Netspar Network for Studies on Pensions, Aging and Retirement

Arie Kapteyn

Federica Teppa

Subjective Measures of Risk Aversion, Fixed Costs, and Portfolio Choice

Discussion Paper 07/2009 - 024

July 15, 2009

Subjective Measures of Risk Aversion, Fixed Costs, and Portfolio Choice

Arie Kapteyn
RAND
Federica Teppa
De Nederlandsche Bank (DNB) and Netspar

July 15, 2009

Abstract

The paper investigates risk attitudes among different types of individuals. We use several different measures of risk attitudes, including questions on choices between uncertain income streams suggested by Barsky *et al.* (1997) and a number of *ad hoc* measures. As in Barsky *et al.* (1997) and Arrondel and Calvo-Pardo (2002), we first analyse individual variation in the risk aversion measures and explain them by background characteristics (both "objective" characteristics and other subjective measures of risk preference). Next we incorporate the measured risk attitudes into a household portfolio allocation model, which explains portfolio shares, while accounting for incomplete portfolios and fixed costs. Our results show that a measure based on factor analysis of answers to a number of simple risk preference questions has the most explanatory power. The Barsky *et al.* (1997) measure has less explanatory power than this "a-theoretical" measure. We provide a discussion of the reasons for this finding. Fixed costs turn out to provide an economically and statistically highly significant explanation for incomplete portfolios.

Jel–Classification: C5; C9; D12; G11

Keywords: Risk Aversion; Portfolio Choice; Subjective Measures; Econometric Models, Fixed Costs

Aknowlegments: We thank Maarten van Rooij for providing us with the data on risky choices for 2004, as well as seminar and conference participants at various institutions for their many useful comments, also on an earlier version of this paper. In this paper, use is made of data of the DNB Household Survey; we thank CentERdata at Tilburg University for supplying them. The views expressed in this paper are those of the authors and do not necessarily reflect those of the institutions they belong to. Any remaining errors are our own responsibility.

1 Introduction

This paper exploits direct measures of risk preferences in a model of household portfolio allocation. There are two main motivations for this. The first one is that if heterogeneity in risk preferences is important then empirical portfolio models should take this into account. The second motivation is that economic theory has a fair amount to say about how risk preferences should influence portfolio allocation. Having direct measures of risk preferences should therefore help us in better testing the validity or predictive power of economic theories of portfolio allocation.

Empirical analyses of portfolio choice of households or individuals appear to indicate that observed choices are often inconsistent with standard asset allocation models. As a consequence, several studies have focused on empirical failures of portfolio theory. The greatest failure is perhaps the fact that the majority of individuals do not hold fully diversified portfolios (Campbell, 2006), although the percentage of households holding risky assets has increased over the last decade (Haliassos and Hassapis, 2002). Potential explanations for the fact that many households do not hold stocks may be the costs of stock market participation (Vissing-Jorgensen, 2002) and peer effects (Hong *et al.*, 2004; Christelis and Georgarakos, 2008). A more recent literature stresses the role of financial literacy: unsophisticated individuals tend to avoid taking direct action over their retirement saving investment decisions (Benartzi and Thaler, 2002; Thaler and Benartzi, 2004; van Rooij *et al.*, 2007).

The sub-optimal degree of international diversification known as "home asset bias" is potentially another empirical failure. Although the ongoing globalization and international integration processes are contributing to a lowering of its magnitude (Baele *et al.*, 2007), this phenomenon has been analyzed extensively (see among others French and Poterba, 1990, 1991; Cooper and Kaplanis, 1994; Glassman and Riddick, 2001; Jermann, 2002). Possible reasons for the over-investment in domestic assets have been identified in different transaction costs between countries (Tesar and Werner, 1994, 1995; Amadi and Bergin, 2006), additional sources of risk for foreign investments (La Porta *et al.*, 1999; Dahlquist *et al.*, 2003; Stulz, 2005), informational costs and asymmetries (Ahearne *et al.*, 2004; Choe *et al.*, 2005; Dvorak, 2005), transparency in international markets (Gelon and Wei, 2005), trust (Guiso *et al.*, 2008), real exchange rate volatility (Fidora *et al.*, 2007), behavioral biases (Strong and Xu, 2003), and investors' financial illiteracy (Graham *et al.*, 2005).

A more fundamental piece of evidence against the rational model of portfolio allocation is provided by Benartzi and Thaler (2001) who find that the allocation of investors is heavily dependent upon the choices offered to them. Roughly speaking, if they are offered n choices they tend to allocate $\frac{1}{n}$ of their investment to each of the choices offered, irrespective of the risk characteristics of the investment opportunities.

Although these findings suggest that the rational model of choice is unable to explain several empirical phenomena, it is often hard to determine in more detail what the underlying cause of disparities between theory and empirical facts may be. The connection between theory and empirical evidence is often tenuous, because too many intervening factors may explain why theoretical predictions are not borne out by data. For this reason some authors have turned to more direct, subjective evidence on preferences to reduce the distance between theory and empirical facts. A prominent example is the paper by Barsky *et al.* (1997) who elicit several pieces of subjective information to improve our understanding of intertemporal choice and portfolio allocation by using US data. More recently, Iezzi (2008) uses a risk aversion measure coming from a subjective question in the Bank of Italy households survey, and explicitly accounts for potential misclassification between the true and the observed risk aversion.

In this paper we also aim to exploit subjective information to construct empirical micro-models of portfolio choice. In contrast with the work by Barsky *et al.* (1997), and Arrondel and Calvo-Pardo (2002) who adapt the same methodology to French data, our model will be a formal structural model of portfolio choice, in which we consider several different measures of risk attitude. One measure is based on hypothetical choices between uncertain income streams in a Dutch household survey, and closely related to the aforementioned work by Barsky *et al.* (1997), and Arrondel and Calvo-Pardo (2002). The Barsky *et al.* (1997) measure has a nice direct interpretation if individuals have CRRA preferences. We will find however, that the measure also has theoretical and empirical problems. Hence we also consider alternative measures of risk attitude. We relate the different measured risk attitudes to observed portfolio choices of households. To deal with incomplete portfolios, we allow for fixed costs of ownership of certain types of assets, which can endogenously generate corner solutions in portfolio allocation. Thus, we formulate and estimate a utility consistent choice model, incorporating subjective measures of risk aversion. The model is closely related to rational portfolio theory and seems to do a reasonable job in describing differences in allocation across individuals who differ in socio-economic characteristics, wealth, and risk attitudes.

The paper is organized as follows. In the next section we describe the data we use in the empirical analysis. In Section 3 we present descriptive statistics on the various risk attitude measures and how they are related. Section 4 describes the portfolio composition of households. Section 5 formulates a simple static asset allocation model with fixed costs. Section 6 presents the empirical results. Section 7 concludes.

2 Data on risk aversion and precaution and how they were collected

The data used in this paper have been collected from the households in the so-called CentERpanel. The CentERpanel is representative of the Dutch population, comprising some 2000 households in the Netherlands. The members of those households answer a questionnaire at their home computers every week. These computers may either be their own or a set-top box provided by CentERdata, the agency running the panel.[1] In the weekend of June 11-15 of 2004 a questionnaire was fielded with a large number of subjective questions on hypothetical choices. The questionnaire was repeated in the weekends of July 9-13 and August 20-24 of 2004 for those panel members who had not responded yet. For this paper we exploit the section involving choices over uncertain lifetime incomes. We merge these data with data from the DNB Household Survey (DHS), previously known as CentER Savings Survey. The DHS collects information on assets, liabilities, demographics, work, housing, mortgages, health and income, and many subjective variables (e.g. expectations, savings motives) from annual interviews with participants in the CentERpanel. Typically the questions for the DHS are asked in May of each year, during a number of consecutive weekends.[2]

In the next subsections we discuss three measures of risk aversion elicited from the respondents in the sample.

2.1 Choices of uncertain lifetime income

Our first measure is based on a number of questions involving *choices over lifetime incomes*. This methodology, taken from Barsky *et al.* (1997), allows us to rank

[1] A set-top box can be hooked up to a TV set and a telephone line to establish Internet access.

[2] More information on CentERdata, the CentERpanel and the DHS is available at their website (http://www.uvt.nl/centerdata/dhs).

individuals with respect to their risk aversion without having to assume a particular functional form for the utility function.

In the Barsky *et al.* (1997) experiment, questions are posed to all respondents in the Health and Retirement Study, consisting of individuals aged over 50. Arrondel and Calvo-Pardo (2002) asked the same questions to a representative sample of French households. In our case, the questions are asked to people who have a paid job (with the exclusion of the self-employed), look for a job either for the first time or after having lost one, and to students.

The structure of the questions is depicted in Figure 1. In the first round, respondents are asked the following question:

Suppose your medical doctor advises you to move house because of an allergy. You follow his advice and it turns out you have to choose between two jobs. Both jobs are comparable in terms of working hours, but one job is more secure than the other.

In the first job there is a guarantee that you will have the same income for the rest of your life. You know that your net income per month will be Y for as long as you work.

In the second job you may earn more, but your income is less secure as well. There is a 50% chance that your income will be twice as high (2xY). But there is also a 50% chance that you will earn much less and that your income will be 0.7xY.

In both cases it is assumed that the income of other household members (if any) remain the same. Also assume that there is no inflation. In other words: in the future the value of the euro is the same as it is now.

Which of the previous mentioned jobs would you prefer?

1 job1: the job that guarantees your income of Y.

2 job2: the job in which there is a 50% chance that you will earn 2xY, but also a 50% chance that it will be $0,7$xY for the rest of your life.

Various quantities in the question vary per respondent, exploiting the computerized nature of the interviews. The quantity Y is the respondent's own selfreported net personal income. 2xY is twice the respondent's income; 0.7xY is personal income times 0.7, etc. This is in contrast to the experiments by Barsky *et al.* (1997) and Arrondel and Calvo-Pardo (2002), in which the incomes were the same for all individuals. Obviously, the question involves a choice between a certain and an uncertain outcome: the former is given by the actual income the respondent receives (Y), the latter is a 50-50 gamble over a good outcome (2xY) and a bad outcome ($0,7$xY).

In the second round each individual is asked a similar question. If she has chosen the certain outcome (Y) in the first round, she now faces another choice where the risky outcome is more attractive. The 50-50 gamble now involves 2xY and 0.8xY (0.8 times income). If she has chosen the risky prospect in the first round, she is now asked to choose between her income for sure and a less attractive gamble, i.e. 50% chance of 2xY and 50% chance of 0.5xY.

Similarly, in the third round the gamble becomes more attractive for those respondents who have once again chosen a certain income stream in the second round (the 50-50 gamble now involves 2xY and 0.9xY), and less attractive for those respondents who preferred the risky choice (the 50-50 gamble now involves 2xY and 0.25xY).

The answers to the questions allow us to identify six groups ranked from most risk averse to least risk averse (or equivalently from least risk tolerant to most risk tolerant; we will generally denote the variable defined by the six classes as "risk tolerance"). Both the Barsky *et al.* (1997) and Arrondel and Calvo-Pardo (2002)'s study involve only two rounds of questions rather than three as ours. For comparison we temporarily combine the two most extreme groups into one. Thus we have four categories of individuals, from I to IV, where the I-group is the union of the 1 and the 2 groups and the IV-group is the union of the 5 and 6 groups. We can then compare the risk tolerance across the three studies. Table 1 gives the results. To facilitate a comparison with the Barsky *et al.* (1997) study we split our sample in two age groups: 50 and younger and over 50.[3]

A comparison between France and the Netherlands on the basis of the complete age range suggests that there is a greater spread of risk aversion in the Netherlands than in France. The Dutch respondents are more heavily represented in the two extreme categories (almost 55% of the Dutch belong to the most risk averse group compared to 43% of the French, whereas 8% of the Dutch belong to the least risk averse group compared to 6% for the French). Summing the percentages of the first two groups and the percentages of the last two groups respectively, suggests that the Dutch are less risk averse than the French (only 74.3% of the Dutch belong to the first two groups compared to 82.5% of the French, whereas 27.5% of the Dutch belong to the last two groups compared to 17.5% of the French).

Considering the sub-samples of respondents over 50, it appears that the Americans may be slightly more risk tolerant than the Dutch. Compared to the Dutch and the Americans, the French appear to be most risk averse.

[3] The Barsky *et al.* (1997) sample consists of respondents over 50.

2.2 Risk attitude measures based on principal components

The DHS-questionnaire contains six direct questions about *investment strategies*. They are reported in the top panel of Table 2. Respondents can express their agreement or disagreement with these statements on a seven point scale (1 means "complete disagreement" and 7 means "complete agreement"). In addition the DHS also contains 16 questions about *savings motives*. The bottom panel of Table 2 reproduces two of them that are related to precautionary motives and uncertainty. As previously, answers can be given on a 1 to 7 scale, where 1 means "very unimportant" and 7 means "very important".

We apply a factor analysis to our sample ($n = 1220$) in order to determine the factor structure of these eight indicators of risk aversion and precaution. The factor analysis is a principal components analysis (PCA), with varimax rotation, where the extraction method is based on eigenvalues greater than 1. It turns out that three underlying factors can explain most of the variance in the answers.[4] The factor loadings are presented in Table 2. The largest factor loadings in each column are highlighted in bold face. We mark by "R" the statements whose score has been reversed to make all the statements consistent. We note that "Factor 1" and "Factor 2" mainly explain the six investment strategies' variables, whereas "Factor 3" mainly explains the precautionary motives and uncertainty variables. In view of the wordings of the eight items, we interpret the first two factors as measures of risk aversion and the third factor as a measure of precaution. We then refer to the three extracted factors as "Riskav1", "Riskav2" and "Precaution", respectively. In general, we would expect "Precaution" to play a role in savings decisions whereas "Riskav1" and "Riskav2" should affect portfolio choice.

2.3 Direct questions on precaution and risk aversion

The third type of subjective measures are derived from an alternative set of questions. As before, the respondents can express their agreement or disagreement with a number of statements on a seven point scale (1 means "complete disagreement" and 7 means "complete agreement"). For our purposes, we select the following ones:

Care1 - Being careful with money is an important character trait

Care2 - I am careless with money.

[4] The cumulative proportion explained by the three retained factors is 66.9 percent.

3 Analysis of the subjective risk aversion measures

Table 3 presents a number of descriptive statistics of the three alternative measures of risk aversion by several demographic and socio-economic characteristics, namely gender, education level and employment status (where available). Panel [A] refers to the risk aversion measure derived from the risky choice over lifetime income questions described in Section 2.1. Note that from now on we once again distinguish six classes of risk tolerance. Moreover, for the purpose of this table, the risk tolerance has been coded from 1 (least risk tolerant) to 6 (most risk tolerant). Panel [B] combines the risk aversion measures stemming from principal components factor analysis explained in Section 2.2. Panel [C] consists of the two direct questions on precaution and risk aversion listed in Section 2.3. As before, we reverse the score in statement "Care2" in order to make it consistent with the other statement.

The *p*-values reported in the table refer to one way analysis of variance of each of the risk attitude measures on the characteristics considered.

In general, for all the three sets of variables the vast majority of the respondents fall in the most risk averse categories. The table suggests that females are slightly more risk averse than males, and have a stronger precautionary motive. The difference in means is (strongly) significant for all the variables but Care1, and rather small. Moreover, better educated individuals are generally less risk averse: the differences in risk tolerance between the three levels of education are statistically significant. However, considering Riskav1 and Riskav2, education has a significant effect for Riskav1 only. The small number of observations for the self employed does not allow for statistically significant differences by employment status.

3.1 How are the subjective measures related?

A simple way to assess how informative the different subjective measures of risk aversion are consists of examining their intercorrelations (Table 4). The first thing to bear in mind is that the three measures extracted from factor analysis (Riskav1, Riskav2 and Precaution) are orthogonal to one another by construction. Risk tolerance is negatively and significantly (at the 1-percent level) correlated with all the other measures but with Precaution (correlation is positive and insignificant) and Care1 (negative and insignificant). These findings are as expected. First, Risk tolerance, Riskav1 and Riskav2 capture risk aversion more directly than any of the other subjective measures: both the high value of the correlation coefficients and the significance level may reflect this. Second, the negative sign is in line with the information conveyed by the measures. Precaution is positively and significantly correlated with both Care1 and Care2.

3.2 Determinants of the subjective risk aversion measures

We consider a number of variables that can potentially affect the risk aversion measures. A full description of their definition is given in Table A1 in the appendix. In addition to the usual background characteristics, like gender, age, education level, we control for both household income and household total financial wealth. Table 5 reports the estimation results from ordered probit and OLS regressions. Females are significantly more risk averse, cautious and careful than males. This is consistent with our priors and with the literature on risk preferences (e.g. Levin *et al.*, 1988; Barsky *et al.*, 1997; Iezzi, 2008). Age is strongly and positively related to both risk aversion and to being careful and cautious. Education mainly plays a significant role for the risk tolerance measure: higher educated individuals are more likely to be risk tolerant. Household income turns out to be totally insignificant in all regressions. Household total financial wealth is negatively related to the first risk aversion measure, but positively affects the level of self-reported carefulness and precaution.

4 Assets and liabilities

The DHS collects extensive information on assets and liabilities. Respondents are asked for ownership and quantity of different categories of assets, both real and financial, and of liabilities and mortgages. Table A2 in the appendix reports data on financial assets at the household level. We group assets in categories that are somewhat homogeneous with respect to their risk profile. We have indicated the category an asset belongs to by *NR* for non-risky assets, by *R* for risky assets and by *O* for other assets. Non-risky assets are checking accounts, savings accounts, deposits, and insurances. Risky assets are defined as the sum of growth and mutual funds, options, stocks and business equity. Other assets are the sum of real estate, mortgage, bonds, money lent out and financial debt. Obviously, the other assets also carry a certain amount of risk. The table presents relative frequencies of ownership, mean values and shares of each asset category in the total portfolio.

Checking accounts, savings accounts, deposits, and insurances are held by 98.5% of the sample. This safe asset makes up more than 60% of average financial wealth in the sample. Business equity is another sizeable component of financial wealth, although it is only held by 4.7% of the sample. Generally, the households in the sample use very little credit: Financial debts amount to slightly more than 10% of total financial assets (and hence financial wealth is close to 90% of total financial assets). Only 5% of the sample holds bonds and/or mortgage bonds, whereas 16% holds stocks; 21.7% of the sample holds growth or mutual funds. Of course the group of stock holders overlaps with the owners of mutual funds or growth funds. 30.4% of the sample households have stocks and/or growth and mutual funds. Clearly, for most households real estate (usually the primary residence) dominates the portfolio. Financial assets are only 23% of total assets and financial wealth is only 20% of total wealth. At the aggregate level, non-risky assets are basically 60% of average financial wealth whereas risky assets represents only 20% of mean financial wealth. Moreover the percentage of people owning risky assets (39.1) is less than half of that of people owning safe assets (98.5). Other assets are widely owned in the sample (81.7%).

Tables 6 and 7 provide a first impression of the explanatory power of the subjective risk aversion measures in determining ownership of risky assets and their shares in total financial wealth. Table 6 shows that both the Barsky *et al.* (1997) risk tolerance measure and the psychometric indexes Riskav1 and Riskav2 contribute significantly to an explanation of ownership of risky assets. The *ad hoc* measures Care1 and Care2 are not statistically significant. Table 7 provides a similar picture with respect to an explanation of shares in total financial wealth, although now also Care2 is significant. The samples used for the three different risk aversion measures differ in size so that a comparison may be contaminated by differences in sample. Table 8 provides goodness of fit measures based on the sample for which all measures are available. This within-sample comparison favors the psychometric indexes, with the Barsky *et al.* (1997) risk tolerance measure and the *ad hoc* measures producing virtually identical goodness of fit.

5 Classical theory of portfolio choice

Our empirical model follows a simple standard model of portfolio choice (see e.g. Gollier, 2004). A consumer's choice problem is to select a vector of risky assets that maximizes expected utility. We provide the derivation of the optimal portfolio for this case. Consider a k-vector of assets, where the first asset is riskless and the remaining $(k-1)$ assets are risky. Let μ be the $(k-1)$-vector of mean excess returns of the risky assets and Σ the variance covariance matrix of the excess returns. Let W be begin of period wealth, r is the riskfree interest rate. The $(k-1)$-vector α

denotes the quantities invested in the risky assets, with stochastic returns given by the vector $\widetilde{x}_0$. Let ι be a $(k-1)$-vector of ones. Then $\iota'\alpha$ is the amount of money invested in the risky assets and $W - \iota'\alpha$ is the amount invested in the riskfree asset (No non-negativity restrictions are imposed). Consumption z is equal to the value of the assets at the end of the period. Thus consumption is:

$$z = (W - \iota'\alpha)(1+r) + \alpha'(\iota + \widetilde{x}_0) = W(1+r) + \alpha'(\widetilde{x}_0 - r) \equiv w_0 + \alpha'\widetilde{x} \qquad (1)$$

where $w_0 = W(1+r)$ and $\widetilde{x} = \widetilde{x}_o - r.\iota$. We assume $\widetilde{x}$ to be normally distributed, so that $\widetilde{x} \sim N(\mu, \Sigma)$. The consumer wants to maximize the expectation of end of period utility subject to (1) by choosing α optimally. We assume a utility function of consumption characterized by constant absolute risk aversion; i.e. the utility of a consumption level z is equal to $u(z) = -\frac{1}{A}\exp(-Az)$, where the parameter A is the coefficient of absolute risk aversion. Inserting (1) in the utility function, neglecting the multiplicative constant A, and taking expectations yields

$$\begin{aligned} V(\alpha) &= -(2\pi)^{-n/2}\left|\Sigma\right|^{-1/2}\int \exp(-A(w_0 + \alpha' x))\exp(-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu))dx \\ &= \exp(-Aw_0 - A\alpha'\mu + \frac{1}{2}A^2\alpha'\Sigma\alpha).(2\pi)^{-n/2}\left|\Sigma\right|^{-1/2} \\ &\quad .\int \exp[-\frac{1}{2}(x-\mu+A\Sigma\alpha)'\Sigma^{-1}(x-\mu+A\Sigma\alpha)]dx \\ &= -\exp(-Aw_0 - A\alpha'\mu + \frac{1}{2}A^2\alpha'\Sigma\alpha) \qquad (2) \end{aligned}$$

Maximizing (2) with respect to α yields:

$$\alpha^* = \frac{1}{A}\Sigma^{-1}\mu \qquad (3)$$

5.1 An econometric model of portfolio choice with fixed costs

As mentioned in the introduction, transaction costs (e.g. brokerage fees, bid-ask spreads, taxes, fund loads, or information costs) have been considered as an explanation of both non-fully diversified portfolios and reduced portfolio size, at least for small investors.[5] In the literature a distinction is made between linear (or variable) and nonlinear (or fixed) transaction costs, as convex optimization methods can solve asset allocation problems in presence of linear costs only. Fixed costs have then been incorporated in portfolio selection models in indirect ways. For example, Mao (1970), Jacob (1974) and Levy (1978) place restrictions on the number of securities

[5] It has been argued that for large investors the fixed costs are trivial (see e.g. Brennan, 1975).

in a model of optimal portfolios. Patel and Subrahmanyam (1982), and Blog *et al.* (1983) place restrictions on the variance covariance matrix of asset returns: the former assume that the pairwise correlation coefficient is the same across all securities, the latter assume a diagonal plus rank-one covariance matrix. More recently, Lobo *et al.* (2007) develop a relaxation method yielding a computable upper bound via convex optimization when separable fixed transaction costs are considered.

Our interest will be in ownership and portfolio shares of a number of asset categories that vary in riskiness. We want to allow for other factors determining portfolio composition than just the distribution of excess returns. To introduce these other factors in a utility consistent way, we replace (2) by

$$V^*(\alpha) = -\exp(-Aw_0 - A\alpha'\mu - A^2 w_0 \alpha'\Sigma z + \frac{1}{2}A^2\alpha'\Sigma\alpha) \tag{4}$$

where z is a vector of taste shifters:

$$z = Bx + \varepsilon \tag{5}$$

where x is a vector of individual (or household) characteristics, B is a parameter matrix, and ϵ an i.i.d. error term. We will interpret ϵ as representing unobservable variations in taste across individuals.

The expression in the exponent of (4) can be written as

$$-Aw_0 - \alpha'(A\mu + A^2 w_0 \Sigma z) + \frac{1}{2}\alpha' A^2 \Sigma\alpha = -Aw_0 - \alpha' b + \frac{1}{2}\alpha'\Psi\alpha \tag{6}$$

where $\Psi \equiv A^2\Sigma$ and $b \equiv A\mu + A^2 w_0 \Sigma z$. This can be written as

$$-Aw_0 - \frac{1}{2}b'\Psi^{-1}b + \frac{1}{2}(\alpha - \Psi^{-1}b)'\Psi(\alpha - \Psi^{-1}b) \tag{7}$$

which is minimal for

$$\widetilde{\alpha} = \Psi^{-1}b = \frac{1}{A}\Sigma^{-1}\mu + w_0 z \tag{8}$$

which is a generalization of (3).

We can insert (8) into the formula for the utility function (4) to obtain the optimal value of the utility function. In principle we can insert any value of the vector α in the utility function and compute its utility value. If portfolios are incomplete then some elements of α will be zero. The utility function can be used to compare the value of different incomplete portfolios and thus we can determine which incomplete portfolio would yield the highest utility. As noted before, a major reason why portfolios are incomplete may be fixed costs. A straightforward interpretation of fixed costs is that it acts like a reduction of w_0. So consider a particular portfolio combination i. In our case i takes on four values: $i = 1$: a household only has safe assets; $i = 2$: a household has safe assets and risky assets; $i = 3$: a household has safe assets and other assets; $i = 4$: a household has all types of assets (safe, risky, and other). Let the corresponding fixed costs of ownership of a particular asset combination be f_i, then we replace w_0 in the utility function by $w_0 - f_i$. By comparing the value of the utility function for each combination i, an individual then can choose the asset combination with the highest utility. Appendix A contains the derivation of the utility for each case and also discusses identification of parameters. Here we just note that by considering the utility of the various asset combinations, we can estimate the parameters in the matrix B in equation (5). Furthermore we can obtain estimates of the quantities $\kappa_1 \equiv \frac{\mu_1}{\sigma_1}$ and $\kappa_2 \equiv \frac{\mu_2}{\sigma_2}$, where σ_1^2 and σ_2^2 are the diagonal elements of Σ. In other words we can interpret κ_1 and κ_2 as standardized (i.e. divided by the standard deviation) expected excess returns of respectively the risky asset and the other asset.

6 Empirical results

In this section we present estimation results for the model explaining the different regimes, for the three different risk aversion measures introduced in Section 2. Since the scale of the measures is not a priori defined (perhaps with the exception of the Barsky *et al.* (1997) measure) and to guarantee that the risk aversion measures are positive for all observations, we adopt the following general specifications for the relative risk aversion measure $R = Aw_0$. The general specification is of the form:

$$R = (1+\gamma)\frac{1-\exp(\Theta)}{1+\exp(\Theta)} \tag{9}$$

where in the case of the Barsky *et al.* (1997) measure

$$\Theta = \exp\left[\beta(1+risk)\right] \tag{10}$$

while for Riskav1 and Riskav2 we specify

$$\Theta = \exp\left[\beta\{1+\lambda riskav1 + (1-\lambda)riskav2\}\right] \tag{11}$$

and for Care1 and Care2 we define

$$\Theta = \exp\left[\beta\{1+\lambda care1 + (1-\lambda)care2\}\right] \tag{12}$$

Moreover, we have normalized the variables Risk aversion, Riskav1, Riskav2, Care1 and Care2 so that they all have mean zero and standard deviation equal to one. Together β and γ determine the variation of the risk aversion measures. Clearly $(1+\gamma)$ determines the range of values that the risk aversion measures can take, while β influences the distribution of the risk aversion measure within that range. The parameter λ weights the components of the risk aversion measures for the factor analysis case and the case where we use the direct measures Care1 and Care2.

Table 9 presents the estimation results for the three different risk aversion measures. Log-income, gender and age are used as taste shifters. The fixed costs for each case are made a function of education level, to allow for the fact that possibly access to the markets for risky or other assets are easier for individuals with a higher education. These individuals may already have more information, or they may be less daunted by the complexities of some of the financial products. To avoid situations where the fixed costs are higher than own wealth, we have have parameterized fixed costs as follows:

$$f_i = \min(w_0, C_i) \tag{13}$$

where $i = F, S, B$, as before. C_i is the full fixed cost of owning asset portfolio i. So, if someone's wealth is less than C_i, then we set $w_0 - C_i = 0$. This will still effectively prevent an individual from acquiring portfolio i. Finally we note that during the Maximum Likelihood estimation, the variance of ϵ_i (cf. 5) converged to zero. So no estimate for the variance-covariance matrix of ϵ_i is presented.

The number of observations varies across specifications. Since the Barsky *et al.* (1997) questions are only asked to respondents who currently have a job, the number of observations is particularly low for the specification using this measure. Furthermore, it turned out that the effect of the age dummy for 65+ for the second taste shifter was not identified, due to insufficient observations in this age category that either had portfolio S or B.

We first concentrate on the results for the direct measures and the factor analytic measures (see columns 3 and 2, respectively). The Barsky *et al.* (1997) results will be discussed below. Recall from equation (5) that *ceteris paribus* a larger value of $x_1'\beta_1$ implies a greater budget share for the risky asset; similarly if $x_2'\beta_2$ is larger that indicates a taste for a bigger share of the other asset; if both $x_1'\beta_1$ and $x_2'\beta_2$ are large then there is a preference for both the risky and the other asset (also see (33), (34), and (29)). Concerning these taste shifters, we note that gender has a minor effect. For x_1 the effect is essentially zero while for x_2 the effect of being female is slightly negative; that is, females are less likely to hold other assets or both the risky and other assets. Not surprisingly, the likelihood of owning risky or other assets appears to increase with age. Income has a highly significant effect on owning the other asset or both the risky and other assets. Its effect on only owning the risky asset is negligible. The paramaters κ_1 and κ_2 are not very accurately estimated, so we refrain from an interpretation.

The value of λ implies that the effects of Riskav1 and Riskav2 are very similar; we also note however that Care2 ("I am careless with money") is more important than Care1 ("Being careful with money is an important character trait"). Turning to the effect of fixed costs, we see that these are highly significant, with relatively minor roles for education. Having a higher education reduces the cost of having other assets (or both other and risky assets) somewhat. This finding is in line with Guiso and Jappelli (2008) who find a positive association between portfolio diversification and level of education. The estimate of ρ is not significantly different from zero, suggesting only weak correlations in the returns of the risky and other assets, at least as perceived by individuals.

Let's now consider the estimates using the Barsky *et al.* (1997) measure. The estimates for this specification are quite different from the estimates for the other two specifications, and not many estimates are statistically significant. Partly this may reflect the limited number of observations. We observe very high fixed costs and almost equally high negative effects of education. This may not tell us much more than that not many households in the sample with a low education are observed to own other assets or both the risky and other assets. In fact, the maximum likelihood indicated a singular Hessian. This is reflected in the fact that no t-value has been computed for the effect of higher education on fixed costs. Indeed, the t-values associated with the fixed costs for owning both the risky and the other assets suggest numerical problems, so we should take these estimates with precaution.

7 Conclusions

We have explored the explanatory power of a number of different subjective measures of risk aversion for the explanation of portfolio choice. The variable Risk tolerance, which has the firmest grounding in economic theory, appears to have little explanatory power. There are a few different possible explanations for this. First of all, the question is quite complicated and many respondents may have a hard time understanding the exact meaning of the question. Secondly, the question conditions on a respondent's current situation. So for instance a risk tolerant individual with a risky portfolio may be induced to choose a safe income stream, since she is already exposed to considerable risk. Conversely, a risk averse individual with a very safe portfolio can afford to choose a riskier income path. In both cases the observed relationship between the measured risk tolerance and portfolio choice is attenuated.

The variables Riskav1 and Riskav2 extracted from a factor analysis of *ad hoc* measures appear to be doing considerably better in terms of explanatory power, both for total assets and for financial assets. Interestingly, in the model for financial assets the measures Care1 and Care2 also contribute significantly to the explanation of portfolio shares. In fact, as Table 8 shows, these *ad hoc* measures seem to do as well as the Barsky *et al.* (1997) measure in explaining portfolio choice. This reinforces the notion that intuitive *ad hoc* measures may be more powerful in explaining portfolio choice than theory based, but complicated, risk tolerance measures.

Thus, the paper provides evidence that individually measured risk preference variables help explain portfolio allocation in line with economic theory. Perhaps surprising to economists, simple intuitive measures of risk preferences appear to be more powerful predictors of portfolio allocation than more sophisticated measures with a firmer basis in economic theory. Yet, for empirical analysis this may be good news. The *ad hoc* questions are much easier to ask and demand considerably less imagination of respondents, which may be exactly why the simple measures work better.

The modeling of incomplete portfolios by introducing fixed costs is relatively straightforward. So far we have restricted ourselves to just three broad group of assets. The set-up can be easily generalized to more than three assets.

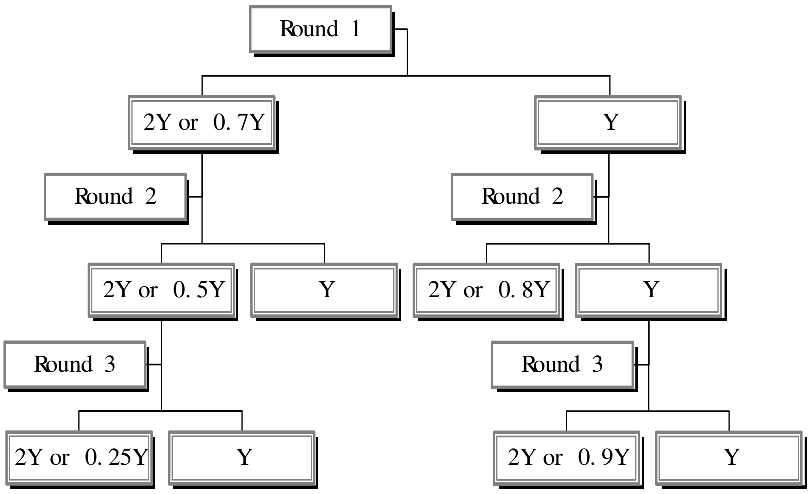

Figure 1: *Gambles over lifetime income. Respondents are asked to choose between a certain outcome (Y denoting the individual's self reported net personal income) and a risky outcome (a gamble over Y with several probabilities). The decision process consists of three consecutive rounds, so that the degree of risk involved in the uncertain outcome depends on the choice made in the previous round.*

Table 1: Risk Tolerance in the US, France and The Netherlands (percentages)

	Total sample		Respondents over 50		
Group	**France**	**Neth.**	**USA**	**France**	**Neth.**
I (most risk averse)	43.1	54.9	64.6	48.6	64.3
II	39.4	19.4	11.6	36.8	16.8
III	11.2	17.8	10.9	8.7	13.0
IV (least risk averse)	6.3	7.9	12.8	5.9	5.9
Total	100	100	100	100	100
N. Obs.	2954	1315	11707	2954	375

Source: DHS, 2004 for NL; BKJS, 1997 for US; Arrondel and Calvo-Pardo, 2002 for France

Table 2: Factor analysis of the risk aversion and the precautionary motives scales using PCA with varimax rotation

Factors	**Factor 1 (Riskav1)**	**Factor 2 (Riskav2)**	**Factor 3 (Precaution)**	**Uniqueness**
Risk aversion Scale				
I think it is more important to have safe investments and guaranteed returns, than to take a risk to have a chance to get the highest possible returns	0.0586	**0.8243**	0.1288	0.3005
I would never consider investments in shares because I find this too risky	0.4136	**0.5743**	-0.0981	0.4895
If I think an investment will be profitable, I am prepared to borrow money to make this investment (R)	**0.7552**	-0.0203	0.0621	0.4254
I want to be certain that my investments are safe	0.0909	**0.8137**	0.2113	0.2849
I get more and more convinced that I should take greater financial risks to improve my financial position (R)	**0.7683**	0.0582	0.1035	0.3957
I am prepared to take the risk to lose money, when there is also a chance to gain money (R)	**0.7602**	0.3018	-0.0248	0.3304
Precautionary motives Scale				
I save to have some savings to cover unforeseen expenses	0.0225	0.1257	**0.8813**	0.2070
I save so I have enough money in my bank account to be sure I will be able to meet my financial liabilities	0.0493	0.0978	**0.8778**	0.2174

Source: DHS, 2004 (N.Obs.: 1220)

- Answers coded in a 1 to 7 scale, from “very unimportant”to “very important”
- (R) denotes the statements whose score has been reversed to make all the statements consistent (from the most to the least risk averse / prudent)
- The largest factor loadings in each column are highlighted in bold face

Table 3: Risk attitude variables by individual background characteristics (means)

	[A]	[B]			[C]		
Characteristic	**Risk tolerance**	**Riskav1**	**Riskav2**	**Precaution**	**Care1**	**Care2 (R)**	**N.Obs.**
Gender							
Male	2.616	-0.169	-0.047	-0.104	5.863	6.161	748
Female	2.478	0.211	0.059	0.130	5.820	6.313	567
p-value	*0.066*	*0.000*	*0.067*	*0.000*	*0.461*	*0.017*	
Education level							
Low	2.132	0.179	-0.003	-0.022	5.939	6.359	170
Middle	2.509	0.007	0.015	0.033	5.832	6.123	251
High	2.719	-0.169	-0.017	-0.013	5.774	6.215	284
p-value	*0.000*	*0.000*	*0.899*	*0.712*	*0.066*	*0.010*	
Employment status							
Employees	not available	-0.137	-0.090	0.009	5.728	6.085	640
Self-employed	not available	-0.227	-0.204	0.015	5.591	6	37
p-value		*0.593*	*0.487*	*0.969*	*0.421*	*0.642*	
Whole sample	2.557	1.05e-09	1.69e-09	2.63e-10	5.840	6.215	1315

Source: DHS, 2004

- The reported p-values refer to one way analysis of variance of each of the risk attitude measures on the characteristics considered

- (R) denotes the statements whose score has been reversed to make all the statements consistent (from the most to the least risk averse)

- [A] risk aversion measure derived from gambles over lifetime incomes; [B] risk aversion measures extracted from factor analysis; [C] risk aversion measures elicited from direct questions

Table 4: Correlations of subjective measures of risk aversion

	Risk tolerance	**Riskav1**	**Riskav2**	**Precaution**	**Care1**
Risk tolerance		-0.210***	-0.153***	0.060	
		(628)	(628)	(628)	
Care1	-0.020	0.056	0.192***	0.312***	
	(656)	(1220)	(1220)	(1220)	
Care2	-0.109***	0.141***	0.126***	0.112***	0.299***
	(656)	(1220)	(1220)	(1220)	(1291)

Source: DHS, 2004

- *** denotes significance at the 1-percent level
- N.Obs. reported in parenthesis

Table 5: Determinants of the alternative measures of risk aversion

	Risk tolerance	**Riskav1**	**Riskav2**	**Precaution**	**Care1**	**Care2**
	Ord.probit	OLS	OLS	OLS	Ord.probit	Ord.probit
Expl. variables	(std.err.)	(std.err.)	(std.err.)	(std.err.)	(std.err.)	(std.err.)
Female	-0.165*	0.393***	0.136**	0.217***	-0.048	0.220***
	(0.085)	(0.059)	(0.060)	(0.062)	(0.065)	(0.071)
35 <= Age < 50	-0.237**	0.298***	0.0646	-0.0879	-0.008	0.270***
	(0.108)	(0.093)	(0.095)	(0.098)	(0.101)	(0.105)
50 <= Age < 65	-0.403***	0.430***	0.266***	-0.144	0.049	0.452***
	(0.119)	(0.095)	(0.097)	(0.100)	(0.104)	(0.108)
Age >= 65	-0.209	0.647***	0.343***	-0.273**	0.437***	0.765***
	(0.785)	(0.105)	(0.107)	(0.110)	(0.116)	(0.124)
Middle educ.	0.259**	-0.058	0.062	0.019	-0.103	-0.128
	(0.113)	(0.074)	(0.075)	(0.077)	(0.081)	(0.088)
High educ.	0.459***	-0.136*	0.061	-0.004	-0.160**	-0.092
	(0.111)	(0.074)	(0.075)	(0.077)	(0.081)	(0.089)
HH income	-0.021	-0.055	0.032	0.025	0.023	0.006
	(0.061)	(0.039)	(0.040)	(0.041)	(0.038)	(0.041)
HH total fin. wealth	-0.002	-0.063***	-0.001	0.053***	0.051**	0.099***
	(0.030)	(0.019)	(0.020)	(0.020)	(0.021)	(0.023)
Constant		0.690*	-0.583	-0.738*		
		(0.40)	(0.41)	(0.42)		
N. Obs.	650	1110	1110	1110	1155	1155
(Pseudo)R2	0.02	0.09	0.02	0.03	0.02	0.03
Log likelihood	-1000.676	-1518.795	-1539.015	-1568.571	-1560.071	-1333.681
p-value test age=0	0.009	0.000	0.001	0.070	0.000	0.000
p-value test educ=0	0.000	0.179	0.597	0.947	0.141	0.339

Source: DHS, 2004

- The omitted dummies serving as reference groups are Age < 35 and Low education, respectively
- Standard errors reported in parenthesis
- ***, **, * denote significance at the 1-percent, 5-percent, 10-percent level, respectively

Table 6: Ownership of risky asset - Probit estimates

Explanatory variables	**[I]** (z-stat)	**[II]** (z-stat)	**[III]** (z-stat)
Female	-0.138	-0.002	-0.168
	(1.17)	(0.02)	(1.90)
Age	-0.009**	0.003	-0.005
	(1.45)	(0.085)	(1.39)
Middle educ.	0.047	0.065	0.049
	(0.30)	(0.55)	(0.44)
High educ.	0.011	0.108	0.107
	(0.07)	(0.92)	(0.99)
HH income	0.062	0.084	0.090
	(0.71)	(1.18)	(1.33)
HH total fin. assets	0.556***	0.519***	0.503***
	(9.95)	(12.45)	(13.22)
Risk tolerance	**0.149*** **		
	(3.30)		
Riskav1		**-0.385*** **	
		(7.83)	
Riskav2		**-0.380*** **	
		(7.80)	
Care1			**-0.002**
			(0.05)
Care2			**-0.058**
			(1.43)
Constant	-6.611***	-6.831***	-5.809***
	(7.02)	(8.93)	(7.78)
N.Obs.	650	1100	1155

Source: DHS, 2004

- The alternative measures of risk aversion are defined in Section 2
- ***, **, * denote significance at the 1-percent, 5-percent, 10-percent level, respectively

Table 7: Risky asset share of financial wealth - OLS estimates

Explanatory variables	**[I]** (z-stat)	**[II]** (z-stat)	**[III]** (z-stat)
Female	-0.040**	-0.007	-0.018
	(2.50)	(0.57)	(1.50)
Middle educ.	0.020	0.013	0.009
	(0.98)	(0.88)	(0.58)
High educ.	0.023	0.02@	0.022
	(1.11)	(1.40)	(1.46)
HH income	-0.009	-0.001	-0.001
	(0.85)	(0.16)	(0.18)
HH total fin. assets	0.038***	0.034***	0.037***
	(6.87)	(8.59)	(9.56)
Risk tolerance	**0.015****		
	(2.38)		
Riskav1		**-0.030*****	
		(4.93)	
Riskav2		**-0.036*****	
		(5.98)	
Care1			**0.004**
			(0.60)
Care2			**-0.011****
			(2.03)
Constant	-0.188***	-0.265***	-0.206***
	(1.59)	(3.18)	(2.49)
N.Obs.	650	1100	1155
R-squared	0.10	0.13	0.09

Source: DHS, 2004

- The alternative measures of risk aversion are defined in Section 2
- ***, **, * denote significance at the 1-percent, 5-percent, 10-percent level, respectively

Table 8: Comparing goodness of fit

Risk measures	**Pseudo R2 - Probit**	**Adjusted R2 - OLS**
Riskav	*0.2865*	*0.1344*
Risk tolerance	*0.2059*	*0.1005*
Care	*0.1984*	*0.0988*
N.Obs.	588	588

Source: DHS, 2004

The alternative measures of risk aversion are defined in Section 2

Table 9: Risky asset share of financial wealth - ML estimations

	Risk tolerance **Coefficient** *(t-value)*	Riskav1 Riskav2 **Coefficient** *(t-value)*	Careful/less **Coefficient** *(t-value)*
Risk aversion	0.048	0.323	0.099
	(0.87)	*(2.54)*	*(1.74)*
x1			
Female	-0.291	0.044	0.009
	(1.66)	*(0.54)*	*(0.13)*
35 <= Age < 50	0.056	-0.141	-0.181
	(0.30)	*(0.73)*	*(1.1)*
50 <= Age < 65	-0.146	0.844	1.004
	(0.71)	*(4.01)*	*(5.32)*
Age >= 65	-0.594	0.776	0.953
	(0.00)	*(3.53)*	*(4.88)*
HH income	0.211	0.006	0.000
	(1.99)	*(0.08)*	*(0.01)*
Constant	-1.683	-0.641	-0.523
	(1.25)	*(0.69)*	*(0.67)*
x2			
Female	0.028	-0.056	-0.047
	(0.02)	*(0.43)*	*(0.35)*
35 <= Age < 50	-0.058	0.351	0.363
	(0.06)	*(1.21)*	*(0.97)*
50 <= Age < 65	-0.116	0.398	0.440
	(0.08)	*(1.50)*	*(1.40)*
Age >= 65	-	0.708	0.707
	(-)	*(2.49)*	*(1.98)*
HH income	0.030	0.455	0.445
	(0.03)	*(4.52)*	*(4.09)*
Constant	-0.282	-5.385	-5.170
	(0.03)	*(4.86)*	*(4.35)*
Fixed costs - safe asset and 1st risky asset			
Middle educ.	2272	1231	1710
	(1.20)	*(0.83)*	*(1.10)*
High educ.	1964	699	800
	(1.21)	*(2.03)*	*(0.58)*
Constant	5924	8601	8723
	(4.23)	*(43.15)*	*(8.11)*
Fixed costs - safe asset and 2st risky asset			
Middle educ.	-82941	1939	1416
	(0.38)	*(0.64)*	*(0.33)*
High educ.	-63995	-1118	-1808
	(0.29)	*(0.43)*	*(0.53)*
Constant	109018	17078	17502
	(0.49)	*(13.95)*	*(5.27)*
Fixed costs - both risky assets			
Middle educ.	-515218	-3014	558
	(68.17)	*(0.38)*	*(0.18)*
High educ.	-509454	-8730	-4778
	(-)	*(1.18)*	*(1.29)*
Constant	540776	31840	25644
	(60.84)	*(4.71)*	*(9.57)*

Table 9 continued: Risky asset share of financial wealth - ML estimations - cont'd

	Risk tolerance **Coefficient** *(t-value)*	Riskav1 Riskav2 **Coefficient** *(t-value)*	Careful/less **Coefficient** *(t-value)*
gamma	4.655	3.708	4.476
	(10.44)	*(7.17)*	*(13.16)*
lambda	-	0.524	0.251
	(-)	*(4.95)*	*(0.61)*
rho	0.583	0.056	0.192
	(0.22)	*(0.16)*	*(0.71)*
kappa1	-0.003	0.275	0.007
	(0.00)	*(0.67)*	*(0.03)*
kappa2	-0.030	0.392	0.146
	(0.01)	*(1.11)*	*(0.36)*
N.Obs.	650	1110	1155
Log-likelihood	-567.11	-1045.776	-1088.97

Source: DHS, 2004

The alternative measures of risk aversion are defined in Section 2

References

[1] Ahearne, A., Griever, W. and Warnock, F. (2004), "Information Costs and Home Bias: an Analysis of US Holdings of Foreign Equities", *Journal of International Economics* 62, 313-336.

[2] Amadi, A. and Bergin, P. (2006), "Understanding International Portfolio Diversification and Turnover Rates", *NBER Working Paper 12473.*

[3] Arrondel, L. and Calvo-Pardo, H. (2002), "Portfolio Choice with a Correlated Background Risk: Theory and Evidence", *Delta Working Paper 2002-16.*

[4] Baele, L., Pungulescu, C. and Ter Horst, J. (2007), "Model Uncertainty, Financial Market Integration and the Home Bias Puzzle", *Journal of International Money and Finance* 26, 606-630.

[5] Barsky, R.B., Juster, F.T., Kimball, M.S., Shapiro, M.D. (1997), "Preference Parameters and Behavioral Heterogeneity: an Experimental Approach in the Health and Retirement Study", *Quarterly Journal of Economics* 112, 537-579.

[6] Bernartzi, S. and Thaler, R. (2001), "Naive Diversification Strategies in Defined Contribution Saving Plans", *American Economic Review* 91, 79-98.

[7] Brennan, M. (1975), "The Optimal Number of Securities in a Risky Asset Portfolio when there are Fixed Costs of Transacting: Theory and Some Empirical Results", *Journal of Financial and Quantitative Analysis* 10(3), 483-496.

[8] Campbell, J. (2006), "Household Finance", *Journal of Finance* 61, 1553-1604.

[9] Choe, H., Kho, B. and Stulz, R. (2005), "Do Domestic Investors Have an Edge? The Trading Experience of Foreign Investors in Korea", *Review of Financial Studies* 18, 795-829.

[10] Cooper, I.A. and Kaplanis, E. (1994), "Home Bias in Equity Portfolios, Inflation Hedging, and International Capital Market Equilibrium", *The Review of Financial Studies* 7 (1), 45-60.

[11] Christelis, D. and Georgarakos, D. (2008), "Investing at Home and Abroad: Different Costs, Different People?", *CSEF Working Paper 188.*

[12] Dahlquist, M., Pinkowitz, L., Stulz, R. and Williamson, R. (2003), "Corporate Governance, and the Home Bias", *Journal of Financial and Quantitative Analysis* 38, 87-110.

[13] Dvorak, T. (2005), "Do Domestic Investors Have an Information Advantage? Evidence from Indonesia", *Journal of Finance* 60, 817-838.

[14] Fidora, M., Fratzscher, M. and Thimann, C. (2007), "Home Bias in Global Bond and Equity Markets: The role of Real Exchange Rate Volatility", *Journal of International Money and Finance* 26, 631-655.

[15] French, K.R. and Poterba, J.M. (1990), "Japanese and US Cross-border Common Stock Investments", *Journal of the Japanese and International Economies* 4, 476-493.

[16] French, K.R. and Poterba, J.M. (1991), "Investor Diversification and International Equity Markets", *American Economic Review* 81, 222-226.

[17] Gelon, R. and Wei, S. (2005), "Transparency and International Portfolio Holdings", *Journal of Finance* 60, 2987-3020.

[18] Glassman, D.A. and Riddick, L.A. (2001), "What Causes Home Asset Bias and How Should It Be Measured?", *Journal of Empirical Finance* 8, 35-54.

[19] Gollier, C. (2004), "The Economics of Risk and Time", MIT Press.

[20] Graham, J., Harvey, C. and Huang, H. (2005), "Investor Competence, Trading Frequency, and Home Bias", *NBER Working Paper 11426.*

[21] Guiso, L., Sapienza, P. and Zingales, L. (2008), "Trusting the Stock Market", *Journal of Finance* 63, 2557-2600.

[22] Guiso, L. and Jappelli, T. (2008), "Financial Literacy and Portfolio Diversification", *European University Institute Working Paper ECO2008/31.*

[23] Haliassos, M. and Hassapis, C. (2002), "Equity Culture and Household Behavior", *Oxford Economic Papers* 54, 719-745.

[24] Hong, H., Kubik, J. and Stein, J. (2004), "Social Interaction and Stock-Market Participation", *Journal of Finance* 59, 137–163.

[25] Iezzi, S. (2008), "Investors' Risk Attitude and Risky Behavior: a Bayesian Approach with Imperfect Information", *Bank of Italy Working Paper 692.*

[26] Jacob, N. (1974), "A Limited-Diversification Portfolio Selection Model for the Small Investor", *Journal of Finance* 29(3), 847-856.

[27] Jermann, U. (2002), "International Portfolio Diversification and Endogenous Labor Supply Choice", *European Economic Review* 46, 507-522.

[28] La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R. (1999), "Corporate Ownership Around the World", *Journal of Finance* 54, 471-515.

[29] Levin, I., Snyder, M. and Chapman, D. (1988), "The Interaction of Experiential and Situational Factors and Gender in a Simulated Risky Decision-Making Task", *Journal of Psychology* 122(2), 173-181.

[30] Levy, H. (1978), "Equilibrium in an Imperfect Market: A Constraint on the Number of Securities in the Portfolio", *European Economic Review* 68(4), 643-658.

[31] Lobo, M., Fazel, M. and Boyd, S. (2007), "Portfolio Optimization with Linear and Fixed Transaction costs", *Annals of Operations Research* 152(1), 376-394.

[32] Mao, J. (1970), "Essentials of Portfolio Diversification Strategy", *Journal of Finance* 25(5), 1109-1121.

[33] Patel, N. and Subrahmanyam, M. (1982), "Optimal Portfolio Selection with Fixed Transaction Costs", *Management Science* 28(3), 303-314.

[34] Rooij van M., Kool, C. and Prast, H. (2007), "Risk-return Preferences in the Pension Domain: Are People Able to Choose? ", *Journal of Public Economics* 91, 701-722.

[35] Strong, N. and Xu, X. (2003), "Understanding the Equity Home Bias: Evidence from Survey Data", *Review of Economics and Statistics* 85, 307-312.

[36] Tesar, L. and Werner, I. (1994), "International Equity Transactions and US Portfolio Choice", in Frankel, J.A. (ed.) *The Internationalization of Equity Markets,* Univ. of Chicago Press, Chicago, 185-227.

[37] Tesar, L. and Werner, I. (1995), "Home Bias and High Turnover", *Journal of International Money and Finance* 14, 467-492.

[38] Vissing-Jorgensen, A. (2002), "Towards and Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures", *NBER Working Paper 8884.*

APPENDIX

Table A1: Definitions of the independent variables

Label	Definition
Female	Female indicator
Age < 35	1 if age less than 35 years (ref. group)
35 <= Age < 50	1 if age between 35 and 50 years
50 <= Age < 65	1 if age between 50 and 65
Age >= 65	1 if age greater than 65
Low educ.	1 if low education (ref. group)
Middle educ.	1 if middle education
High educ.	1 if high education
HH income	Household income in logs
HH total financial wealth	Household total financial wealth in logs

Table A2: Households' assets and liabilities in 2004

Assets / Liabilities	Type	% own	Mean	% fin.ass.	% total ass.
Checking, dep., ins., etc.	*NR*	98.5	32,300	62.6	14
Growth, mutual funds	*R*	21.7	4,890	10.4	2.1
Bonds	*O*	5	2,277	4.4	0.1
Stocks	*R*	16	3,918	7.5	1.7
Options	*R*	1.4	89	0.2	0.03
Money lent out	*O*	8	976	0.2	0.4
Business equity	*O*	4.7	7,682	14.7	3.3
TOTAL FINANCIAL ASSETS		-	52,132	100	22.6
Real estate	*O*	64	178,625	-	77.4
TOTAL ASSETS		-	230,757	-	100
Financial debt	*O*	-	5,606	10.8	2.4
Mortgage	*O*	52.2	54,121	-	23.5
NET FINANCIAL WEALTH		-	46,526	89.2	20.2
Non-risky assets	*NR*	98.5	32,300	62.6	14
Risky assets	*R*	39.1	8,896	17.1	3.9
Other assets	*O*	81.7	189,560	176.1	56.2
NET WORTH		-	171,030	-	74.1

Source: DHS, 2004 (N.Obs.: 1549)

- All amounts are expressed in Euro's

- NR denotes "Non-risky"; R denotes "Risky"; O denotes "Other"

A.1 Derivation of regime utilities and choice probabilities for the case $k = 3$

By inserting (8) into the formula for the utility function (4) we obtain as the value of the utility function in the (interior) optimum:

$$\begin{aligned} V^*(\tilde{\alpha}) &= -\exp(-Aw_0 - \frac{1}{2}b'\Psi^{-1}b) = \\ & -\exp(-Aw_0 - Aw_0 z'\mu - \frac{1}{2}\mu'\Sigma^{-1}\mu - \frac{1}{2}w_0^2 A^2 z'\Sigma z) \end{aligned} \tag{14}$$

We note that for this expression to make sense the expression must by increasing in w_0. The expression in the exponent is quadratic in w_0. For V^* to be increasing in w_0 we need

$$w_0 > -\frac{1 + z'\mu}{Az'\Sigma z} \tag{15}$$

Now consider the value of the utility function for incomplete portfolios, i.e. where some of the elements of α are zero. It will be convenient to define a selection matrix that selects the elements of the portfolio that are zero. Let S' (k_1 x $(k-1)$) and D' (k_2 x $(k-1)$) be selection matrices with $k_1 + k_2 = k - 1$, i.e. $\begin{bmatrix} S' \\ D' \end{bmatrix}$ is a permutation of I_{k-1}, the $(k-1)$ x $(k-1)$ identity matrix. The matrix S' selects the elements of α which are zero and D' selects the elements of α which are non-zero. Some useful properties of S and D are:

$$S'S = I_{k_1} \quad D'D = I_{k_2} \quad SS' + DD' = I_{k-1} \quad D'S = 0 \tag{16}$$

Using this notation the exponent of the utility function (7) can be written as

$$\begin{aligned} & -Aw_0 - \frac{1}{2}b'\Psi^{-1}b + \frac{1}{2}\left[(\alpha - \Psi^{-1}b)' \begin{pmatrix} S & D \end{pmatrix} \begin{pmatrix} S' \\ D' \end{pmatrix} \Psi \begin{pmatrix} S & D \end{pmatrix} \begin{pmatrix} S' \\ D' \end{pmatrix} (\alpha - \Psi^{-1}b)\right] = \\ & -Aw_0 - \frac{1}{2}b'\Psi^{-1}b + \frac{1}{2}\left[\begin{pmatrix} S' \\ D' \end{pmatrix}\alpha - \begin{pmatrix} S' \\ D' \end{pmatrix}\Psi^{-1}b\right]' \begin{bmatrix} S'\Psi S & S'\Psi D \\ D'\Psi S & D'\Psi D \end{bmatrix} \left[\begin{pmatrix} S' \\ D' \end{pmatrix}\alpha - \begin{pmatrix} S' \\ D' \end{pmatrix}\Psi^{-1}b\right] \end{aligned} \tag{17}$$

If we impose that $S'\alpha = 0$, we can again write equation (17) as a quadratic expression in the unrestricted asset quantities $D'\alpha$. Consider the quadratic expression in (17) and adopt temporarily the following simplified notation:

$$\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}' \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = q_1' A_{11} q_1 + 2q_2' A_{21} q_1 + q_2' A_{22} q_2 =$$

$$\left(q_2 + A_{22}^{-1} A_{21} q_1\right)' A_{22} \left(q_2 + A_{22}^{-1} A_{21} q_1\right) + q_1' \left(A_{11} - A_{12} A_{22}^{-1} A_{21}\right) q_1 \qquad (18)$$

This expression is minimized for $q_2 = -A_{22}^{-1} A_{21} q$, while the minimum attained is $q_1' \left(A_{11} - A_{12} A_{22}^{-1} A_{21}\right) q_1$. Returning to the original notation and inserting this into (17), we obtain for the value of the utility function

$$\begin{aligned} U(\alpha) &= -\exp(-Aw_0 - \frac{1}{2} b' \Psi^{-1} b + \frac{1}{2} b' \Psi^{-1} S \left[S' \Psi S - S' \Psi D (D' \Psi D)^{-1} D' \Psi S\right] S' \Psi^{-1} b) = \\ & \quad -\exp(-Aw_0 - \frac{1}{2} b' \left[\Psi^{-1} - \Psi^{-1} S (S' \Psi^{-1} S)^{-1} S' \Psi^{-1}\right] b) \end{aligned} \qquad (19)$$

Now consider a regime i with non-zero risky assets $D_i' \alpha$, while the vector of assets $S_i' \alpha$ is zero. Let the corresponding fixed costs of ownership of that particular asset combination be f_i . A fixed cost can be seen as a reduction of w_0. Note that w_0 also appears in the expression for b (see below (6)), so we also index b by the regime by writing b_i. The utility of regime i is therefore

$$-A(w_0 - f_i) - \frac{1}{2} b_i' [\Psi^{-1} - \Psi^{-1} S_i (S_i' \Psi^{-1} S_i)^{-1} S_i' \Psi^{-1}] b_i \qquad (20)$$

For completeness, we also write (20) in terms of the original parameters:

$$-A(w_0 - f_i) - \frac{1}{2} (\mu + A(w_0 - f_i) \Sigma z)' [\Sigma^{-1} - \Sigma^{-1} S_i (S_i' \Sigma^{-1} S_i)^{-1} S_i' \Sigma^{-1}] (\mu + A(w_0 - f_i) \Sigma z) \qquad (21)$$

In this expression the vector z is a random variable and hence (21) is random. For concreteness we restrict ourselves to the the case where $k = 3$. This is the case for which we will estimate the model, although in principle one can easily generalize our approach to $k > 3$. For simplicity, let us define the following quantities:

$$\Gamma \equiv \Sigma^{-1}, c_i = A(w_0 - f_i), y \equiv \Sigma z,\ G_i \equiv \Sigma^{-1} - \Sigma^{-1} S_i (S_i' \Sigma^{-1} S_i)^{-1} S_i' \Sigma^{-1} \qquad (22)$$

It is useful to note that

$$\Gamma^{-1} = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{12} & \gamma_{22} \end{pmatrix}^{-1} = \frac{1}{\gamma_{11}\gamma_{22} - \gamma_{12}^2} \begin{pmatrix} \gamma_{22} & -\gamma_{12} \\ -\gamma_{12} & \gamma_{11} \end{pmatrix} = \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \qquad (23)$$

Now consider the four possible cases:

A - *Only the safe asset is in the portfolio.* We indicate this case by a subscript A. We have that $S_A = I_2$ and $G_A = 0$. Furthermore, by assumption $f_A = 0$. So the corresponding expression in (21) is equal to $-Aw_0$.

B - *The safe asset and the first risky asset are in the portfolio.* We indicate this case by a subscript F. We have that $S'_F = (0,1)$, $G_F = \begin{pmatrix} \gamma_{11} - \frac{\gamma_{12}^2}{\gamma_{22}} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 \\ 0 & 0 \end{pmatrix}$. The corresponding expression in (21) is equal to $-c_F - \frac{1}{2\sigma_1^2}(\mu_1 + c_F y_1)^2$. The notation is such that the subscript 1 for μ and y denote the first element of the corresponding vectors; however the subscipt F for f and c refer to the fact that this is case B where the first risky asset is in the portfolio.

C - *The safe asset and the second risky asset are in the portfolio.* We indicate this case by a subscript S. We have that $S'_S = (1,0)$, $G_S = \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{\sigma_2^2} \end{pmatrix}$. The corresponding expression in (21) is equal to $-c_S - \frac{1}{2\sigma_2^2}(\mu_2 + c_S y_2)^2$. As before, the subscript 2 for μ and y denotes the second element of the corresponding vectors, whereas the subscipt S for f and c refer to the fact that this is case S where the second risky asset is in the portfolio.

D - *Both risky assets are in the portfolio.* We indicate this case by subscript B. We have that S_B is the zero matrix. Hence $G_B = \Gamma$ and the corresponding expression in (21) becomes $-c_B - \frac{1}{2}(\mu + c_B\Sigma z)'\Gamma(\mu + c_B\Sigma z) = -c_B - \frac{1}{2}(\mu + c_B y)'\Gamma(\mu + c_B y)$.

Denote the utilities of the four regimes by $\widetilde{U}_A$, $\widetilde{U}_B$, $\widetilde{U}_C$, $\widetilde{U}_D$, respectively. As in Van Soest (1995) we add random error terms to the four utilities as follows

$$U_i = \widetilde{U}_i + \varsigma_i, \quad i = A, B, C, D \tag{24}$$

where the ζ_i are independent and extreme value distributed. Thus the probability that we will observe case i is equal to

$$P_i = \frac{\exp\{\widetilde{U}_i\}}{{}_{j=1}^{4} \exp\{\widetilde{U}_j\}} \tag{25}$$

Note that $\widetilde{U}_i$ depends on the random variables ϵ_k ($k = 1, 2$), cf. equation (5). We have to integrate these random variables out, but that can be done in the usual way through simulation.

Let n index an observation: $n = 1, ...N$. Then the general form of the likelihood conditional on the error terms is

$$L =_{n=1}^{N} \ln\{P_{ni}|\epsilon_{ni}\} \tag{26}$$

Suppose that for each observation we draw R times from the distribution of ϵ_{ni}, then we approximate the log-likelihood by

$$L_S = \frac{1}{R}{}_{r=1}^{R}{}_{n=1}^{N} \ln\{P_{ni}|\widehat{\epsilon}_{nir}\} \tag{27}$$

where $\widehat{\epsilon}_{nir}$ is the r-th drawing for observation n.

It is useful to reparameterize the model so as to obtain better insight into the identification properties. Define the following quantities:

$$Q \equiv \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}; \quad P = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \tag{28}$$

Thus $\Sigma = PQP$, $\Gamma = P^{-1}Q^{-1}P^{-1}$, $y = \Sigma z = PQPz$. Consider the various cases. Begin with case D:

$$\begin{aligned} -c_B - \frac{1}{2}(\mu + c_B y)'\Gamma(\mu + c_B y) &= -c_B - \frac{1}{2}\left(P^{-1}\mu + c_B P^{-1}y\right)' Q^{-1}\left(P^{-1}\mu + c_B P^{-1}y\right) = \\ & -c_B - \frac{1}{2}\left(P^{-1}\mu + c_B QPz\right)' Q^{-1}\left(P^{-1}\mu + c_B QPz\right) \end{aligned} \tag{29}$$

We note that

$$P^{-1}\mu = \begin{pmatrix} \frac{\mu_1}{\sigma_1} \\ \frac{\mu_2}{\sigma 2} \end{pmatrix}, QPz = \begin{pmatrix} \sigma_1 z_1 + \rho\sigma_2 z_2 \\ \rho\sigma_1 z_1 + \sigma_2 z_2 \end{pmatrix} \tag{30}$$

Recall that $z = \Lambda x + \epsilon$. Thus we can write $QPz = QP\Lambda x + QP\epsilon \equiv Bx + u$, with $B = QP\Lambda$ and $u = QP\epsilon$. Thus we can also write

$$QPz = \begin{pmatrix} x'\beta_1 + u_1 \\ x'\beta_2 + u_2 \end{pmatrix} \tag{31}$$

in rather obvious notation. So we can write (29) as

$$-c_B - \frac{1}{2}\left[\begin{pmatrix} \frac{\mu_1}{\sigma_1} \\ \frac{\mu_2}{\sigma 2} \end{pmatrix} + c_B \begin{pmatrix} x'\beta_1 + u_1 \\ x'\beta_2 + u_2 \end{pmatrix}\right] \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}^{-1} \left[\begin{pmatrix} \frac{\mu_1}{\sigma_1} \\ \frac{\mu_2}{\sigma 2} \end{pmatrix} + c_B \begin{pmatrix} x'\beta_1 + u_1 \\ x'\beta_2 + u_2 \end{pmatrix}\right] \tag{32}$$

Using this notation we can write case B as

$$-c_F - \frac{1}{2\sigma_1^2}(\mu_1 + c_F y_1)^2 = -c_F - \frac{1}{2}\left(\frac{\mu_1}{\sigma_1} + c_F\left[x'\beta_1 + u_1\right]\right)^2 \tag{33}$$

And case C:

$$-c_S - \frac{1}{2\sigma_2^2}(\mu_2 + c_S y_2)^2 = -c_S - \frac{1}{2}\left(\frac{\mu_2}{\sigma_2} + c_S\left[x'\beta_2 + u_2\right]\right)^2 \tag{34}$$

A.2 Identification

Assume we can identify the vectors β_1 and β_2. Recall that $c_i = A(w_0 - f_i)$. Assume for a moment that f_i is known so that $(w_0 - f_i)$ is observable. We don't have a measure of A but rather a proxy. Say that we have a measure $\widetilde{A}$ and we assume $A = \delta\widetilde{A}$, where δ needs to be estimated. This introduces an indeterminacy in the term $c_i\,[x'\beta_2 + u_2]$. The only identifying information about δ is then contained in the term $-c_i$. That seems a rather weak basis for identifiability. In principle, the terms $\frac{\mu_1}{\sigma_1}$ and $\frac{\mu_2}{\sigma_2}$ are identified as intercepts and if we are willing to assume values for μ_1 and μ_2 then we can infer σ_1 and σ_2. There is no harm however in replacing the terms $\frac{\mu_1}{\sigma_1}$ and $\frac{\mu_2}{\sigma_2}$ by new parameters κ_1 and κ_2, say. Finally, ρ is identified from (32). We also note that the signs of coefficients κ_1 and κ_2 and $\beta_1$and β_2 are not identified. In (33) and (34) these parameters appear in quadratic terms so that we can change their signs without affecting the corresponding expressions for the utility levels. In (32) we can change the sign of κ_1 and $\beta_1$without affecting the level of the utility function, if we also change the sign of ρ. A similar observation holds for κ_2 and β_2.