics are reported in parentheses.

Panel A: Decision to exit the stock market and active choice in PPM system

	Exit June-Dec 2000 (EXITSIM)							Exit Jan-June 2001 (EXITAFTER)						
	First Stage		2SLS		Probit			First Stage		2SLS		Probit		
	Est.	t-stat	Est.	t-stat	Est.	t-stat	ME	Est.	t-stat	Est.	t-stat	Est.	t-stat	ME
Choice			0.4046	(1.31)	2.2024	(2.88)	0.4726			-0.0835	(-0.29)	-0.7281	(-0.59)	-0.1334
Past Participation Dummy	-0.0061	(-0.91)	0.0525	(3.15)	0.1701	(2.76)	0.0578	-0.0082	(-1.22)	-0.0114	(-0.73)	-0.0719	(-0.83)	-0.0101
	F-test of instruments		Hansen OIR Test		Wald Exogeneity Test			F-test of instruments		Hansen OIR Test		Wald Exogeneity Test		
	F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val		F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val	
	2.28	0.103	0.981	0.322	36.72	0.000		2.33	0.0977	2.96	0.085	0.170	0.169	

Panel B: Decision to Exit the Stock Market and Number of Funds Chosen in PPM System

	Exit June-Dec 2000 (EXITSIM)							Exit Jan-June 2001 (EXITAFTER)						
	First Stage		2SLS		Probit			First Stage		2SLS		Probit		
	Est.	t-stat	Est.	t-stat	Est.	t-stat	ME	Est.	t-stat	Est.	t-stat	Est.	t-stat	ME
Number of funds			0.0396	(1.96)	0.3858	(3.02)	0.0607			-0.0243	(-1.05)	-0.1462	(-1.98)	-0.0199
Past Participation Dummy	0.0542	(1.47)	0.0465	(3.30)	0.3368	(3.56)	0.0564	0.0590	(1.74)	-0.0062	(-0.39)	-0.0463	(-0.51)	-0.0063
	F-test of instruments		Hansen OIR Test		Wald Exogeneity Test			F-test of instruments		Hansen OIR Test		Wald Exogeneity Test		
	F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val		F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val	
	26.660	0.000	0.154	0.695	7.800	0.005		25.480	0.000	1.911	1.669	5.320	0.021	

Panel C: Decision to Exit the Stock Market and Fraction of Equity Chosen in PPM System

	Exit June-Dec 2000 (EXITSIM)							Exit Jan-June 2001 (EXITAFTER)						
	First Stage		2SLS		Probit			First Stage		2SLS		Probit		
	Est.	t-stat	Est.	t-stat	Est.	t-stat	ME	Est.	t-stat	Est.	t-stat	Est.	t-stat	ME
Equity			0.2415	(1.77)	1.9007	(5.46)	0.4385			-0.1114	(-0.76)	-0.6713	(-1.74)	-0.0915
Past Participation Dummy	-0.0107	(-1.29)	0.0517	(3.57)	0.2976	(3.57)	0.0715	-0.0131	(-1.90)	-0.0109	(-0.70)	-0.0687	(-0.76)	-0.0093
	F-test of instruments		Hansen OIR Test		Wald Exogeneity Test			F-test of instruments		Hansen OIR Test		Wald Exogeneity Test		
	F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val		F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val	
	9.350	0.000	0.431	0.511	18.350	0.000		9.380	0.000	1.900	0.168	1.100	0.294	

Note: All controls from Table 1 and an intercept are also included in the model.

These results hold for both for the period during the introduction of PPM system (July-December, 2000) and afterwards as well (January-June 2001). [8]

We now consider the decision to exit the market. When endogeniety is properly controlled, it is interesting to note that the results for the period July-December 2000 now show a strong positive impact of our proxies for activity in PPM on the decision to exit the stock market. It is consistent with some participants in the stock market (probably, bubble era entries) being induced by the PPM publicity campaigns to review stock market risks (or perhaps they simply looked at their holdings which dropped 20% in value during that period) and decided to leave. The decision to leave the stock market after the introduction of the PPM system proves to be unrelated to the PPM behavior. Results in Table 6 are weaker than those reported in Table 4; further the 2SLS estimates lose significance, as do the estimates based on the Choice variable (Panel A). The estimates based on the Number of Funds and the Fraction of Equity are, however, still negative and significant.

Overall, these findings again support a "learning effect." In other words, exposure to decisionmaking about risky choices educates people. Those who otherwise would not have made risky decisions are thereby induced to participate in the stock market.

4.4 A Counterfactual Test

Next we perform a counterfactual experiment to test whether our main explanatory variables -- Choice, Number of Funds, and Equity – might proxy for some individual-specific characteristics that are not necessarily related to the

[8] It is worth noticing that the tests of the goodness of instruments show indeed that our instruments do a good job. We report the first-stage estimates and a joint F-test of significance of the instruments; they are jointly significant at the first stage. For the 2SLS linear probability model we report Hansen overidentifying restrictions test. For IV-probit estimates we report Wald test of exogeneity. For 2SLS estimates, the test of overidentified restrictions fails to reject the model.

introduction of the PPM system. That is, it might be that people who make a deliberate selection of pension assets, or people that invest most of their contributions in equity, are also the ones who would in any case participate more in the stock market, regardless of whether PPM is in force.

To address this issue, we re-estimate the same model as before, but now over a different period. Specifically we select a time prior to the introduction of the PPM, so entry and exit decisions now refer to December 1999-June 2000. The goal is to use the same model as before. We hypothesize that if the decision to participate in the stock market is related to the introduction of the new pension scheme, as opposed to individual-specific characteristics, we would anticipate that the very same variables which explain stock market participation later on would also explain it beforehand. Results in Table 7 support our hypothesis. That is, we find no relationship between our main explanatory variables and the decision to enter or exit the stock market. For the control variables, instead, the same relationship holds. This suggests that it was the introduction of the new Swedish retirement scheme that affected stock market behavior, as opposed to participants' characteristics. The findings also provide strong evidence of causality from PPM choice to subsequent investment behaviors.

Table 7: Placebo Test: Effect of PPM Choice on Entry Decision before PPM System Introduction: IV Estimates.
This table reports the result of 2SLS linear probability model and IV-Probit estimates of a decision to enter (Panel A) and exit the stock market the stock market as a function of PPM choice. The entry and exit referred to the period from December 1999 to June 2000 (prior to PPM introduction). We used as instruments Age and squared Age of the investor. We report the results of first stage estimate and joint F-test of significance of the instruments. For 2SLS linear probability model we report Hansen overidentifying restrictions test. For IV-probit estimates we report Wald test of exogeneity. Past Participation Dummy is defined as equal to 1 if the person held equity in our sample in 1998:12 or 1999:06.The control variables are as in Tables 1 and 3. t-statistics are reported in parentheses.

Panel A: Decision to Enter Stock Market from December 1999-June 2000

	First Stage		2SLS		Probit		First Stage		2SLS		Probit		First Stage		2SLS		Probit	
	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
Choice			0.0091	(0.20)	0.1060	(0.12)												
Number of funds									0.0028	(0.56)	0.0448	(0.65)						
Equity Past Participation															0.0139	(0.65)	0.2591	(0.86)
Dummy	0.0509	(0.88)	0.3610	(7.78)	1.4901	(16.40)	0.0916	(0.36)	0.3612	(7.74)	1.4903	(18.30)	0.0485	(0.87)	0.3608	(7.73)	1.4835	(18.94)
	F-test of instruments		Hansen OIR Test		Wald Exogeneity Test		F-test of instruments		Hansen OIR Test		Wald Exogeneity Test		F-test of instruments		Hansen OIR Test		Wald Exogeneity Test	
	F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val	F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val	F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val
	30.310	0.000	1.506	0.220	0.000	0.988	219.110	0.00	1.311	0.252	0.090	0.762	153.720	0.000	1.246	0.264	0.140	0.711

Panel B: Decision to Exit the Stock Market from December 1999-June 2000

	First Stage		2SLS		Probit		First Stage		2SLS		Probit		First Stage		2SLS		Probit	
	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat	Est.	t-stat
Choice			0.0949	(1.29)	0.0983	(1.28)												
Number of funds									0.0201	(1.07)	0.3612	(1.87)						
Equity Past Participation															0.1166	(1.10)	1.6020	(1.14)
Dummy	0.0193	(0.49)	0.0298	(2.56)	0.3442	(4.73)	0.2074	(1.20)	-0.5894	(-27.96)	-2.1812	(-3.51)	-0.0016	(-0.03)	-0.5868	(-27.02)	0.3687	(1.27)
	F-test of instruments		Hansen OIR Test		Wald Exogeneity Test		F-test of instruments		Hansen OIR Test		Wald Exogeneity Test		F-test of instruments		Hansen OIR Test		Wald Exogeneity Test	
	F-stat.	p-val	Hans J	p-val	$\chi 2$	p-val	F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val	F-stat.	p-val	Hansen J	p-val	$\chi 2$	p-val
	4.750	0.009	0.553	0.457	2.470	0.116	19.040	0.009	0.626	0.429	2.250	0.134	9.150	0.000	0.478	0.489	5.940	0.015

Note: All controls from Table 1 and an intercept are also included in the model.

5. Discussion and Conclusion

This chapter has focused on how the introduction of a defined contribution retirement system affects investors' propensity to participate directly in the stock market. Our unique evidence on investor patterns before and after the Swedish PPM was introduced permits us to focus on the decision to invest in stocks. Results show how this pattern changes, once investors were permitted to participate in the new pension system. Specifically, we showed that introducing the chance to invest in retirement funds increased peoples' tendency to enter the stock market. We also show that investors who made a deliberate choice of their pension asset allocation also boosted stock market participation once the plan was in place. Investors who previously did not participate in the stock market turn out to have a higher likelihood of entering once they are presented with the new pension scheme.

What this means, we argue, is that requiring workers to invest in mutual funds can act as a triggering device which induces them to enter the stock market as well; being induced to choose among different pension funds appears to "educate" participants about the stock market. The fact that investors do not treat their retirement account investments as close substitutes for direct equity investment also implies that the adoption of a capitalization-based system does not necessarily crowd out direct investment. This may be of interest to policymakers in the current debate about moving from a pay-as-you-go to a fully funded defined contribution system.

Chapter 6. Benefits of Contribution: Individual Asset Allocation and Performance in a Defined Contribution Pension System

Anders Karlsson and Lars Nordén

1. Introduction

Currently, there is a trend for countries to move away from defined benefit pension systems towards partially defined contribution plans. There are several reasons for this trend, the perhaps most important being to finance the worldwide phenomenon of aging populations. The key issue in moving to a defined contribution plan is to make individuals not only more conscious of their own pension schemes, but also to let them bear investment risks previously borne by governments or employers (Bodie and Crane, 1998). A growing interest in how non-professional investors handle this new responsibility is emerging in the financial research. The recent shift in the pension system in Sweden, from defined benefit to defined contribution, enables an evaluation of individual asset allocation and portfolio performance of a large diversified sample of non-professional investors.

In order to clarify which asset classes are preferred, and then measure individual performance, the asset classes available in the Swedish pension system (PPM) must first be identified. Therefore, the analysis is carried out in three steps. In the first step, a factor analysis is performed in order to explore the latent factors, or actual asset classes, that are driving the returns of the mutual funds. The findings show that although there appears to be a wide range of choices available to individuals, the large amount of mutual funds can be represented with only a few orthogonal factors or distinct asset classes. By using the ten most important factors, more than 90

percent of the total variance of returns for the original set of 465 mutual funds can be accounted for. In other words, allowing for roughly ten percent noise, about ten orthogonal asset classes are available for investors to choose among in the initial round of investment in the defined contribution part of the Swedish pension system.

The factors are identified in terms of real world asset classes or indices. The by far most important factor is easily identified as an overall, world market index. Moreover, among the others factors are found covering equity from Japan, the Far East, and different types of fixed income securities. Apparently, there are clearly a lot of redundant mutual funds present in the initial choice set. Also, a choice among say ten orthogonal factors, or distinctly different indices, would facilitate an easier and more efficient individual asset allocation than the actual choice between 465 different, or sometimes not so different, mutual funds.

After having identified the factors within the Swedish defined pension contribution system the analysis is turned to individuals' asset allocation. Using a sample of individuals taking part in the defined contribution pension portfolio formation in the year 2000, the individuals' loadings and communalities with respect to the different factors are analyzed. Moreover, the individuals' factor exposures are also related to a number of demographic and socio-economic variables in order to find out whether asset allocation differs with respect to individual characteristics. The results show that more sophisticated individuals have a higher probability of making an active choice, and the result of this activity is to reduce the loadings on the overall market factor and domestic Swedish bond factors.

Finally, the performance of the individuals' portfolios is investigated over the first four years since the introduction of the new Swedish pension system. Jensen's alpha from a regression of an individual's monthly excess return on excess returns on a set of market indices, related to the orthogonal factors, is used to measure performance.

The results indicate that sophisticated investors are more likely to choose worse performing funds than non-sophisticated investors. Also, men and older investors are more likely to pick well performing funds than women and younger investors.

The contributions to previous research of this study are several. First, there are very few studies on the investment opportunities available for individuals in defined contribution plans (see Blake et al., 2004). The analysis of the investment opportunities of the Swedish defined contribution plan identifies 13 core assets classes among the available 464 mutual funds. Second, diversities in asset allocation are investigated with respect to individuals' characteristics. The factor analysis of the offered set of mutual funds makes it possible to investigate individuals' choices of orthogonal asset classes, i.e. actual diversification, rather than "naïve" diversification of simply investing in several, possibly highly correlated, mutual funds.[1] Third, the performance of the individuals' portfolios within the pension plan is evaluated. Here the analysis of Blake et al. (2005) is extended to an individual level. Again, individual performance is analyzed in detail, highlighting differences with respect to individual demographic and socio-economic characteristics. Therefore, it is possible to identify groups of superior performance relative other groups, who benefit from the shift from the old defined benefit to the new defined contribution pension system. Hence, the results have several policy implications, both on an individual investor level, e.g. for individual pensioners in terms of asset allocation and performance, and on a larger economy-wide scale, for policy makers dealing with the construction of pension schemes.

The rest of the study is organized into five sections. The following section briefly presents the Swedish pension system, with emphasis on individual choice in the defined contribution part. Section 3 outlines the factor analysis framework for

[1] See Benartzi and Thaler (2001).

extracting latent factors from the mutual funds available to the individuals. In section 4, we relate the factor analysis to the individuals' asset allocation and their factor loadings, whereas in section 5 we evaluate the performance of the individuals' portfolios. The study ends in section 6 with some concluding remarks.

2. The Swedish pensions system: a mixture of defined benefit and contribution

The new pension system was introduced in Sweden in the autumn of 2000 and consists of three parts. The first and largest part is the income pension, which is based on 16 percent of the annual income and is used to finance those who are retired today. The amount paid in also serves as a base in calculating future pension payments. The second part, the premium pension, is based on 2.5 percent of the annual income. In the first round in 2000, 2.5 percent of the previous four years of income was invested. This amount was allocated at each individual's discretion.

Each individual was presented with an investment opportunity set of 464 funds[2] and invited to choose between one and five funds.[3] If no choice was made, the allotted money was invested in the Seventh Swedish Pension Fund run by the government. This default alternative is an equity fund and cannot be chosen once the investor has made an active choice. The resulting investment portfolio can be altered as often as the individual investor wishes. The accrued amount will be paid out on a monthly basis to the individual at the time of her or his retirement. The third part of the system is a guaranteed pension level designed to ensure that no retiree will be completely without pension payments at the time of her or his retirement, regardless of her or his previous income. In total, 18.5 percent of the annual income for each individual is invested to finance this system, and all annual income from the age of

[2] 464 funds were available in the 2000 brochure. The 2003 brochure contains more than 600 funds.
[3] The Swedish pension system is described in further detail at www.ppm.nu and www.pension.nu. See also Engström and Westerberg (2003), Karlsson (2005) and Säve-Söderbergh (2003).

16 is included. However, an individual earning more than 7.5 income base amounts[4] per year will only be accredited an upper limit of 7.5 income base amounts, although he/she will still pay 18.5 percent of his/her income to finance the pension system.

During autumn 2000 all participants in the Swedish pension system were provided with a brochure containing 464 mutual funds with accompanying information on risk, historical returns, fees, and a few words briefly describing each fund. Table 1 provides an extract from the brochure, with information on one randomly chosen fund available for the investors as an example. Apart from the information exemplified in Table 1, the funds are also categorized at three different levels in the brochure (see Table 3).

Figure 1: Extract from the PPM Investor Information Folder, fund example

Fund number	Fund name, Management company	Information regarding the funds	Fund fee (%)	Percentage return 99-12-31 (after fees)						Total risk
				In the year					Last 5	(last 3
				95	96	97	98	99	years	years)
191080	Baring Global Emerging Markets Baring International Fund Managers (Ireland) Ltd	Emerging markets' equity and equity related assets	1.59	-32	10	25	-25	77	25.3	32 (Red)

The percentage return for the last five years equals the compounded annual growth rate of return for the years 1995 through 1999. The total risk corresponds to an annualised percentage standard deviation of three-year monthly historical fund returns. The total risk is also categorised into five different classes, and colours, with respect to standard deviation; Class 1: very low risk, dark green, percentage standard deviation in the range 0-2; Class 2: low risk, light green, 3-7; Class 3: average risk, yellow, 8-17; Class 4: high risk, orange, 18-24; Class 5: very high risk, red, 25-.

Source: Premiepensionsmyndigheten (2000).

[4] For the year 2000, one income base amount equals SEK 38,800.

3. Asset classes in the current Swedish pension system

In order to evaluate individuals' asset allocation and performance, the investment opportunities available to the individuals at the time of the initiation of the partial defined contribution system in Sweden are first investigated. In the year 2000, individuals could choose among 465 different mutual funds, including the default alternative, with different asset allocation approaches and fund managers. The analysis aims at exploring the latent factors, or actual asset classes, driving the returns, and thus the investment performance, of the mutual funds. A set of latent factors is extracted from the correlation matrix of mutual fund returns. Then, comparing the factor loadings for the different underlying mutual funds identifies the factors.

3.1 Factor analysis: how many factors are covered by the current system?

In order to perform the factor analysis the correlation matrix of the mutual fund returns is estimated. Monthly data on mutual fund price quotes and dividends are collected during the period from December 2000 through December 2004. Monthly log returns, including dividends, are calculated for each mutual fund over the sample period, resulting in 465 return series, each containing 44 monthly observations. During the four-year sample period a number of mutual funds have ceased to exist for different reasons. The main purpose with the factor analysis is to evaluate the number of different asset classes available to the individual investors in 2000, when the defined contribution part of the pension system was initiated, and the initial choices of mutual funds were made. Hence, in order to retain an investor perspective, and to avoid a selection bias in the factor analysis, it is important to keep track of all changes in the initial set of mutual funds over the sample period.

Over the four-year period, 58 mutual funds were terminated, and the invested money was transferred into another fund managed by the same company. In this case, monthly log returns from December 2000 are calculated using the initial mutual fund, and are then simply rolled over to the new mutual fund during the termination month, continuing to calculate log returns. Another 186 funds were terminated, where it was up to each individual investor involved to redistribute the invested money. Here, the investors could choose to invest the money from the terminated fund into any fund available at the termination date, including new funds not available in 2000. If no choice was made, the money was transferred to the default fund, which is a government run equity fund. For simplicity, the return series of terminated funds are set equal to the default fund returns from the termination date. Finally, four funds are excluded from the analysis due to lack of data. As a result, 449 mutual fund returns are used as input into the factor analysis, including the default fund.

A principal component factor analysis is performed using the 449 times 449 correlation matrix of fund returns, where the purpose is to identify the common factors that are responsible for the correlations among the mutual fund returns. The following basic factor model for the mutual fund returns is assumed:

(1) $$R_{i,t} = \alpha_{i,1}F_{1,t} + \alpha_{i,2}F_{2,t} + \ldots + \alpha_{i,m}F_{m,t} + \varepsilon_{i,t}$$

where $R_{i,t}$ denotes the standardized return on mutual fund i in period t, $F_{j,t}$ is the common factor j return, where j = 1, …, m, $\alpha_{i,j}$ is the loading of return i on factor j, and $\varepsilon_{i,t}$ is a factor return, unique to mutual fund i, with mean zero and variance equal to σ_i^2. Since the factor analysis is carried out using a correlation matrix, it is convenient to express the factor model according to equation (1) in terms of standardized returns.

Factor analysis rests on the assumption that the total variance of mutual fund returns can be decomposed into two components; the variance that is common with each factor, the commonality of the fund return with each factor, and the unique fund return variance. From equation (1), the variance of the return on mutual fund i is obtained as:

$$(2) \quad Var(R_{i,t}) = \alpha_{i,1}^2 Var(F_{1,t}) + \alpha_{i,2}^2 Var(F_{2,t}) + \ldots + \alpha_{i,m}^2 Var(F_{m,t}) + Var(\varepsilon_{i,t}) =$$

$$\alpha_{i,1}^2 + \alpha_{i,2}^2 + \ldots + \alpha_{i,m}^2 + \sigma_i^2$$

where the second equality follows from the standardization that the variance of each factor j equals one. Using equation (2), the square of each loading is referred to as the shared variance between the fund and each factor returns, whereas σ_i^2 corresponds to the unique, idiosyncratic fund return variance. That is, the shared variance between a fund and a factor returns is the fund's communality with the factor. The communality is used as a measure of the degree to which the fund is a good and reliable measure of the factor. The sum of the squared loadings equals the total communality, i.e. the part of the fund return variance that is shared with all m factors.

Initially, the principal component factor analysis produces an equal number of latent orthogonal factors, as there are mutual fund return series. However, the aim with the analysis is to reduce the amount of relevant factors, and to keep the m most important ones, namely the factors that can explain a large part of the variation among the returns. Moreover, the rest of the factors are treated as noise, or according to equation (1) as unique factors, not common to all mutual fund returns. Table 2 presents the initial factor solution. Here, m = 23 factors are retained, together responsible for more than 97 percent of the variation among the returns. Each of the 23 factors is associated with an eigenvalue larger than one, i.e. sum of

squared factor loadings, which is the most common rule of thumb used as an aid in selecting the appropriate number of factors.

Table 2: Factor analysis, initial and rotated solution

Factor	Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	307.81	68.555	68.5546	257.62	57.377	57.3766
2	34.359	7.6524	76.2069	31.668	7.0530	64.4296
3	20.602	4.5883	80.7952	30.850	6.8707	71.3003
4	14.000	3.1180	83.9132	24.522	5.4615	76.7617
5	10.611	2.3633	86.2765	23.891	5.3209	82.0826
6	8.6618	1.9291	88.2056	13.109	2.9196	85.0022
7	6.5648	1.4621	89.6677	9.1348	2.0345	87.0367
8	4.8645	1.0834	90.7511	7.6295	1.6992	88.7360
9	3.9579	0.8815	91.6326	5.3101	1.1827	89.9186
10	3.6782	0.8192	92.4518	4.3428	0.9672	90.8858
11	3.2362	0.7208	93.1725	3.7569	0.8367	91.7226
12	2.6210	0.5837	93.7563	3.5350	0.7873	92.5099
13	2.4844	0.5533	94.3096	3.3459	0.7452	93.2551
14	2.3359	0.5203	94.8298	2.7545	0.6135	93.8685
15	2.0904	0.4656	95.2954	2.2293	0.4965	94.3650
16	1.8268	0.4069	95.7023	2.1874	0.4872	94.8522
17	1.6572	0.3691	96.0713	1.9699	0.4387	95.2909
18	1.4795	0.3295	96.4009	1.9499	0.4343	95.7252
19	1.3349	0.2973	96.6982	1.9178	0.4271	96.1524
20	1.2471	0.2777	96.9759	1.7693	0.3941	96.5464
21	1.0987	0.2447	97.2206	1.7507	0.3899	96.9363
22	1.0639	0.2369	97.4576	1.6802	0.3742	97.3105
23	1.0114	0.2252	97.6828	1.6715	0.3723	97.6828

The initial results of the factor analysis are very powerful. With 23 orthogonal factors it is possible to explain more than 97 percent of the variation in the original 449 mutual fund returns. Hence, with more than 97 percent accuracy, it is possible to represent the mutual funds with only 23 factors. As a result, the apparent wide range of choices available to an investor in the defined contribution part of the Swedish pension system can be reduced to a much more narrow choice among only 23 uncorrelated factors or asset classes.

3.2 Identifying the factors: which asset classes are covered?

The results from the factor analysis are useful only if the factors can be identified in terms of the asset classes each factor represents. First, the real economic meanings of the factors are of concern. Indeed, if the factors can be interpreted in terms of actual economic and/or financial variables it lends credibility to the factor analysis, and increases the confidence that economic influences are extracted rather than random noise. Second, the factors must be properly identified to be of use in the subsequent analysis, where first individual factor loadings are investigated in terms of asset allocation, and then, the performance of the individuals' portfolios evaluated.

To facilitate the interpretation of the factors an orthogonal factor rotation is performed using the varimax rotation method. The varimax method is an orthogonal rotation procedure of the initial solution to the factor analysis from Table 2 that minimizes the number of fund returns with high loadings on each factor.[5] Table 2 presents the rotated factor solution. Note that the rotated factor solution consists of the same amount of 23 factors, explaining the same fraction, 97 percent, of the return variation among the mutual funds. However, each individual factor is left with a different fraction of the total explanatory power.

Turning to the actual interpretation of the rotated factor solution, Table 3 presents average total communalities and factor loadings for the mutual funds, divided into the fund categories presented to the individual investors. All average total communalities are very high, with an overall average equal to 97.59 percent. This means that the 23 retained factors can explain more than 97 percent of the return variance for an average mutual fund. The first rotated factor has a reasonably straightforward interpretation as an overall (global) equity market portfolio.

[5] For details, see Sharma (1996).

Table 3: Average communalities and factor loadings for the mutual funds

Category 1	Category 2	Category 3	#Funds	Communality	F1	F2	F3	F4	F5
Equity	Sweden	Sweden (normal)	28	0.9910	0.9449	0.0877	0.1294	-0.1155	0.1155
		Sweden small cap	6	0.9590	0.8294	0.1374	0.2857	-0.1762	0.0498
		Sweden index	7	0.9948	0.9518	0.0701	0.1137	-0.1053	0.1160
	Regional	Swedish equity and foreign equity	11	0.9933	0.9389	0.1677	0.1566	-0.0760	0.1527
		Nordic countries	12	0.9788	0.9044	0.1347	0.1658	-0.0798	0.1164
		Europe	36	0.9829	0.8662	0.1490	0.1590	0.0445	0.1461
		Euroland	8	0.9843	0.8879	0.1362	0.1457	0.0219	0.1292
		Europe small cap	9	0.9751	0.7258	0.2247	0.3757	0.0688	0.0844
		Europe index	7	0.9898	0.8979	0.1274	0.1198	0.0446	0.1413
		North America and USA	26	0.9810	0.8163	0.2158	0.2312	0.0337	0.2201
		Asia and Far East	18	0.9665	0.5860	0.2625	0.6176	0.0650	0.1291
		Global	32	0.9827	0.8311	0.2597	0.2326	0.0335	0.1849
		New markets	21	0.9714	0.6349	0.2712	0.5311	0.0098	0.0246
	Countries	Japan	20	0.9753	0.3086	0.8467	0.1556	0.1186	0.0657
		UK	6	0.9530	0.7041	0.2619	0.3251	0.1634	0.1207
		Other countries	19	0.9497	0.6838	0.2079	0.3135	0.0514	0.1101
	Industry	IT and Communication	19	0.9855	0.8447	0.1212	0.2016	-0.0811	0.0387
		Pharmacutical	7	0.9560	0.5714	0.0978	0.0409	0.0793	0.1594
		Other industries	16	0.9482	0.6645	0.2385	0.2850	0.0651	0.1076
Mixture	Mixture	Swedish equity and fixed income	3	0.9697	0.8985	0.0785	0.0555	-0.1544	0.0277
		Swedish equity, Swedish and foreign fixed income	28	0.9863	0.9318	0.1628	0.1494	-0.0557	0.0653
		Foreign equity and fixed income	22	0.9767	0.6063	0.2314	0.1772	0.4410	0.0690
Generation	Generation	Pension in less than 10 years	5	0.9891	0.9193	0.1969	0.1896	-0.0006	0.1262
		Pension in less than 20 years	6	0.9940	0.9248	0.1983	0.2036	-0.0092	0.1579
		Pension in more than 20 years	21	0.9956	0.9289	0.1909	0.1949	-0.0208	0.1597
Fixed inc.	Fixed inc.	Sweden, short maturity	15	0.9496	-0.1614	-0.0730	-0.1003	0.1469	-0.3477
		Sweden, long maturity	15	0.9867	-0.1702	-0.0527	-0.0311	0.1361	-0.7750
		Europe and Euroland	18	0.9668	-0.1100	0.0433	0.0283	0.7297	-0.2030
		Others	15	0.9678	-0.0432	0.2319	0.1391	0.5068	-0.0440
Default			1	0.9974	0.9355	0.1854	0.1989	-0.0672	0.1103
All funds			457	0.9759	0.6661	0.1802	0.1928	0.0650	0.0499

Table 3: Average communalities and factor loadings for the mutual funds (cont.)

Category 1	Category 2	Category 3	F6	F7	F8	F9	F10	F11	F12	F13
Equity	Sweden	Sweden (normal)	0.0736	-0.0350	-0.0397	0.0423	-0.0302	0.0643	0.0021	0.0103
		Sweden small cap	0.0629	0.0494	-0.0756	0.0556	0.1019	0.1091	-0.0108	-0.0493
		Sweden index	0.1147	-0.0751	-0.0640	0.0605	-0.0496	0.0582	0.0014	0.0220
	Regional	Swedish equity and foreign equity	0.0844	-0.0017	0.0185	0.0252	-0.0083	0.0253	0.0253	0.0201
		Nordic countries	0.0812	-0.0032	-0.0452	0.0552	-0.0141	0.0350	0.0242	-0.0219
		Europe	0.1125	0.2386	-0.0174	0.0559	-0.0296	0.0100	0.0549	-0.0039
		Euroland	0.1056	0.2736	-0.0148	0.0567	-0.0242	0.0061	0.0183	-0.0299
		Europe small cap	0.2125	0.2218	-0.1353	0.0716	-0.0036	0.2088	0.0152	0.0320
		Europe index	0.0910	0.2856	0.0185	0.0667	-0.0355	-0.0108	0.0620	-0.0081
		North America and USA	0.0647	0.0355	0.2151	-0.0128	0.1005	-0.0014	0.0321	0.0953
		Asia and Far East	0.1308	0.0436	0.0688	0.0142	0.0506	0.0425	0.0054	0.0490
		Global	0.0872	0.1079	0.1297	0.0135	0.0507	-0.0016	0.0531	0.0516
		New markets	0.0525	0.0498	0.1097	-0.0139	0.0339	0.0333	0.2511	-0.0030
	Countries	Japan	0.0664	0.0374	0.0332	0.0406	0.0109	0.0295	0.0265	0.0116
		UK	0.1271	0.2667	0.0039	0.0547	-0.0638	0.0646	0.0526	0.1128
		Other countries	0.1232	0.1425	-0.0226	0.0615	-0.0037	0.0585	0.0301	0.0154
	Industry	IT and Communication	0.0326	-0.1085	0.0344	-0.0587	0.3212	0.0057	0.0930	0.0719
		Pharmacutical	0.0978	0.0503	0.1071	0.0784	-0.0340	-0.0255	-0.0016	0.3533
		Other industries	0.1463	0.2015	0.0240	0.0504	-0.0335	0.0298	0.0530	0.0223
Mixture	Mixture	Swedish equity and fixed income	0.0098	-0.0624	-0.0695	0.0588	-0.0586	0.0279	0.0031	-0.0348
		Swedish equity, Swedish and foreign fixed income	0.0538	-0.0233	0.0158	0.0429	0.0037	0.0352	0.0165	0.0277
		Foreign equity and fixed income	0.0412	0.1345	0.0515	0.0387	0.0193	0.0261	0.0341	0.0260
Generation	Generation	Pension in less than 10 years	0.0604	0.0052	0.0891	0.0430	0.0018	0.0330	0.0211	0.0528
		Pension in less than 20 years	0.0616	0.0257	0.0643	0.0071	-0.0006	0.0153	0.0137	0.0466
		Pension in more than 20 years	0.0683	0.0350	0.0709	0.0128	0.0038	0.0223	0.0131	0.0449
Fixed inc.	Fixed inc.	Sweden, short maturity	-0.7078	-0.0210	0.0105	0.0606	-0.0142	0.0049	-0.0014	-0.0322
		Sweden, long maturity	-0.1407	-0.0022	-0.0311	0.0402	0.0028	0.0153	-0.0022	-0.0121
		Europe and Euroland	-0.0565	0.0130	-0.0104	0.0507	-0.0185	0.1100	-0.0244	-0.0067
		Others	-0.0597	-0.0172	0.3534	0.1821	-0.0125	-0.0145	0.0405	0.0001
Default			0.1078	0.0460	-0.0198	0.0200	0.0098	0.0330	0.0271	0.0719
All funds			0.0413	0.0644	0.0308	0.0433	0.0092	0.0351	0.0311	0.0298

All general equity funds, domestic as well as foreign, have high loadings on this first factor, whereas the specialty equity funds and fixed income funds on average show lower corresponding loadings. The second factor is interpreted as a Japan related equity factor, due to the high loadings for the Japan country equity funds, whereas the third is labeled as an Asian or Far East equity factor, since the Asia and Far East regional equity funds show high loadings.

Mutual funds in the categories Europe and Euroland, and others, fixed income together with foreign equity and fixed income have high loadings on the fourth factor. Therefore, factor four is interpreted as a European fixed income factor. Factors five and six are associated with high negative loadings for Swedish fixed income funds with long and short maturity respectively. These factors are labeled as short in long- and short-term bonds or alternatively, long in corresponding bond yields.

The six most important rotated factors together account for 85 percent of the variance in the mutual fund returns (see Table 2). These factors are also relatively easy to interpret given the average factor loadings in Table 3. From factor seven and onwards, the interpretation becomes somewhat more awkward. As an additional aid in the interpretation information of the fund with the highest absolute loading on each factor is presented in Table 4. Here, a summary of the final identification of each of the 13 most important factors is also displayed.

In Table 3 European and UK equity funds are seen to load high on factor seven. Moreover, from Table 4, the highest loading on factor seven belongs to a European property fund. Given this information, this factor is interpreted as a European real estate factor. Factor eight is relatively straightforwardly interpreted as a US bond factor. Factor nine appears to affect only two mutual funds in our sample, namely the two fixed income funds from the Norwegian company Industrifinans. Given the

fact that each of the two loadings is quite high, almost 0.80, this factor is associated with a Norwegian fixed income dimension. Finally, based on loadings information from Tables 3 and 4, factors ten through 13 are associated with information technology stocks, high yield bonds, eastern European equity, and biotech stocks respectively.

Table 4: Factor identification: mutual funds with the highest factor loading for each factor

Factor	Highest absolute loading	Mutual fund with highest absolute loading	Factor identification
1	0.9687	Nordbanken Allemansfond Beta	World equity market portfolio
2	0.9323	JPM Japan Equity Fund	Japan equity
3	0.7645	Skandia Fond Aktiefond Far East	Far East equity
4	0.9208	BL - Short Term Euro	Euroland fixed income
5	-0.9263	Alfred Berg Obligationsfond	Swedish long term fixed income
6	-0.8896	EP Likviditetsfond Sverige	Swedish short term fixed income
7	0.5069	MSDW SICAV European Property Fund	European real estate
8	0.7291	MSDW SICAV US Bond	US Bonds
9	0.7882	Industrifinans Obligasjon/Obl. Utland	With Industrifinans only
10	0.4361	UBS (Lux) Equity Fund - Technology	IT
11	0.7336	Fleming European High Yield Bond Fund	High yield bonds
12	0.5178	Nomura Global Fund - Eastern European Sub-Fund	Eastern Europe
13	0.6734	Pictet G.S.F Compartiment Biotech	Biotech

To summarize the results from the factor analysis it is seen that by using e.g. the ten most important rotated factors, it is possible to account for more than 90 percent of the total variance of returns for the set of 465 mutual funds. In other words, allowing for roughly ten percent noise, there are about ten orthogonal asset classes available for investors to choose among in the initial round of investment for the defined contribution part of the Swedish pension system. In practice it is of course not possible to invest in the latent orthogonal factors. Nevertheless, it is argued that a similar set of choices can be obtained by replacing the rotated factors with real world indices, or indeed the 13 funds listed in Table 4, according to the interpretations above. In any case, there are clearly a lot of redundant mutual funds present in the initial choice set. Remember that each individual could choose to

invest in a maximum of five different mutual funds. A choice among say ten orthogonal factors, or distinctly different indices, would facilitate an easier and more efficient individual asset allocation than the actual choice between 465 different, or sometimes not so different, mutual funds.

4. Individual asset allocation: who holds what?

After identifying the factors driving the mutual fund returns, the focus is turned to the actual asset allocation choices made by individuals in the first round of the new Swedish pension system. Using a sample of individuals taking part in the 2000 defined contribution pension portfolio formation the individuals' loadings and communalities are analyzed with respect to the different factors. Individuals' factor exposures are related to a number of demographic and socio-economic variables using regression analysis, in order to find out who holds what, and whether asset allocation and diversification differ with respect to individual characteristics.

4.1 Data on individual choices and characteristics

The data come from the first round of investment choices made in the new Swedish pension system, coupled with a number of surveys on demographic and economic variables. The data constitutes a sample from a cross section of individuals in the Swedish work force. The first pension investments in the new pension system, in autumn 2000, involved 4.4 million individuals. Their investment choices are linked with individual demographic data collected by Statistics Sweden for the year 2000.[6] Statistics Sweden surveys 15,000 households that represent a cross section of the whole population in Sweden. This compiled data set makes it possible to study investment behavior in great detail. For each individual there is information on the

[6] Data sources from Statistics Sweden are, HEK 2000; a report on household economy, IoF 2000; income report and SUN 2000; educational status. These three reports are for the total population in Sweden. They are linked to a survey on 15,000 households reporting in-depth wealth statistics.

amount invested, which funds and how many funds the individual has invested in. Also, the age, gender, education, occupation, disposable income and net wealth for the same individual are included in the data set. From the 15,651 individuals with complete individual information in the data set, 10,375 individuals (66.4%) made an active investment decision. The remaining 5,276 individuals (33.7%) did not make an active investment decision. Instead, they are assigned to the default alternative: the Seventh Swedish Pension Fund, which is an equity fund run by the government. Based on the information regarding the individuals' choices, the default alternative is treated as an entirely passive choice. Even if an individual considered the default fund to be the optimal choice, and acted accordingly, he/she shows up as making a passive choice in the data set.

4.2 Individual factor loadings and asset allocation

The initial portfolio for each individual in the sample can contain positions in a maximum of five mutual funds, where each mutual fund loads on the common factors according to equation (1). Hence, for each individual k, the individual portfolio return in period t is characterized as:

$$(3) \qquad r_{k,t} = \sum_{p=1}^{h} w_{k,p} R_{p,t} = \sum_{p=1}^{h} w_{k,p} (\alpha_{p,1} F_{1,t} + \ldots + \alpha_{p,m} F_{m,t} + \varepsilon_{p,t})$$

where h = 1, …, 5 denotes the number of mutual funds chosen by individual k, and the weight $w_{k,p}$ is defined as the relative amount of money spent on fund p by individual k. The variance of the individual return can be written as:

$$(4) \qquad Var(r_{k,t}) = \sum_{p=1}^{h} w_{k,p}^2 Var(R_{p,t}) + \sum_{\substack{p=1 \\ p \neq q}}^{h} \sum_{q=1}^{h} w_{k,p} w_{k,q} Cov(R_{p,t} R_{q,t}) =$$

$$\sum_{p=1}^{h} w_{k,p}^2(\alpha_{p,1}^2+\ldots+\alpha_{p,m}^2+\sigma_p^2)+\sum_{\substack{p=1\\p\neq q}}^{h}\sum_{q=1}^{h} w_{k,p}w_{k,q}(\alpha_{p,1}\alpha_{q,1}+\ldots+\alpha_{p,m}\alpha_{q,m}+\sigma_{p,q})$$

where σ_p^2 denotes the unique return variance of fund p, and $\sigma_{p,q}$ is the covariance between fund p and q returns, that cannot be accounted for by the m most important factors. For each individual k, ignoring the unique variance and covariance terms in equation (4), the total communality is defined as follows:

(5) $$\overline{\alpha}_k = \sum_{p=1}^{h} w_{k,i}^2(\alpha_{i,1}^2+\alpha_{i,2}^2+\ldots+\alpha_{i,m}^2)+\sum_{\substack{p=1\\p\neq q}}^{h}\sum_{q=1}^{h} w_{k,p}w_{k,q}(\alpha_{p,1}\alpha_{q,1}+\ldots+\alpha_{p,m}\alpha_{q,m})$$

and the communality for each individual on a certain factor j is:

(6) $$\overline{\alpha}_{k,j} = \sum_{p=1}^{h} w_{k,p}^2\alpha_{p,j}^2+\sum_{\substack{p=1\\p\neq q}}^{h}\sum_{q=1}^{h} w_{k,p}w_{k,q}\alpha_{p,j}\alpha_{q,j}$$

In equation (5) $\overline{\alpha}_k$ is a measure of the exposure for all m factors for individual k, whereas in equation (6) $\overline{\alpha}_{k,j}$ is a corresponding measure of the individual's exposure to factor j only, where $j<m$. Both measures are calculated for individuals making an active choice, as well as for individuals with the passive choice of the default fund alternative. Note that all individuals with the default choice have the same exposure to the factors according to equation (5) and (6).

Differences among individuals are analyzed with respect to the factor communalities, and thus asset allocation. However, it is first important to take into account the sample selection issue that there are "passive" individuals in the sample, the ones not making an active choice or with preference for the default alternative. To simply ignore the "passive" individuals would induce a selection bias if their

characteristics prove to be different from those of the "active" individuals. Hence, it is necessary to jointly model the factor communalities and the likelihood of making an active choice. This is accomplished within a nested type of model. In essence, the model presumes that each individual jointly considers two investment choices. The first is the choice of whether to be active or passive, and the second, given that the individual decides to make an active choice, is to choose the desired loading on each factor. The model is estimated using the two-step procedure according to Heckman (1976), where first a probit model is used to estimate the likelihood of making an active choice, and second, a regression analysis is performed to model individuals' factor communalities, taking the likelihood of activity into account.

First, consider the choice of activity for an individual. Let z be a nominal variable with two outcomes: z = 1 if the individual chooses to make an active investment decision and z = 0 if he or she chooses to be passive. Define $\Pr(z=1)$ and $\Pr(z=0)$ as the individuals' probability of making an active or passive choice respectively. For each individual k the choice of activity is modeled according to:

(7) $$z_k = \boldsymbol{\gamma}'\mathbf{w}_k + \xi_k$$

where $\mathbf{w}_k$ is a vector of explanatory variables for the activity choice of individual k, $\boldsymbol{\gamma}$ is a vector of coefficients measuring the effect of each explanatory variable on the activity choice, and ξ_k is a residual term. The coefficients in equation (7) are estimated using the maximum likelihood probit estimation technique. Accordingly, $\Pr(z_k = 1) = \Phi(\boldsymbol{\gamma}'\mathbf{w}_k)$ and $\Pr(z_k = 0) = 1 - \Phi(\boldsymbol{\gamma}'\mathbf{w}_k)$, where $\Phi(.)$ denotes the standard normal cumulative distribution function.

Second, the individual communality on each factor $\overline{\alpha}_{k,j}$ from equation (6) is related to a set of explanatory variables using regression analysis. The regression analysis consists of individuals with an active choice only. Hence, to take the activity choice

into account, it is necessary to perform a conditional regression analysis with the dependent variable $\overline{\alpha}_{k,j}|z_k = 1$. Conditioning on the variables that are thought to help explaining individuals' communalities, and using the second step in Heckman's (1976) estimation procedure, the regressions are formulated as:

(8) $$\overline{\alpha}_{k,j} = \boldsymbol{\beta}'_j \mathbf{x}_k + \beta_{\lambda,j} \lambda_k(\hat{\boldsymbol{\gamma}}'\mathbf{w}) + \eta_{k,j}$$

Equation (8) forms a system of m regression equations, where $\mathbf{x}_k$ is a vector of explanatory variables, $\boldsymbol{\beta}_j$ is a vector of regression coefficients in equation j, including a constant term $\beta_{0,j}$ and slope terms $\beta_{q,j}$, relating the factor-specific communality j to explanatory variable q, and $\eta_{k,j}$ is a corresponding error term. The function $\lambda_k(\hat{\boldsymbol{\gamma}}'\mathbf{w}) = \phi(\hat{\boldsymbol{\gamma}}'\mathbf{w})/\Phi(\hat{\boldsymbol{\gamma}}'\mathbf{w})$ is known as the inverse Mills ratio, or the hazard function, for the normal distribution from the probit estimation of equation (7). Heckman (1976) motivates the inclusion of $\lambda_k(\hat{\boldsymbol{\gamma}}'\mathbf{w})$ as an explanatory variable in equation (8). Given that only individuals who have made an active choice ($z_k = 1$) are used in the regressions according to equation (8), the regression coefficients in $\boldsymbol{\beta}_j$ now can be consistently estimated without incurring a selection bias.

In equation (8), for each individual k, the asset allocation choices between the m different factors, and the individual factor communalities, are expected to be interrelated. Hence, the error terms are likely to be correlated across equations. To take this cross-equation correlation into account, the m regression equations are estimated simultaneously using Zellner's (1962) SUR technique.

In the probit estimation of equation (7), the vector $\mathbf{w}_k$ contains a set of individual characteristics as explanatory variables. The inclusion of explanatory variables in the first pass analysis of individual activity is based on the results of Engström and Westerberg (2003), and Karlsson and Nordén (2004). The vector of explanatory

variables $\mathbf{x}_k$ in equation (8) includes all variables in $\mathbf{w}_k$, plus an additional set suitable for the SUR model, but not for the probit model.

Both individual activity and asset allocation are related to the level of investor sophistication (Grinblatt and Keloharju, 2001, Karlsson and Nordén, 2004). Investor sophistication is represented by four sets of variables in both the probit and the SUR regression analysis: i) level of education, less than high school, high school or more than high school education, where dummy variables for less and more than high school education are included, EDU_1 and EDU_2 respectively; ii) the (natural log of the) amount of money invested in the pension system (MONEY), where it is argued that a large amount of money should cause the investor to pay closer attention to his or her investment choice; iii) the natural log of disposable income (INCOME); and iv) the natural log of net wealth (WEALTH). These variables are assumed to be positively correlated with investor sophistication and. Moreover, following previous evidence, more sophisticated individuals ought to invest in more diversified portfolios with respect to the different factors.

Related to the investor sophistication issue are individuals' total portfolios of financial holdings, apart from the investment in the defined contribution pension fund system. Accordingly, the dummy variables RISKY = 1 if an individual owns risky assets (stocks or other mutual funds) prior to the pension investment, and zero otherwise, and NONRISKY = 1 if an individual has prior holdings of risk-free assets (bonds or other fixed income securities), and zero otherwise, are included in the SUR regression analysis. To some extent, an individual can be regarded as more sophisticated with respect to asset allocation if he or she has prior experience with assets like stocks, mutual funds or bonds. Hence, the two dummy variables are incorporated in the probit regression model as well.

Individuals' occupation also influences their asset allocation decisions, in particular when the decisions are related to pension investments. Karlsson and Nordén (2004) find it more likely for an individual to be home biased if he or she has a high level of job security. Such an individual is more likely to stay employed, and still earn an income, even if domestic markets go down. Also, the return on investments will increase if the domestic market goes up, thus hedging the individual's purchasing power. In Sweden, an individual working in the public sector usually has a high level of job security and the risk of unemployment is relatively small.[7] Hence, a different asset allocation behavior is expected for government employees than for individuals who are privately or self-employed. To account for individuals' occupation, dummy variables are included, in both the probit and the SUR model, for private employment (OCC_2), self-employment (OCC_3), and unemployment (OCC_4), to separate from the base case individuals who are government employees.

Included is also a gender dummy variable MEN = 1 if the individual is a man, and zero if she is a woman, a dummy variable MARRIED = 1 if the individual is married, and zero if he or she is unmarried, and an interaction term MEN_MAR = 1 if the individual is a married man, and zero otherwise. Barber and Odean (2001) find evidence suggesting that men are more overconfident than women, and also relatively more likely to take risks. Moreover, Barber and Odean (2001) argue that marriage might weaken the gender effect. Given these results, differences in asset allocation and factor communalities with respect to marital status and gender are expected, in particular with respect to individuals' choices between risky and not so risky factors according to the SUR model in equation (8). Given the results from Engström and Westerberg (2003), that gender and marital status are important

[7] According to statistics from Statistics Sweden and The National Board of Labor Markets, in the year 2000, the percentage of employees loosing their jobs was 1 percent in the private sector and 0.1 percent in the public sector.

explanatory sources for the activity choice, the three dummy variables, related to gender and marital status, are included in the probit analysis as well. Finally, individual age is included in both the probit and the SUR model estimation. Age is directly related to the investment horizon, which is known to affect asset allocation decisions (Karlsson, 2005).

As additional explanatory variables in the SUR model, four dummy variables are used to represent the number of chosen mutual funds for each individual within the pension system (each investor can choose one to five funds). Hence, $D_2 = 1$ if two funds are chosen and zero otherwise, $D_3 = 1$ if three funds are chosen and zero otherwise, etc., where the base case is to choose one fund in the regression model. Benartzi and Thaler (2001) indicate that the complicated reality of portfolio diversification may cause inexperienced investors to diversify in a naïve manner, believing that many assets diversify better than fewer assets. This is not always true in the Swedish pension system, where the investment opportunity set contains a lot of mutual funds but only a few asset classes, or orthogonal factors. Nevertheless, it is reasonable to assume that it is more likely for an individual to load on several factors the larger amount of funds he or she chooses.

The SUR model also contains a control for the percentage transactions cost paid by each individual for the contribution pension fund investment (FEE) and a proxy for the risk associated with each individual investment. Two measures of portfolio risk are sued. The first measure (STDEV) directly uses the numerical value of the annualized standard deviation of three-year monthly historical portfolio returns for the three years 1997 through 1999. The portfolio standard deviation is calculated by taking each portfolio's weighted average returns for the past 36 months and then calculating the standard deviation of this average return series, thus capturing covariance in returns. The second measure (RISKCAT) is simply each individual's

weighted average category of risk, according to the classification based on standard deviation (see Table 1).[8] Finally, the variable RETURN is calculated based on the fund information as exemplified in Table 1. The analysis uses the compounded annual return for the three years 1997 through 1999. For each individual, the return is calculated as the weighted average for all funds in the portfolio. One motivation for including historical returns in the regression analysis is to control for possible momentum effects (Chan et al., 1996), that individuals choose mutual funds, and thus factors, with positive historical returns, hoping for future positive returns as well.

The estimation results from both the first pass probit regression and the second pass SUR regression are presented in Table 5. The first column of Table 5 contains the estimated coefficients and p-values in the probit model according to equation (7), where all 15,651 observations are used in the estimation. Evidently, most of the coefficients associated with the explanatory variables related to investor sophistication are significant, and consistent with the same story, namely that more sophisticated investors have a higher likelihood of making an active choice. Individuals with less than high school education (variable EDU_1) show a significantly lower likelihood of being active than the benchmark individuals with high school education. Moreover, more wealthy individuals, as measured with the variables MONEY and WEALTH, but not with INCOME, have a higher likelihood of making an active choice. Individuals with previous experience with risky assets are more likely to make an active choice, whereas previous holdings of non-risky assets like bonds are not important for the activity choice. Occupation matters to some extent as self-employed and unemployed individuals show a significantly lower likelihood of making an active choice.

[8] The empirical analysis is concentrated on the RISKCAT measure of risk, because the model fit is better using this measure rather than STDEV. However, the regression results are virtually the same irrespective of which measure of risk is uses as explanatory variable.

Table 5: Two-step Heckman regression (SUR) results for individual factor communalities

	z_k	$\overline{\alpha}_{k,1}$	$\overline{\alpha}_{k,2}$	$\overline{\alpha}_{k,3}$	$\overline{\alpha}_{k,4}$	$\overline{\alpha}_{k,5}$	$\overline{\alpha}_{k,6}$	$\overline{\alpha}_{k,7}$	$\overline{\alpha}_{k,8}$	$\overline{\alpha}_{k,9}$	$\overline{\alpha}_{k,10}$	$\overline{\alpha}_{k,11}$	$\overline{\alpha}_{k,12}$	$\overline{\alpha}_{k,13}$
Constant	-1.4530	107.40	5.4119	-2.9839	0.4966	3.8815	-0.1723	-0.1769	0.3476	1.3144	-5.0750	-1.8433	-0.4221	-8.3490
	0.0000	0.0000	0.0750	0.3776	0.7135	0.8256	0.9645	0.8974	0.7900	0.2854	0.0006	0.0000	0.6073	0.0070
EDU_1	-0.1273	1.5157	-0.0758	-0.1486	-0.0708	-0.6410	-0.5260	0.0251	0.0492	-0.0039	0.0138	-0.0736	0.0011	-0.2011
	0.0000	0.0187	0.6174	0.3789	0.2947	0.0661	0.0064	0.7144	0.4493	0.9492	0.8510	0.0013	0.9792	0.1932
EDU_3	-0.0459	0.4701	0.1856	0.1083	-0.0187	-0.5432	-0.1740	-0.0646	-0.1451	0.0566	-0.1158	-0.0120	0.0163	0.0595
	0.0712	0.1653	0.0200	0.2220	0.5973	0.0030	0.0862	0.0725	0.0000	0.0796	0.0228	0.3178	0.4505	0.4635
MONEY	0.1945	-0.9487	-0.1274	0.2060	-0.0157	0.7235	0.5452	-0.0534	-0.0828	-0.0231	-0.1172	0.1283	-0.0568	0.2688
	0.0000	0.2772	0.5353	0.3676	0.8637	0.1255	0.0369	0.5650	0.3480	0.7812	0.2400	0.0000	0.3057	0.1988
INCOME	0.0016	-0.1559	0.0027	0.0223	0.0063	-0.0104	0.0274	0.0060	0.0037	0.0106	-0.0014	0.0011	-0.0026	0.0173
	0.7894	0.0148	0.8597	0.1908	0.3584	0.7668	0.1587	0.3833	0.5739	0.0859	0.8489	0.6290	0.5370	0.2665
WEALTH	0.0096	-0.0786	-0.0131	0.0153	0.0007	0.0623	0.0264	-0.0034	-0.0085	-0.0038	-0.0093	0.0081	-0.0056	0.0178
	0.0000	0.0929	0.2330	0.2111	0.8824	0.0137	0.0596	0.4997	0.0716	0.3945	0.0805	0.0000	0.0584	0.1106
RISKY	0.3054	-2.6164	-0.1716	0.3446	0.0483	1.3841	1.0785	0.0227	0.0026	-0.1302	-0.0299	0.2217	-0.0731	0.4294
	0.0000	0.0476	0.5808	0.3191	0.7270	0.0527	0.0064	0.8712	0.9846	0.3008	0.8428	0.0000	0.3838	0.1748
NONRISKY	0.0363	-0.1152	-0.0393	0.1458	-0.0509	-0.0193	0.1497	-0.0079	-0.0540	0.0027	0.0067	0.0213	0.0047	0.1230
	0.1404	0.7056	0.5844	0.0678	0.1111	0.9067	0.1011	0.8083	0.0798	0.9257	0.8466	0.0493	0.8095	0.0924
OCC_2	0.0217	-0.3074	0.0694	0.0962	0.0446	0.1849	-0.0074	-0.0218	-0.0831	0.0665	-0.0751	0.0321	-0.0097	0.1264
	0.4218	0.3003	0.3205	0.2158	0.1515	0.2494	0.9335	0.4891	0.0056	0.0186	0.0267	0.0024	0.6056	0.0756
OCC_3	-0.2089	2.7531	0.4632	0.4376	-0.0937	-2.0478	-0.8688	-0.0472	-0.0043	0.0987	0.0781	-0.1799	0.1360	-0.8593
	0.0001	0.0144	0.0803	0.1375	0.4266	0.0008	0.0099	0.6930	0.9697	0.3574	0.5435	0.0000	0.0572	0.0014
OCC_4	-0.2228	1.8231	0.5067	-0.3636	-0.0352	-0.5972	-0.9685	0.0269	0.0015	0.0405	-0.1627	-0.1563	0.0824	-0.4439
	0.0000	0.1176	0.0646	0.2332	0.7729	0.3432	0.0055	0.8277	0.9902	0.7150	0.2214	0.0002	0.2659	0.1118
MEN	-0.1431	-0.3895	0.3539	0.1501	0.0543	-0.6610	-0.0734	0.0929	0.1768	0.0777	0.0442	-0.1326	0.0634	-0.4033
	0.0013	0.6448	0.0752	0.4975	0.5394	0.1481	0.7718	0.3009	0.0384	0.3342	0.6468	0.0000	0.2379	0.0464
MARRIED	0.2331	-1.8280	-0.1172	0.4658	-0.0198	0.9427	1.0292	0.0014	0.0021	-0.0501	-0.1038	0.1468	-0.0654	0.2660
	0.0000	0.0902	0.6443	0.0992	0.8610	0.1062	0.0014	0.9901	0.9848	0.6256	0.3997	0.0001	0.3404	0.3035

Table 5 (cont.): Two-step Heckman regression (SUR) results for individual factor communalities

	z_k	$\overline{\alpha}_{k,1}$	$\overline{\alpha}_{k,2}$	$\overline{\alpha}_{k,3}$	$\overline{\alpha}_{k,4}$	$\overline{\alpha}_{k,5}$	$\overline{\alpha}_{k,6}$	$\overline{\alpha}_{k,7}$	$\overline{\alpha}_{k,8}$	$\overline{\alpha}_{k,9}$	$\overline{\alpha}_{k,10}$	$\overline{\alpha}_{k,11}$	$\overline{\alpha}_{k,12}$	$\overline{\alpha}_{k,13}$
MANMAR	0.0633	0.0200	-0.1514	-0.0431	0.0053	0.4746	-0.1450	-0.0849	-0.1417	-0.0615	0.0349	0.0901	0.0087	0.2239
	0.2075	0.8806	0.3334	0.8047	0.9392	0.1871	0.4663	0.2294	0.0350	0.3322	0.6458	0.0001	0.8379	0.1601
AGE	-0.0067	0.0429	0.0049	-0.0162	0.0001	-0.0362	-0.0339	0.0034	0.0003	0.0021	0.0146	-0.0054	0.0044	0.0003
	0.0000	0.1513	0.4869	0.0383	0.9759	0.0253	0.0002	0.2850	0.9135	0.4701	0.0000	0.0000	0.0220	0.9701
D2	-	-2.4729	-0.6763	-0.8628	0.2455	1.8413	0.8092	0.1594	0.2553	0.0850	-0.1741	0.0690	-0.0395	-0.0156
		0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0005	0.0000	0.0396	0.0004	0.0000	0.1519	0.8803
D3	-	-5.1606	-0.8622	-1.1240	0.3144	3.6718	1.1434	0.3715	0.3891	0.0399	-0.0530	0.0673	-0.0552	-0.1167
		0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.2933	0.2448	0.0000	0.0295	0.2221
D4	-	-4.9928	-0.8487	-1.2057	0.4012	2.9725	1.3780	0.4937	0.4562	0.0248	-0.0084	0.1255	-0.0509	-0.2119
		0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.5339	0.8605	0.0000	0.0556	0.0345
D5	-	-6.0833	-0.5177	-0.9278	0.4027	3.0565	1.3938	0.5934	0.4666	0.0566	-0.0916	0.1254	0.0061	-0.0018
		0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1357	0.0440	0.0000	0.8086	0.9849
FEE	-	-22.012	0.9489	4.2832	0.0162	0.3128	1.2457	0.6101	0.7671	0.4298	0.8761	0.1104	0.6728	3.9228
		0.0000	0.0000	0.0000	0.7464	0.2254	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
RISKCAT	-	-2.5687	0.6657	1.9574	0.0078	-3.4036	-2.8167	0.5065	0.5568	-0.2455	0.2937	0.0156	0.4511	1.4163
		0.0000	0.0000	0.0000	0.8156	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1671	0.0000	0.0000
RET	-	8.7115	-1.3420	-2.7826	-0.1192	-1.1221	0.3546	-0.4099	-0.5173	0.0013	2.3849	-0.0604	-0.2813	-1.3872
		0.0000	0.0000	0.0000	0.0004	0.0000	0.0002	0.0000	0.0000	0.9658	0.0000	0.0000	0.0000	0.0000
λ_k	-	-18.297	-1.8338	2.6203	0.6407	10.720	7.1617	-0.1727	-0.3986	-0.4074	-0.2156	1.4283	-0.4244	2.9925
		0.0225	0.3313	0.2121	0.4455	0.0135	0.0029	0.8393	0.6227	0.5938	0.8139	0.0000	0.4051	0.1194
$\overline{R}^2$	0.0641	0.2996	0.0742	0.2442	0.0442	0.1558	0.1486	0.1323	0.1516	0.0437	0.5759	0.0533	0.1893	0.2292

However, there is no significant difference between privately employed individuals and individuals employed by the government. Finally, gender, marital status and age matters for the choice, where it seems like young, married women are more likely to choose actively than older, unmarried men. The gender result is rather surprising, and runs counter to the expectations based on Barber and Odean (1998).

The rest of Table 5 contains the estimated coefficients and p-values from the SUR model according to equation (8).[9] The model consists of m = 13 equations, where the 13 most important factors are retained. The explanatory variables are the same as in the probit equation, plus a set of control variables for the number of actively chosen funds, the funds' transactions costs, risk level, historical returns, and the inverse Mills ratio from the probit regression. The analysis focuses on the SUR results for the coefficients representing individual characteristics. Note however that most of the control variables are associated with significant coefficients, and thus are important.

Regarding the coefficients for the variables that are proxies for investor sophistication, individuals with a low education level (EDU_1) have a significant tendency for loading relatively higher on the first, overall market factor (the regression equation for the dependent variable $\overline{\alpha}_{k,1}$, in the second column of Table 5). Moreover, individuals with high income, high wealth, and with previous holdings of risky assets, have significantly smaller loadings on the first factor. This result is consistent with the significantly negative coefficient for the inverse Mills ratio (λ_k) in the regression equation for the first factor. This coefficient represents an indirect effect on the loading on the first factor from the explanatory variables in the probit model. Hence, more sophisticated individuals have a higher probability of

[9] Note that each reported regression coefficient in Table 5 equals 100 times the corresponding estimated coefficient.

making an active choice, and the result of this activity is to reduce the loading on the overall market portfolio.

The low education dummy variable (EDU_1) is associated with significantly negative coefficients in the regression equations for communalities for factor five and six (equations for $\overline{\alpha}_{k,5}$ and $\overline{\alpha}_{k,6}$). In the same equations, individual wealth (WEALTH) and previous experience of risky assets (RISKY) are associated with significantly positive coefficients. These results are consistent with the idea that more sophisticated individuals have relatively lower loadings on Swedish bonds than less sophisticated individuals.[10] Moreover, the coefficients for the Mills ratio are significantly positive in the equations for $\overline{\alpha}_{k,5}$ and $\overline{\alpha}_{k,6}$. Hence, the more active sophisticated individuals are using the activity not only to reduce the allocation to the market portfolio, but also to reduce the allocation to Swedish long- and short-term bonds.

Given the results that sophisticated investors tend to have relatively less of their holdings allocated to the market portfolio or to bonds than less sophisticated investors, the next step is to analyze in which asset classes they have relatively larger holdings. In Table 5, in the regression equation for $\overline{\alpha}_{k,11}$, i.e. the equation for individual communalities on the high yield bond factor, significantly negative coefficients for the EDU_1 variable, and significantly positive coefficients for the WEALTH and RISKY variables, are observed. In addition, these three equations have significantly positive coefficients for the λ_k variable. Hence, the sophisticated investors appear to first, according to the probit analysis, be more likely to actively choose, and second, to use the active choice to add some high yield bonds to their portfolios.

[10] According to the results in Table 3 and 4, Swedish bond funds have significantly negative loadings on factors five and six. Therefore, if an investor has a negative loading on either factor five or six it should be interpreted as a positive loading on actual bonds.

For the gender and marital status variables, the results show a significant higher tendency for men to load higher on the US bond factor (eight), and lower on the yield factor (eleven) and the biotech factor (13) than women. Interestingly, men's loadings on factors eight and eleven are more different than women's for single rather than married men. The interaction term between the male and marriage dummy variables have a significantly negative coefficient in the $\overline{\alpha}_{k,8}$ equation and a significantly positive coefficient in the $\overline{\alpha}_{k,11}$ equation. Hence, one asset allocation effect seems to be to make men to choose asset classes more in line with women's wishes. It might not be an overconfidence related issue, but the marital effect is similar to the argument in Barber and Odean (2001) that marriage might weaken a gender effect, in this case with respect to asset allocation.

Finally, at the five percent significance level, older individuals tend to load significantly higher on domestic long- and short-term bonds (lower loadings on factors five and six), lower on Far-East equities (factor three) and high yield bonds, which are consistent with a lower risk-taking at an older age, or the time diversification idea (see e.g. Karlsson, 2005). However, older individuals also load significantly higher on the IT and Eastern Europe factors (ten and eleven). These results contest the time diversification idea.

5. Individual performance

How well do the individual portfolios perform, and is it possible to see any systematic differences among individuals' performance in the Swedish pension plan? To answer these questions, Jensen's alpha is computed for each individual portfolio over the four-year period, from the initiation of the new pension plan in 2000, through 2004. Moreover, the alphas are compared across individuals taking

their demographic and socio-economic characteristics into account.[11] Jensen's alpha for each individual is estimated in the following regression:

$$(9) \qquad r_{k,t} - r_{f,t} = a_k + \sum_{s=1}^{S} b_{k,s}(I_{s,t} - r_{f,t}) + e_{k,t}$$

where $r_{k,t}$ is the portfolio return for individual k in period t, $r_{f,t}$ is the risk-free rate of return in period t, a_k is Jensen's alpha for individual k, $I_{s,t}$ is the return on index s in period t, $b_{k,s}$ is the sensitivity of individual k to index s, and $e_{k,t}$ is the residual for individual k in period t.

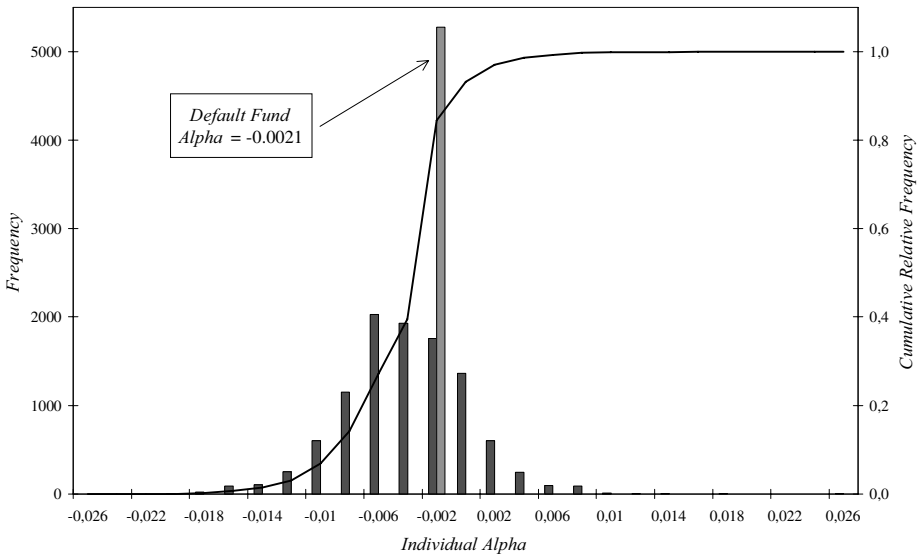
Jensen's alpha, a_k in equation (9), measures the return the individual earns in excess of what he/she would have earned if he/she held a portfolio with broad market indices with the same risk.[12] Having identified the factors generating the mutual fund returns over the sample period, the factor identification from Table 4 is used to specify appropriate indices in equation (9). For all individuals alike, a six-index model is used, with the MSCI World index (to represent Factor 1 in Table 4), the MSCI Japan index (Factor 2), the MSCI Far East, excluding Japan (Factor 3), the Serfiex DEMI Euro Zone T-bill index (Factor 4), Handelsbanken Swedish 5-10 Years Government Bond index (Factor 5), and the Merrill Lynch Euro High Yield index (Factor 11). Monthly returns on the Swedish one-month Treasury bill rate (Factor 6) are used as a proxy for the risk-free interest rate.

[11] Note that the performance of each individual's portfolio is evaluated, given the initial composition in 2000. Thus, the possibility of individuals dynamically changing their portfolios over the sample period is not taken into account.

[12] The overwhelming evidence is that alpha on average is negative for mutual funds; see e.g. Blake et al. (1993), Grinblatt and Titman (1996), Jensen (1968), Sharpe (1966), and Wermers (2000). Hence, a negative individual alpha on average is not inconsistent with individuals picking better performing funds; see Blake et al. (2005). Note also that we analyze the relative performance for different types of individuals, not individual performance per se.

Figure 1 displays a frequency diagram for Jensen's alpha for all individuals, including people with the passive default choice. All passive individuals have a monthly alpha equal to -0.0017, whereas the average alpha for the active individuals equals -0.0043, with a standard deviation equal to 0.0040. Hence, on average the active individuals have a significantly worse performance than people in the default fund.[13]

Figure 1: Individual performance measured as Jensen's alpha (monthly basis)



To investigate individual differences with respect to performance, it is important to take the passive individuals properly into account. The results from the probit choice model according to equation (7) indicate significant differences between active and passive individuals. Therefore, the passive individuals are considered as a group

[13] A *t*-test of the hypothesis that the average alpha for the active individuals is not different from alpha for the default fund results in a *t*-statistic equal to -64.3, and thus a rejection of the hypothesis at any reasonable significance level.

with a common performance (alpha = -0.0071), and divide the active individuals into the following groups based on performance, $a_k \leq -0.002$, $-0.002 < a_k \leq 0.002$, and $a_k > 0.002$. Let y_k be a nominal variable with J = 4 categories defined as $y_k = 1$ if $-0.002 < a_k \leq 0.002$, $y_k = 2$ if $a_k \leq -0.002$, $y_k = 3$ if $a_k > 0.002$, and $y_k = 4$ to represent the default. Moreover, let $\Pr(y_k = m \mid \mathbf{w}_k)$, m = 1, …, 4, be the conditional probability for individual k of observing the outcome m given the explanatory variables $\mathbf{w}_k$. Following Theil (1969), the multinomial logit model is used to estimate the probabilities for individual k as:

(10)
$$\Pr(y_k = m \mid \mathbf{w}_k) = \frac{1}{1 + \sum_{j=2}^{J} \exp(\mathbf{w}_k \boldsymbol{\beta}_j)} \qquad \text{for m = 1}$$

$$\Pr(y_k = m \mid \mathbf{w}_k) = \frac{\exp(\mathbf{w}_k \boldsymbol{\beta}_m)}{1 + \sum_{j=2}^{J} \exp(\mathbf{w}_k \boldsymbol{\beta}_j)} \qquad \text{for m > 1}$$

The constraint $\boldsymbol{\beta}_1 = 0$ for m = 1 is made to ensure that the probabilities are identifiable. Note that the base category in the multinomial logit model is an alpha close to zero ($-0.002 < a_k \leq 0.002$). The reasons are twofold; first, this formulation allows an evaluation of the probability of having better or worse performance than the middle, close to zero-alpha base case. Second, the alpha for the passive individuals is entailed in the base category, which to some extent isolates the probability of being passive from the performance issue.

Table 6 presents the results from the multinomial logit estimation. For each explanatory variable, a Wald test statistic is reported, which is χ^2-distributed under the null hypothesis that the coefficients in the probability equations for y = 2 and y = 3 are jointly equal to zero. Thus, the test does not include the corresponding coefficient in the default probability equation.

Table 6: Results from the multinomial logit model

	Pr(y = 2)	Pr(y = 3)	Pr(y = 4)	χ^2-statistic
CONSTANT	1.5345	-0.0348	4.3186	-
	(0.0158)	(0.9672)	(0.0001)	
EDU_1	-0.1232	0.1505	0.1421	7.346
	(0.0659)	(0.2374)	(0.0361)	(0.0254)
EDU_3	0.1957	0.1661	0.2236	11.58
	(0.0007)	(0.1801)	(0.0003)	(0.0031)
MONEY	0.0645	-0.2414	-0.2955	19.29
	(0.0954)	(0.0017)	(0.0001)	(0.0001)
INCOME	-0.0072	-0.0357	-0.0099	2.056
	(0.6154)	(0.1540)	(0.4933)	(0.3577)
WEALTH	0.0178	0.0240	-0.0021	14.76
	(0.0002)	(0.0238)	(0.6772)	(0.0006)
RISKY	0.4545	0.1000	-0.1797	73.96
	(0.0001)	(0.3929)	(0.0016)	(0.0001)
NONRISKY	-0.0550	-0.2966	-0.1171	6.777
	(0.3188)	(0.0093)	(0.0454)	(0.0338)
OCC_2	-0.0374	0.2827	-0.0503	6.315
	(0.5263)	(0.0363)	(0.4238)	(0.0425)
OCC_3	0.7228	1.0605	0.9332	26.36
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
OCC_4	0.0757	0.4903	0.4399	6.724
	(0.4362)	(0.0097)	(0.0001)	(0.0347)
MALE	-0.2038	0.7245	0.1339	18.51
	(0.0723)	(0.0024)	(0.2320)	(0.0001)
MARRIED	-0.0683	0.0350	-0.4319	0.743
	(0.4514)	(0.8704)	(0.0001)	(0.6897)
MAR_MALE	0.2481	-0.3287	0.0522	8.369
	(0.0449)	(0.2089)	(0.6740)	(0.0152)
AGE	-0.0303	0.0087	-0.0096	189.9
	(0.0001)	(0.0939)	(0.0002)	(0.0001)

Table 5 contains results from the estimation of the multinomial logit model in Equation (10). The dependent variable y_k, has four possible outcomes (m = 1, …, 4), where each of the first three corresponds to an "active" choice, and an alpha in a range according to $y_k = 1$ if $-0.002 < a_k \leq 0.002$, $y_k = 2$ if $a_k \leq -0.002$, $y_k = 3$ if $a_k > 0.002$, and $y_k = 4$ represents the "passive" default. $\Pr(y_k = m \mid \mathbf{w}_k)$ is the probability of observing outcome m given $\mathbf{w}$. The variables in $\mathbf{w}$ are EDU_1 (a dummy variable equal to 1 if the education level of the individual is below high school), EDU_3 (dummy variable equal to 1 if the education level is above high school, the base case represents high school education), MONEY (total initial investments in pensions funds for each individual in SEK), INCOME (individual's disposable income in SEK), WEALTH (market value in SEK of financial assets and real estate holdings, net of debt for each individual), RISKY (dummy variable equal to 1 if the individual has risky holdings), NONRISKY (dummy variable equal to 1 if the individual has holdings with low risk), OCC_2 through OCC_4 are dummy variables for individuals occupation (OCC_2 represents private sector employment, OCC_3 self-employment, OCC_4 unemployment, whereas the base case represents government employment), MALE (dummy variable equal to 1 if the individual is male), MARRIED (dummy variable equal to 1 if the individual is married), MAR_MALE (MALE× MARRIED) and AGE (in years). The probabilities are modeled as:

$$\Pr(y_k = m \mid \mathbf{w}_k) = \frac{1}{1 + \sum_{j=2}^{J} \exp(\mathbf{w}_k \boldsymbol{\beta}_j)} \quad \text{for } m = 1, \quad \Pr(y_k = m \mid \mathbf{w}_k) = \frac{\exp(\mathbf{w}_k \boldsymbol{\beta}_m)}{1 + \sum_{j=2}^{J} \exp(\mathbf{w}_k \boldsymbol{\beta}_j)} \quad \text{for } m > 1$$

where $\boldsymbol{\beta}_1 = 0$ for the first outcome. The model is estimated using the maximum likelihood technique outlined in Berndt et al. (1974), and the heteroskedasticity-consistent covariance matrix according to White (1980). The estimated coefficients are presented for each probability and explanatory variable, with *p*-values in parentheses. Each χ^2-statistic results from a Wald test for the hypothesis that each explanatory variable does not affect the likelihood of outcomes $y_k = 2$ and $y_k = 3$, relative the first outcome $y_k = 1$, and is χ^2-distributed with four degrees of freedom.

Each test statistic should be interpreted as a test for an effect of each explanatory variable on the probability of performing either better or worse than the group of individuals with an alpha close to zero. As in the Heckman analysis above, the default probability equation is used for control purposes only. Most of the results for the probability equation Pr($y_k = 4$) are consistent with the results from the first pass probit regression in Table 5. However, note the opposite signs of the coefficients, as Pr($y_k = 4$) refers to the probability of passive choice, whereas in the probit model in equation (7) is for the probability of an active choice.

From the probability equations of main interest, corresponding to worse or better performance than the base category, the results in Table 6 show that well educated individuals have a significantly higher probability of a worse performance. Moreover, at the ten percent significance level, there is a lower probability for individuals with less than high school education to perform worse than the base category. Among the other proxies for investor sophistication it is noted that individuals with a lot of money invested in the premium pension have a significantly higher probability of performing worse (only at the ten percent level), and a lower probability of performing better than the base category. Likewise, previous experience with risky (non-risky) assets increases (decreases) the probability of having a worse (better) performance than the benchmark case. These results are consistent with the idea that more sophisticated individuals perform significantly worse than less sophisticated people. As for the remaining two proxies for

sophistication, income is not a significant determinant of alpha, whereas wealth is associated with significantly positive coefficients in both the worse- and better-performance categories. Hence, more wealthy individuals have a higher probability of performing worse and better than less wealthy individuals, who in turn are more likely to have an alpha close to zero.

Privately employed (and unemployed) individuals, represented by the dummy variable OCC_2 (OCC_4) in Table 6, show a significantly higher likelihood than government employees of performing better than the close-to-zero alpha category, at the five percent significance level. Moreover, self-employed individuals are significantly more likely to perform both worse and better than government employees.

Men have a significantly higher (lower) probability to be in the category with better (worse) performance than the zero-alpha category, at the one (ten) percent significance level. Hence, the results are consistent with a significantly higher likelihood for men to perform better than women. Interestingly, the gender effect that men perform better than women appears to be higher for single men than married men. The coefficient for the interaction term between the male and marriage dummy variables is significantly positive, at the five percent level, in the probability equation for the category with worse performance than the close-to-zero category. Although no significant difference is observed between married and unmarried men regarding the likelihood of being in the performance category with highest alpha, the gender/marriage results are rather straightforward. Men have better-performing portfolios than women in general, but single men perform even better than married men. One perhaps controversial way of interpreting these results is that married men are under “bad influence” from their spouses, where this influence tends to remove some of the “male performance edge”.

Finally, a significant age effect is noted in Table 6. Older individuals have a significantly higher (lower) probability of having a higher (lower) alpha than younger individuals, at the ten percent (any) significance level.

6. Concluding remarks

Due to changes in the age structure in many industrialized countries, we are currently experiencing a trend where countries tend to shift pension systems from a defined benefit system to a defined contribution plan. These shifts may have far-reaching consequences for retirees. The key issue in moving from a defined benefit type of pension system to a defined contribution plan is to make individuals more responsible for their own pension investments, and to a larger extent than before, let them bear the actual investment risk. This study analyzes consequences for different types of individuals in the recent redesign of the Swedish pension system. Since the redesign in 2000, the Swedish pension system is a partially defined contribution plan, in which individuals set aside 2.5 percent of their annual income at their own discretion, and another 16 percent into a traditional defined benefit type of account. The basic intent is to analyze individuals' asset allocation and performance in the defined contribution part of the pension system.

The analysis is performed in three steps. First, the investment alternatives available to the investors when the partial defined contribution pension system was launched in 2000 are studied. All available investment alternatives (465 mutual funds) are broken down into orthogonal factors using a factor analysis. The results from the factor analysis indicate that 23 latent factors account for more than 90 percent of the total variation among the 465 fund returns. The most important latent factors are identified in terms of real-world actual asset classes. The most important factor represents the world market portfolio, as approximated with e.g. the MSCI world

index. Other important factors represented in the menu of mutual funds include various bond factors and industry-, country-, or sector-based equity factors.

In the second part of the analysis, the individual investment choices are investigated in terms of the latent factors or inherent asset classes from the factor analysis. In doing so, the idea is to identify which factors are most common in the portfolios of different demographic groups of individuals. Since approximately one third of the population have chosen to invest in a "passive" default alternative, rather than making an active choice of mutual funds, or asset classes, the analysis faces a potential selection bias. The procedure according to Heckman (1979) is used to take the selection bias into account. Thus, the individual choice of activity is first modeled with a probit model. Then, the asset allocation choice is modeled within a seemingly unrelated regression (SUR) framework, taking the probability of making an active choice into consideration. The results show that more sophisticated individuals have a higher probability of making an active investment choice. The default alternative consists of a fund, which is almost perfectly correlated with the world market portfolio. For the active individuals, the results show that more sophisticated individuals have significantly lower loadings on the market portfolio and Swedish long- and short- term bonds, and significantly higher loadings on high yield bonds. One interpretation of the results is that more sophisticated individuals have higher risk in their pension plan portfolios than less sophisticated individuals.

In the third and final part of the study, the performance of the individuals' pension plan portfolios is studied. Performance is measured with Jensen's alpha for each individual, which is obtained from a time series regression of individual monthly returns on a set of market indices representing the most important factors from the factor and asset allocation analysis in the first two parts of the study. The results indicate that sophisticated investors are more likely to choose worse performing

funds than non-sophisticated investors. Also, men and older investors are more likely to pick well performing funds.

References

Adler, M. and B., Dumas, 1983, International portfolio choice and corporate finance: a synthesis. Journal of Finance 38, 925-984.

Agnew, J. and P., Balduzzi and A., Sunden, 2003, Portfolio choice and trading in a large 401(k) plan, American Economic Review 93, 193–215.

Agnew, J., 2002, Inefficient choices in 401(k) plan: evidence from individual level data, Working paper.

Alba, J. and A., Chattopadhyay, 1985, Effects of contest and post-category cues on recall of competing brands. Journal of Marketing Research 22, 340-349.

Arrow, K., 1965, Aspects of the Theory of Risk Bearing (Helsinki: Yrjö Jahnsonin Saato)

Asch, S., 1946. Forming impressions of personality. Journal of Abnormal and Social Psychology 41, 258-290.

Barber, B. and T. Odean, 2001, Boys will be boys: gender, overconfidence, and common stock investment, Quarterly Journal of Economics 141, 261-292.

Barberis, N. and A. Shleifer, 2003, Style investing, Journal of Financial Economics 68, 161–200.

Barberis, N., A. Shleifer, and J. Wurgler, 2005, Comovement, Journal of Financial Economics 75, 283–317.

Barberis, N., and M. Huang, 2001, 'Mental Accounting, Loss Aversion, and Individual Stock Returns.' Journal of Finance, 56: 1247-1292.

Baxter, M., and U. J. Jermann, 1997, The international diversification puzzle is worse than you think, American Economic Review 87, 170–180.

Benartzi, S. and R. Thaler, 2001, Naïve diversification strategies in defined contribution saving plans, American Economic Review, 91(1), 79-98.

Benartzi, S., 2001, Excessive extrapolation and the allocation of (401)k accounts to company stock, Journal of Finance 56, 1747–1764.

Berndt, E., B. Hall, R. Hall, and J. Hausmann, 1974, Estimation and inference in non-linear structural models. Annals of Economic and Social Measurement 3, 653-665.

Black, F., 1974. International capital market equilibrium with investments barriers. Journal of Financial Economics 1, 337-352.

Blake, C., E. Elton and M. Gruber, 2005, Participant reaction and the performance of funds offered by 401(k) plans, Working Paper, New York University.

Blake, C., E. Elton and M. Gruber, 2004, The adequacy of investment choices offered by 401(k) plans, Paper presented at the 2004 meeting of the European Finance Association.

Blake, C., E. Elton and M. Gruber, 1993, The performance of bond mutual funds, Journal of Business.

Bodie, Z., 1995, On the Risk of Stock in the Long Run, Financial Analysts Journal, 51, 3, 68-76.

Bodie, Z., R. Merton, and W. Samuelson, 1992, Labor Supply Flexibility and Portfolio Choice in a Lifecycle Model. Journal of Economic Dynamics and Control, 16, 3, 427-449.

Borjas, G., 2003, Welfare reform, labor supply, and health insurance in the immigrant population. Journal of Health Economics 22, 933-958.

Bornstein, R., 1989, Exposure and affect: Overview and meta-analysis of research, 1968-1987. Psychological Bulletin 106, 265-289.

Brown, S. and W. Goetzmann, 1995, Attrition and mutual fund performance. Journal of Finance 50, 679-698.

Brown, S., W. Goetzmann, and S. Ross., 1995, Survival. Journal of Finance, 50, 853-873.

Calvet,L, J. Campbell and P. Sodini., 2006, Down or Out: Assessing The Welfare Costs of Household Investment Mistakes, Harvard Institute of Economic Research Discussion Paper No. 2107 Available at SSRN: http://ssrn.com/abstract=881768

Camerer, C. and M. Weber, 1992, Recent Development in Modeling Preferences: Uncertainty and Ambiguity, Journal of Risk and Uncertainty 5, 325–370.

Camerer, C., 1995, Individual decision making. In: Kagel, John H., and Alvin E. Roth (Eds.), Handbook of Experimental Economics, Princeton University Press, Princeton, NJ, 587–703.

Campbell, B., 1969, The existence of evoked set and determinants of its magnitude in brand choice behavior. Columbia University, unpublished dissertation.

Campbell, J. and L. Viceira, 2002, Strategic Asset Allocation – Portfolio Choice for Long Term Investors, Clarendon Lectures in Economics, Oxford University Press

Carhart, M., 1997, On the persistence of mutual fund performance. Journal of Finance 52, 57-82.

Chan, L., N. Jegadeesh and J. Lakonishok, 1996, Momentum strategies, Journal of Finance 51, 1681-1783.

Choi, J., D. Laibson, B. Madrian and A. Metrick, 2004, Employee's investment decisions about company stock, NBER Working Paper, 10228.

Choi, J., D. Labison, B. Madrian and A. Metrick, 2002, Defined contribution pensions: plan rules, participant decisions, and the path of least resistance, in Poterba, J., Ed.: Tax policy and the economy, MIT Press.

Choi, J., D. Laibson, B. Madrian and A. Metrick, 2001, For better or for worse: default effects and the 401(k) savings behavior, NBER Working Paper, 8651.

Christoffersen, S., 1997, Fee waivers in money market mutual funds, Mimeo.

Cocco, J., F. Gomes and P. Maenhout, 2005, Consumption and Portfolio Choice over the Life Cycle, The Review of Financial Studies, 18(2), 491-530.

Cooper, I. and E. Kaplanis, 1994, Home bias in equity portfolios, inflation hedging, and international capital market equilibrium, Review of Financial Studies 7, 45–60.

Coval, J. and T. Moskowitz, 2001, The geography of investment: informed trading and asset prices, Journal of Political Economy 109, 811–841.

Coval, J. and T. Moscowitz, 1999, Home bias at home: local equity preference in domestic portfolios. Journal of Finance 54, 2045-2073.

Cromwell, H., 1950. The relative effect on audience attitude of the first versus the second argumentative speech of a series. Speech Monographs 17, 105-122.

Cronqvist, H. and R. Thaler, 2004, Design choices in privatized social-security systems: Learning from the Swedish experience, American Economic Review 94:2, 424–428.

Cronqvist, H., 2006, 'Advertising and Portfolio Choice,' Ohio State University, Fisher College of Business Working Paper.

Dahlquist, M., L. Pinkowitz, R., Stulz and R. Williamson, 2002, Corporate governance, investor protection, and the home bias. Working paper, Ohio State University.

DeMarzo, P., R. Kaniel, and I. Kremer, 2004, Diversification as a Public Good: Community Effects in Portfolio Choice, Journal of Finance 59, 1677-1715.

Dhar, R. and N. Zhu, 2002, Up close and personal: An individual level analysis of the disposition effect. Yale International Center for Finance Working paper 02-20.

Dimson, E., P. Marsh and M. Staunton, 2002, Triumph of the Optimists: 101 Years of Global Investment Returns , Princeton University Press.

Dorn, D. and G.Huberman, 2005, Talk and action: what individual investors say and what they do. Review of Finance 9, 437-481.

Duflo, E. and E. Saez, 2002, Participation and investment decisions in a retirement plan: the influence of colleagues' choices", Journal of Public Economics, 85, 121-148.

Duncan, M. and B. Murdock, 2000, Recognition and recall with precuing and postcuing. Journal of Memory and Language 42, 301-313.

Eisenhauer, J., and M. Halek, 1999, Prudence, Risk Aversion, and the Demand for Life Insurance, Applied Economics Letters 6, 239-242.

Engström, S. and A. Westerberg, 2003, Which Individuals Make Active Investment Decisions in the New Swedish Pension System? Journal of Pension Economics and Finance, 2, 3, 225-245.

Fischhoff, B., P. Slovic and S. Lichtenstein, 1977, Knowing with certainty: the appropriateness of extreme confidence. Journal of Experimental Psychology: Human Perception and Performance 3, 552-564.

Fox C., D. Bardolet and Daniel Lieb, 2004, Partition dependence in decision analysis, resource allocation, and consumer choice, Mimeo.

Frankel, L. and B. Solov, 1963, Does recall of an advertisement depend on its position in the magazine? Journal of Advertising Research 2, 28-32.

French, K. and J. Poterba, 1991, Investor diversification and international equity markets. American Economic Review 81, 222-226.

Friedman, M., 1959, The Demand for Money: Some Theoretical and Empirical Results, Journal of Political Economy, 67, 327-351.

Gherig, T., 1993, An information based explanation of the domestic bias in international equity markets. Scandinavian Journal of Economics 95, 97-109.

Glanzer, M. and A. Cunitz, 1966, Two storage mechanisms in free recall. Journal of Verbal Learning and Verbal Behavior 5, 351-360.

Greene, W., 2000, Econometric analysis, Prentice Hall.

Grinblatt, M. and M. Keloharju, 2001, How distance, language, and culture influence stock holdings and trades, Journal of Finance 56, 1053–1073.

Grinblatt, M. and S. Titman, 1992, The persistence of mutual fund performance. Journal of Finance 47, 1977-1984.

Grinblatt, M. and S. Titman, 1989, Mutual fund performance: an analysis of quarterly portfolio holdings, Journal of Business 39, 393-411.

Hanssens, D. and B. Weitx, 1980, The effectiveness of industrial print advertisements across product categories. Journal of Marketing Research 17, 294-306.

Hansson B. and M. Persson, 2000, Expected Utility Maximization and Time Diversification. Working paper, Lund University.

Hauser, J. and B. Wernerfelt, 1990, An evaluation cost model of consideration sets. The Journal of Consumer Research 16, 393-408.

Heath, C. and A. Tversky, 1991, Preference and belief: ambiguity and competence in choice under uncertainty. Journal of Risk and Uncertainty 4, 5—28.

Heckman J., 1979, Sample Selection Bias as a Specification Error, Econometrica, 47, 153-161.

Hellwig, M., 1982, Rational Expectations Equilibrium with Conditioning on Past Prices: A Mean-Variance Example, Journal of Economic Theory, 278—312.

Hellwig, M., 1980, On the Aggregation of Information in Competitive Markets, Journal of Economic Theory, 477—498.

Hendricks, D., J. Patel and R. Zeckhauser, 1993, Hot hands in mutual funds: Short-run persistence of performance. Journal of Finance 48, 93-130.

Hjerm, M., 2005, What the future may bring: Xenophobia among Swedish adolescents. Acta Sociologica 48, 292-307.

Hong, H., J. Kubik and J. Stein, 2004, Social Interaction and Stock Market Participation, Journal of Finance, 59(1), 137–163.

Huberman, G. and P. Sengmuller, 2002, Company stock in 401(k) plans, Working Paper, Columbia University.

Huberman, G. and W. Jiang, 2006, Offering versus choice in 401(k) Plans: Equity exposure and number of funds. Journal of Finance 61, 763-801.

Huberman, G. and W. Jiang, 2004, The 1/N heuristic in 401(k) plans, Working Paper, Columbia University.

Huberman, G., 2001, Familiarity breeds investment, Review of Financial Studies 14, 659–680.

Ibbotson, R. and W. Goetzmann, 1994, Do winners repeat? Journal of Portfolio Management 20, Winter, 9-18.

James, C. and J. Karceski, 2003, Captured money? Differences in the performance characteristics of retail and institutional mutual funds, Working Paper, University of Florida.

Jensen, M., 1968, The performance of mutual funds in the period 1945-1964, Journal of Finance 23, 389-416.

Kacperczyk, M., C. Sialm, and L. Zheng, 2004, On the industry concentration of actively managed equity mutual funds, Journal of Finance forthcoming.

Kahneman, D., 1973, Attention and Effect, Prentice Hall, New Jersey.

Kang, J. and R. Stulz, 1997, An analysis of foreign portfolio equity ownership in Japan, Journal of Financial Economics 46, 3–28.

Karlsson, A. and L. Nordén, 2007, Home sweet home: home bias and international diversification among individual investors, Journal of Banking and Finance 31, 317-333.

Karlsson, A., M. Massa, and A. Simonov, 2006, Portfolio choice and menu exposure. Working paper, Stockholm University School of Business.

Karolyi, A. and R. Stulz, 2002, Are financial assets priced locally or globally? in Constantinides, G., Harris, M., and Stulz, R. (Eds.), Handbook of the Economics of Finance. North-Holland.

Kilka, M. and M. Weber, 2000, Home bias in international stock return expectations. Journal of Psychology and Financial Markets 1, 176--192.

Klayman, J., J. Soll, C. González-Vallejo and S. Barlas, 1999, Overconfidence: it depends on how, what and whom you ask, Organizational Behavior and Human Decision Processes, 79:3, 216-247.

Kritzman, M. and D. Rich, 1998, Beware of Dogma, The Journal of Portfolio Management, 24, 66-77.

Kritzman, M., 1994, What Practitioners need to know… About Time Diversification, Financial Analysts Journal, 50, 1, 14-18.

Latane, H., 1963, Income Velocity and Interest Rates: A Pragmatic Approach, Review of Economics and Statistics, 42, 445-449.

Lee, W., 1990, Diversification and Time, Do Investment Horizons Matter?, The Journal of Portfolio management, 4, 60-69.

Lewis, K., 1999, Trying to explain home bias in equities and consumption. Journal of Economic Literature 37, 571-608.

Levy, H. and M. Sarnat, 1970, International diversification of investment portfolios. American Economic Review 60, 668-675.

Levy, H., 1994, Absolute and Relative Risk Aversion: An Experimental Study, Journal of Risk and Uncertainty 8, 289-307.

Long, J., 1997, Regression models for categorical and limited dependent variables. Advanced Quantitative Techniques in the Social Sciences Series 7, Sage Publications Inc.

Lund, F., 1925, The psychology of belief: A study of its emotional and volitional determinants. Journal of Abnormal and Social Psychology 20, 183-191.

Madrian, B. and D. Shea, 2001, The power of suggestion: inertia in 401(k) participation and savings behavior. Quarterly Journal of Economics 116, 1149-1187.

Massa, M. and A. Simonov, 2006, Hedging, familiarity, and portfolio choice. Review of Financial Studies 19, 633-685.

Meltzer, A., 1963, The Demand for Money: The Evidence From Time Series, Journal of Political Economy, 71, 219-246.

Merton, R., 1987, A Simple Model of Capital Market Equilibrium with Incomplete Information, Journal of Finance, 42, 483-510.

Miller, G., 1956, The magical number seven, plus or minus two: Some limits on our capacity for processing information. The Psychological Review 63, 81-97.

Morgan Stanley Capital International, 2006, Morgan Stanley Capital Indices, www.ms.com.

Mullainathan, S., 2002, Thinking through categories, MIT Mimeo.

Mullainathan, S., and A. Shleifer, 2002, Coarse thinking and persuasion, MIT Mimeo.

Müller-Peters, A., 1998, The significance of national pride and national identity to the attitude toward the single European currency: A Europe-wide comparison. Journal of Economic Psychology 19, 701-719.

Nedungadi, P., 1990, Recall and consumer consideration sets: Influencing choice without altering brand evaluations. Journal of Consumer Research 17, 263-276.

Oberlechner, T. and C. Osler, 2003, Overconfidence in currency markets, Working Paper, Brandeis University.

Pashler, H., J. Johnston, 1998. Attentional limits in dual-task performance, in Pashler, H., (Ed), Attention, Psychology Press.

Peng, L. and W. Xiong, 2002, Time to Digest and Volatility Dynamics, Working Paper.

Poterba, J., J. Rauh, S. Venti, and D. Wise, 2003, Utility Evaluation of Risk in Retirement Saving Accounts, NBER working paper series, working paper 9892.

Pratt, J., 1964, Risk Aversion in the Small and in the Large, Econometrica, 32, 122-136.

Premiepensionsmyndigheten, 2000, Fondkatalog för din premiepension.

Rosnow, R., 1966, What ever happened to the "Law of Primacy." Journal of Communication 16, 10-31.

Rydgren, J., 2004, The logic of xenophobia. Rationality and Society 16, 123-148.

Samuelson, P., 1994, The long term case for equities and hoe it can be oversold, journal aof portfolio management, 21, issue 1, 15-24.

Samuelson, P., 1989, The judgment of economic science on rational portfolio management: Indexing, timing, and long horizon effects, journal of portfalio management 16, 4-12.

Samuelson, P., 1963, Risk and Uncertainty: A Fallacy of Large Numbers, Scientia, 98, 108-113.

Selden, R., 1956, Monetary Velocity in the United States, in Studies in the Quantity Theory of Money, M. Friedman (ed.), Univeristy of Chicago Press, Chicago, Ill.

Sharma, S., 1996, Applied multivariate techniques, John Wiley & Sons, Inc.

Sharma, S., T. Shimp and J. Shin, 1995, Consumer Ethnocentrism: A test of antecedents and moderators. Journal of the Academy of Marketing Science 23, 26-37.

Sharpe, W., 1966, Mutual fund performance, Journal of Business 39, 119-138.

Shiller, R., 2003, From efficient markets theory to behavioral finance, Journal of Economic Perspectives, 17:1, 83-104.

Siegel, F. W., and J. Hoban, 1982, Relative Risk Aversion Revisited, Review of Economics and Statistics 64, 481-487.

Siegel, J., 1994, Stocks for the Long Run. Irwin

Simon, H., 1955, A behavioral model of rational choice. Quarterly Journal of Economics 69, 99-118.

Sims, C., 2000, Implications of Rational Inattention, Working Paper, Princeton University.

Sirri, E. and P. Tufano, 1998, Costly search and mutual fund flows. Journal of Finance 53, 1589-1622.

Smith, T. and S. Kim, 2006, World opinion: National pride in comparative perspective. International Journal of Public Opinion Research 18, 127-136.

Strong, N. and X. Xu, 2003, Understanding the equity home bias: evidence from survey data, Review of Economics and Statistics 85, 307–312.

Stulz, R., 1981, On the effects of barriers to international investment. Journal of Finance 36, 923--934.

Sunden, A., 2000, How Will Sweden's New Pension System Work? Issue Brief No. 3, Chestnut Hill, MA: Boston College, Center for Retirement Research.

Sundin, A. and S. I. Sundqvist, 2002, Owners and Power in Sweden's Listed Companies, SIS Ägarservice AB, 1986–2002.

Szpiro, G., 1986, Measuring Risk Aversion: An Alternative Approach, Review of Economics and Statistics 68, 156-159.

Säve-Söderbergh. J., 2003, Gender, risk and portfolio choices of retirement wealth, Working Paper, The Swedish Institute for Social Research, Stockholm University.

Tajfel, H. and J. Turner, 1979, An integrative theory of intergroup conflict. In: Austin, G., Worchel, S. (Eds). The social psychology of intergroup relations. Wadsworth, Monterey, pp. 33-47.

Tajfel, H., 1979, Individuals and groups in social psychology. British Journal of Social and Clinical Psychology 18, 183-190.

Tajfel, H., 1978, Differentiation between social groups. Academic Press, London.

Teo, M. and S-J. Woo, 2003, Style effects, Mimeo.

Tesar, L. and I. Werner, 1995, Home bias and high turnover, Journal of International Money and Finance 14, 467–92.

Theil, H., 1969, A multinomial extension of the linear logit model, International Economic Review 10, 251-259.

Thorly, S., 1995, The Time-Diversification Controversy, Financial Analysts Journal, 51, 3, 18-22.

Tversky, A. and D. Kahneman, 1974, Judgment under uncertainty: Heuristics and biases/ Science 185, 1124-1131.

Van de Ven, W. and B. Van Praag, 1981, The Demand for Deductibles in Private Health Insurance: A Probit Model with Sample Selection. Journal of econometrics, 17: 229-252.

Van Nieuwerburh, S. and L. Veldkamp, 2004, Information immobility and the home bias puzzle, Mimeo.

Wermers, R., 2000, The value of active mutual fund management, Journal of Financial and Quantitative Analysis.

White, H., 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817-838.

Vissing-Jørgensen, A., 2003, Perspectives on behavioral finance: does "irrationality" disappear with wealth? Evidence from expectations and actions, Working Paper, Northwestern University.

Wright, P., 1975, Consumer choice strategies: Simplifying vs. optimizing. Journal of Marketing Research 12, 60-67.

Zajonc, R., 1968, Attitudinal effects of mere exposure. Journal of Personality and Social Psychology, Monograph Supplement 9, 1-27.

Zellner, A., 1962, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, Journal of American Statistical Association 57, 348-368.