MANNHEIM RESEARCH INSTITUTE FOR THE ECONOMICS OF AGING

HOW DOES HOUSEHOLD PORTFOLIO DIVERSIFICATION VARY WITH FINANCIAL SOPHISTICATION AND ADVICE?

Hans-Martin von Gaudecker

238-2011

L13, 17_D-68131 Mannheim_Phone +49 621 181-2773/1862_Fax +49 621 181-1863_www.mea.uni-mannheim.de

How does household portfolio diversification vary with financial sophistication and advice?

Hans-Martin von Gaudecker *

March 22, 2011

Abstract

Economic theory suggests that households should invest their financial wealth in a combination of cash and a well-diversified equity portfolio. Yet, many households' equity investments are strongly concentrated in a few assets. Attempts to explain this discrepancy have included low levels of cognitive skills and/or financial knowledge; and poor or misguided financial advice. In order to investigate these claims empirically, I construct detailed portfolios for the respondents to a Dutch household survey. The data allow me to estimate the portfolios' risk-return properties without resorting to assumptions about characteristics of specific asset classes. Controlling for a large number of covariates, my results show that the combination of low numerical-financial skills and not seeking advice from other persons is strongly associated with the largest losses from underdiversification, whereas financial knowledge does not seem to have much of an effect.

JEL codes: D14, D12, G11

Keywords: Household portfolios, diversification, financial literacy, financial advice.

*Universität Mannheim, Department of Economics; L 7, 3-5; 68131 Mannheim; Germany; `hmgaudecker[at] uni-mannheim[dot]de`; tel. +49 621 181 1879. Several of the data manipulations and preliminary estimations were performed by Mark Prins for his MSc Thesis at the VU University Amsterdam and in a subsequent research assistantship to the author. Funding from Netspar is gratefully acknowledged. Special thanks to him for an excellent programming job and many fruitful discussions. I appreciate helpful comments received from seminar and conference participants at Frankfurt, at the 2nd SAVE conference in Deidesheim, and the 2011 Netspar Pension Workshop in Amsterdam, in particular Tabea Bucher-Koenen, Dimitris Georgarakos, Michael Haliassos, Olivia Mitchell, Stephen Zeldes, and Michael Ziegelmeyer. Furthermore, I would like to thank CentERdata for providing the string variables necessary to create the individual return series; Mauro Mastrogiacomo for passing on code to prepare the DHS data; and Maarten van Rooij, Annamaria Lusardi, and Rob Alessie for providing the data and code used to construct the financial literacy scores in van Rooij *et al.* (forthcoming).

1 Introduction

Going back to at least Markowitz (1952), the canonical model of portfolio choice predicts (a) that households will hold a positive share in risky assets and (b) that the risky component will consist of a well-diversified portfolio, optimising its risk-return characteristics. The earlier empirical studies based on microeconomic data demonstrated that a large fraction of households do not hold any risky assets (e.g. Guiso *et al.*, 2002; Haliassos and Bertaut, 1995; Mankiw and Zeldes, 1991). This finding stimulated the development of a number of theoretical models which can account for this fact. Popular explanations include transaction costs (Vissing-Jørgensen, 2002), background risk (Heaton and Lucas, 2000), or behavioural economic theories (Barberis *et al.*, 2006). See Campbell (2006) for an overview.

Largely due to a lack of suitable data, the prediction of a well-diversified portfolio was hardly challenged until recently. Deviations from this recommendation have been documented first by Blume and Friend (1975), but most forcefully by Calvet *et al.* (2007). The latter employ extraordinarily detailed administrative data, which is only available in very selected countries. The first, minor, contribution of this paper is to demonstrate that their results are replicable to a large degree with survey data when households are asked for the specific items in their portfolios. Theoretical models that predict a low number of stocks in the portfolio are still rare,[1] so underdiversification would generally be considered an investment mistake. The main contribution of this paper is to document the pattern of how losses from underdiversification vary in the population. Compared to the administrative data of Calvet *et al.* (2007) or Grinblatt *et al.* (forthcoming), I have much more information about households and individuals, including various measures of financial literacy, the most important source of financial advice, risk attitudes, education, income and wealth.

Recently, there has been an increased interest in the lack of financial skills as a driver of poor financial decisions. The output measure has most often been undersaving (Bayer *et al.*, 2009; Cole and Shastry, 2009; Hilgert *et al.*, 2003; Lusardi and Mitchell, 2007a,b), although recent applications include stock market participation (van Rooij *et al.*, forthcoming), overindebtedness (Lusardi and Tufano, 2009), and mortgage delinquency (Gerardi *et al.*, 2010). Very recently, several authors have also connected financial literacy and portfolio diversification (e.g. Bilias *et al.*, 2009; Graham *et al.*, 2009; Guiso and Jappelli, 2009; Kimball and Shumway, 2010), but the available data constrains their choice of portfolio measures. I compare my results to some of those measures and show that there is a substantial benefit to using the more detailed data.

My results indicate that the majority of households reaches reasonable levels of diversification. Compared to investing in the benchmark portfolio, the median loss from underdiversification on the financial portfolio is limited to 29 basis points per year. However, the distribution has a fat right tail, where losses become very substantial. The worst outcomes are associated most with the combination of low levels of financial skills and relying on one's own financial judgement (as opposed to seeking advice from professionals or family/friends). This pattern holds regardless of the covariates controlled for. Financial knowledge does not appear to have an effect at any part of the distribution. The pattern suggests that policies targeting

[1] Very recently, van Nieuwerburgh and Veldkamp (2010) provided a rationale for low diversification based on information costs. My results show that such an explanation may justify some of the underdiversification seen in the data. However, it is unlikely to stand behind the portfolios incurring the largest losses. The same goes for the "beat the Jones" argument of Roussanov (2010).

either numerical-financial skills or the availability of advice may be effective in ameliorating the worst investment outcomes.

2 Data and empirical strategy

As discussed in the introduction, recent research has made clear that a significant fraction of households hold widely under-diversified portfolios (Calvet *et al.*, 2007). That and several other studies (Calvet *et al.*, 2009a,b; Massa and Simonov, 2006) are based on administrative records for the Swedish population, where banks were required to report the details of individuals' portfolios to the tax authorities. To a lesser extent, the same is true for Finland, where detailed stock holdings and an indicator of mutual fund ownership are available (Grinblatt *et al.*, 2011, forthcoming). To the best of my knowledge, such requirements are not in place anywhere outside Scandinavia and even in Sweden they have ceased to exist with the abolishment of the wealth tax in 2007. Since Sweden is unusual in a number of ways – most importantly for the topic at hand, a very high stock market participation rate with a stockholder pool that differs markedly from the one in other countries (Christelis *et al.*, 2010a) – it is important to find ways for conducting similar analyses in other regions. Furthermore, the use of administrative data limits the range of covariates that can be used to explain portfolio holdings to those collected by the government for administrative purposes. While the number of variables is substantial in Sweden, the content often does not exactly cover what a researcher would like to know.

A very popular alternative for investigating individual investment behaviour is to obtain data from discount brokers (Barber and Odean, 2000, 2001; Goetzmann and Kumar, 2008; Hackethal *et al.*, 2011; Ivković *et al.*, 2005, 2008; Korniotis and Kumar, forthcoming; Odean, 1998). An important advantage over the administrative data described before is that these datasets not only contain the portfolio composition at a certain date per year, but all trades in the observation period. Consequently, many of the just-cited studies have focused on the implications of suboptimal trading behaviour for portfolios' performance. These datasets are arguably less than optimal to study diversification issues because often only directly held stocks are observed in detail.[2] Furthermore, it is unknown (a) how much of households' portfolios the individuals' observed accounts cover and (b) to what extent holders of discount brokerage accounts are representative of the population of interest. Tang *et al.* (2010) pursue a related research strategy in comparing the actual performance of U.S. 401(k) pension plans with the optimal strategy under the investment menu offered by the pension provider. They demonstrate large losses from underdiversification, which almost exclusively stem from participants' choices. While the results are suggestive, one cannot know from such data whether at least part of the inefficiencies might be undone outside the tax-deferred accounts: Bergstresser and Poterba (2004) show that half of all individuals who own equity through retirement accounts also own equity outside of these accounts.

The most widespread instrument of empirical social science research is the household survey. The U.S. Survey of Consumer Finances has arguably been the most important source of knowledge about household saving and portfolio choice since its inception more than 3 decades ago (see, for example, the literature reviewed in Campbell, 2006). Important

[2] For example, Goetzmann and Kumar (2008) use a dummy for mutual fund holdings as an explanatory variable in a regression explaining the underdiversification of the stock portfolio.

recent contributions focussing on diversification issues include Christelis *et al.* (forthcoming) or Polkovnichenko (2005). Arguably the main strength of the SCF and general-purpose datasets with a strong module on financial matters[3] is that they contain a wealth of background information in addition to diversification proxies such as the number of directly held stocks, whether the household invests in mutual funds, and asset allocation shares. The main drawback of such surveys is the quality of such diversification measures[4] – some investors achieve good diversification results with a low number of stocks while some mutual funds concentrate their investments in very specific sectors. Calvet *et al.* (2009b, Online Appendix) find that among several potential diversification measures that can be constructed with prototypical survey data, the share of funds in the risky portfolio performs best. Whether the correlation of 0.49 between the fund shares and their favoured measure of diversification (for details on this measure, see Section 2.3 below) is high or low depends on the question at hand. It might well be reasonable as a control variable when the focus is on other questions; but if diversification issues play the central role in an analysis, one would hope for better measures.

In this study, I combine several strengths of the various approaches by constructing detailed portfolios for the respondents of the Dutch Central Bank Household Survey (DHS). I describe this survey in the first part of this section, emphasising measures of financial wealth. Linking individual portfolio components to historical return series allows me to calculate diversification statistics that are measured in meaningful economic quantities. After describing the linking procedure and the diversification measures, I sketch a production function framework for explaining investment outcomes. Finally, I outline the variables that serve as inputs into this function, most notably those regarding financial sophistication and advice for financial decision-making.

2.1 Financial wealth variables in the CentERpanel / DHS

I use data from the CentERpanel, a Dutch household survey that is administered via the Internet. In order to avoid selection problems due to lack of Internet access, respondents without a computer are equipped with a set-top box for their television set (and with a TV if they do not have one). Respondents are reimbursed for their costs of using the Internet. The panel consists of more than 1,500 households who are representative of the Dutch population in terms of observable characteristics. It has rich background information on important demographic and socio-economic variables. The CentERpanel was the role model for the RAND American Life Panel, which has emerged as another workhorse in the area of household financial decision-making (Hung *et al.*, 2009; Hung and Yoong, 2010; Lusardi and Mitchell, 2007b).

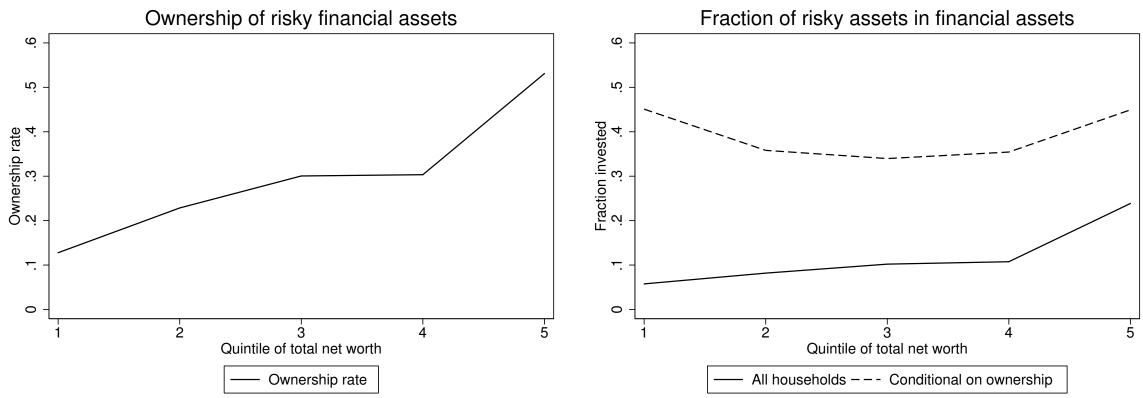
The CentERpanel hosts the Dutch Central Bank Household Survey (DHS), which contains particularly detailed information on financial matters. For this reason, it has been used extensively to describe the portfolio choice behaviour of Dutch households, excellent examples are Alessie *et al.* (2002, 2004, 2006); Dimmock and Kouwenberg (2010); Korniotis and Kumar (forthcoming). My analysis is cross-sectional, but in order to increase the sample size I make

[3]Some important examples are the HRS or PSID in the U.S. or the SHARE, ELSA, or ECHP datasets in Europe. Bilias *et al.* (2010); Christelis *et al.* (2010a,b); Juster *et al.* (1999); or Lusardi and Mitchell (2007a) are some exemplary studies using and describing these datasets for related questions.

[4]This characteristic is shared by other questionnaire-type approaches that I am aware of, such as tailor-made surveys (Kimball and Shumway, 2010), commercial investor surveys (Graham *et al.*, 2009), or hybrids thereof (Guiso and Jappelli, 2006, 2009).

use of the 2005 and 2006 waves, which contain information on the financial portfolios at the end of the previous year. Because of this, I label them with the years 2004 and 2005 in the remainder of the paper. Table A.1 in the Appendix contains an overview of the relevant asset and debt categories. The 35 entries in the first column are clearly too many to analyse for my purposes and the remaining columns show how I aggregate them into manageable numbers. The most important distinctions are in the upper part of the columns labelled "Level 2", differentiating between risky and safe financial assets, and "Level 1", which breaks up risky financial assets into three categories: Mutual funds, directly held stocks, and bonds and options. Throughout the analysis, I exclude households with less than 1,000 Euros in financial assets (8.6% of the sample), leaving 2,661 observations on 1,607 households.

Figure 1: Ownership rates of and fractions invested in risky assets, by total net worth



Source: CentERpanel, own calculations.

Figure 1 shows risky asset ownership along with unconditional and conditional shares by total net worth.[5] The left panel shows the standard pattern of rising ownership rates in wealth (Guiso *et al.*, 2002), starting from about 13% in the lowest wealth quintile to more than 50% in the highest wealth quintile. As usual, the rise is most pronounced in the highest wealth class. A similar pattern can be seen for the shares invested in risky assets, when averaging over all households, including the non-participants. These rise from 6% in the lowest wealth quintile over 10-11% in the upper-middle quintiles to 24% in the highest. The second line in the right panel reveals that this pattern is mostly driven by ownership rates. Conditional on ownership, risky asset holdings follow a U-shaped pattern with 45% in the two extreme quintiles and 34-36% in the middle quantiles. This contrasts with the pattern of a steep wealth gradient in the risky asset share among participants in Sweden (Calvet *et al.*, 2007, 2009b) and highlights the importance of establishing results along those studies' lines for other countries.

2.2 Detailed portfolio components

A unique feature of the dataset is that individuals are not only asked for the number of stocks and mutual funds they posses, but also to report the names and quantities held in each

[5]Total net worth is defined as total assets minus total debt in the last column of Table A.1.

of those.[6] In particular, they are asked for the details of their 10 (5, 5) largest positions of stocks (mutual funds, growth funds[7]). These three items make up the largest share of all risky assets, which furthermore consist of company or mortgage bonds and options (see Table A.1 for details).

Table 1: Descriptive statistics on coverage of risky portfolio by components with time series

	Variable name	# raw	# obs	Mean	p_5	Median	p_{95}
(1)	Total number of households	2661	1607				
(2)	Owners of risky financial assets	837	528				
(3)	Owners of shares/funds	791	500				
(4)	Raw report of individual items	789	498	3.33	1	2	10
(5)	Raw report (hh. in final sample)	648	408	3.69	1	2	11
(6)	Matched report of individual items	648	408	2.96	1	2	9
(7)	Fraction of shares/funds covered	648	408	.898	.278	.989	1
(8)	Fraction of risky fin. assets covered	648	408	.842	.183	.989	1
(9)	Fraction of quantities imputed	648	408	.081	0	0	.75
(10)	Length of time series of returns	269	269	138	57	128	235
(11)	Total expense ratio, mutual funds	170	170	1.3	.35	1.26	1.87

Source: CentERpanel, Datastream, Euroinvestor, own calculations. Numbers in column "# raw" refer to all observations, those in column "# obs" are adjusted for clustering at the household level, as are the remaining statistics. In the case of the time series of asset returns, the number of observations refers to the number of different assets. Returns are observed at a monthly frequency. The total expense ratio is expressed as an annual percentage of the asset value.

The second and third row of Table 1 reveal that of the 528 households who own any risky financial assets, 5% do not own any shares or mutual funds, but only bonds or options. Only two of the remaining households did not provide the names of any of their assets. Several reports were difficult to interpret, leaving 408 different households for whom I could match return series to the larger part of the portfolios' components. Row 5 in Table 1 shows that the mean household in the final sample holds 3.7 different items. These numbers are close to those found for the U.S. (e.g. Bilias *et al.*, 2009; Polkovnichenko, 2005) or Sweden (Calvet *et al.*, 2007).

The next row shows that I can match close to 3 items on average to historical returns on Datastream and Euroinvestor (not all funds were available on Datastream). The true rate of matches is even higher than the 80% implied by these numbers because in some households, multiple individuals answer the questionnaires and name the same portfolio items. These are

[6] The names of individual stocks and mutual funds are not part of the data that is available by default. They may be obtained from CentERdata for a small administrative charge.

[7] Growth funds are essentially the same as mutual funds, except for the fact that they reinvest any dividends and interest they receive from their investments. The distinction is made in the questionnaire due to different tax treatments. I do not maintain this distinction and refer to both as mutual funds.

consolidated in row 6, but not in rows 4 or 5. A similarly positive picture emerges when inspecting the fraction of assets covered by the portfolio components for which a return series is available. The average coverage rate is about 90% with a very left-skewed distribution – the median is at 98.9%. Adding bonds and options to the denominator reduces average rates by 5 percentage points and only affects the lower tail of the distribution. In the analysis below, I assume that the unobserved part of the risky portfolio behaves the same way as the observed part and exclude households with coverage rates below 30%, having checked robustness to various levels. Some individuals stated the name of a portfolio component, but did not provide information on the amount held. I imputed this information by assuming that the difference between total portfolio holdings and portfolio holdings attributable to specific assets is equally distributed among all reported assets. This implies that the coverage figures mentioned before are potentially overstated. However, row 9 of Table 1 reveals that this affects only 8% of the portfolio balances, and that it is concentrated among much less than half of all households. Furthermore, half of those who do not provide quantity information on an individual portfolio component have only this one item in their portfolio. Hence, no bias of the diversification results, which are independent of portfolio size, would arise from these households.

The bottom part of Table 1 shows that households reported ownership of 269 different assets; of which 170 are mutual funds and 99 are shares. I use the maximum available period for the returns from January 1990 to June 2009, or 235 months, for the analysis. Several assets are observed for shorter periods of time, leading to an average (median) of 138 (128) months. Calvet *et al.* (2007) abstract from mutual fund fees in their main analysis and explore the robustness of their result to incorporating the exact mutual fund fees for the 10 most popular mutual funds and applying average fees to the rest. I take the opposite path, incorporating mutual fund fees in the main part of the paper and checking robustness to excluding them. I could find information on the fees that 140 of these charge via Morningstar or a fund's prospectus. For another 20 mutual funds, I assigned the fee of similar funds managed by the same company. I imputed the fees for the remaining 10 funds from the distribution of available fees. The last line of Table 1 shows that fees are in the usual range with an annual average of 130 basis points. In the estimations which contain mutual fund fees, I subtract 30 basis points from the benchmark index, which approximates the fees charged by index funds replicating common benchmarks.

2.3 Construction of diversification measures

In order to reduce the return series data to single measures of portfolio efficiency, I follow the strategy of Calvet *et al.* (2007) rather closely and merely sketch it here in order to keep the paper self-contained. The interested reader is referred to Calvet *et al.* (2007), including the corresponding Online Appendix, for further details. In a first step, I decompose total portfolio risk into a systematic and an idiosyncratic component. All returns are framed as excess returns over the risk-free rate, which is approximated by the money market rate. The portfolio risk decomposition is based on a regression of the household portfolio's excess return $r^e_{h,t}$ on a benchmark's excess return $r^e_{b,t}$:

$$r^e_{h,t} = \alpha_h + \beta_h \cdot r^e_{b,t} + \varepsilon_{h,t}$$

I take the MSCI Europe index as the benchmark, the results are robust to using the excess returns of the AEX or the (unhedged) MSCI World Index instead. The decomposition of the household's total portfolio risk σ_h^2 into a systematic σ_b^2 and an idiosyncratic $\sigma_{h,\text{idios.}}^2$ component is then given by:

$$\sigma_h^2 = \beta_h^2 \cdot \sigma_b^2 + \sigma_{h,\text{idios.}}^2 \tag{1}$$

The advantage of this decomposition is that it is purely statistical, i.e. it does not involve any assumptions about asset pricing. The main drawback is that while a large amount of idiosyncratic risk-taking is a sign of inefficient investing, its magnitude is difficult to interpret.

Constructing diversification measures with a meaningful scale requires an estimate of expected returns. Directly estimating expected returns in each dataset would be problematic because of the short return histories for some assets; and because the time series cover different time spans.[8] Again, I follow Calvet *et al.* (2007) and assume that assets are priced according to a CAPM, where I take the MSCI Europe to proxy the efficient market portfolio. This choice seems natural for a member of the Eurozone. Net of the 30 basis points annual fee, the benchmark has an annual excess return $\mu_b = 5.75\%$ over the 1983-July 2009 period. Along with the standard deviation $\sigma_b = 16.7\%$ this leads to a Sharpe ratio $S_b = \mu_b/\sigma_b$ of 35%. Imposing the CAPM leads to the following regression for all assets $a = 1, 2, \ldots 269$:

$$r_{a,t}^e = \beta_a \cdot r_{b,t}^e + \varepsilon_{a,t}.$$

Given the betas of all assets and the portfolio weights for each household, it is straightforward to calculate the expected returns μ_h of the household portfolios. A first measure of diversification loss is the relative Sharpe ratio loss:

$$RSL_h = 1 - \frac{S_h}{S_b} \tag{2}$$

The relative Sharpe ratio loss relates the Sharpe ratio of the household portfolio to that of the benchmark. It equals zero for an efficient portfolio and one for a portfolio where all risk is idiosyncratic.

While the relative Sharpe ratio loss has a number of attractive features (see Calvet *et al.*, 2007), its usefulness is confined to risky assets. A poorly diversified risky portfolio will not lead an investor far astray from the efficient frontier if the share in risky assets is sufficiently low. The independence of RSL_h of the risky asset share thus is not necessarily desirable. Calvet *et al.* (2007) therefore define the return loss, which is the average return a household loses by not choosing a position on the efficient frontier with the same level of risk. I skip its derivation and directly report a (slightly simplified) version that is useful for decomposing it into various components:

$$RL_h = \mu_b \cdot \omega_h \cdot \beta_h \cdot \left(\frac{RSL_h}{1 - RSL_h} \right) \tag{3}$$

The return loss of the household portfolio is the product of the expected excess return on the market portfolio (which does not vary in the population), the risky asset share ω_h, the beta,

[8] To see this, assume that one observes two assets with identical moments. Data for the first is available in the 2000-2005 period and for the second from 2002 to 2007. The first asset would likely have a much lower estimated alpha because the market conditions were worse during the earlier period. Pricing assets via the CAPM avoids this problem as long as the correlation with the index does not change with market conditions.

and a nonlinear transformation of the relative Sharpe ratio loss.[9] In the mean-variance plane, the return loss is the vertical distance between the efficient frontier and the location of the household portfolio.

2.4 An investment production function

One of the big advantages of the CentERpanel/DHS data is that it allows, for the first time, to relate detailed diversification outcomes to covariates that are not typically available in administrative data. For example, the Swedish administrative data of Calvet *et al.* (2007, 2009b) contain measures of wealth, income, employment, age, household size, education, and immigration status. One of their main findings is that wealthier households invest both more aggressively and more efficiently. The data do not allow to discern whether this is because these households are able to buy better advice; or whether they take better financial decisions by themselves. The policy conclusions would be very different: In the former case, one would target the supply of investment advice. In case investor sophistication is the key, financial education programs could be of help (Tang *et al.*, 2010).

In order to clarify concepts, it is useful to think of the investment process in terms of a simple production function. The output is a measure of efficient investment, e.g. one of those considered in the previous subsection. A certainly non-exhaustive list of important inputs identified in the literature are financial literacy/knowledge, cognitive abilities and education, the source of financial advice, risk aversion, age, gender, and several others described below. I approximate the production function by a linear equation

$$Y = X'b + u. \tag{4}$$

The investment outcome Y is observed and relevant for the household as a whole, but many of the inputs in the vector X concern individuals. The DHS contains a variable asking about who takes financial decisions in the household on a five point-scale. If both partners agree on a financial decider, I use the inputs for this person. In case of ties (e.g. both partners stating that they have equal say), I use the inputs from the member identified as the household head. The results of Smith *et al.* (2010) provide some evidence that this approach is sensible. Analysing the correlation between cognitive skills and various economic outcomes for older households, separately for each partner, they show that numeracy of the financial respondent in the HRS data is by far the most important correlate.

Previewing the results, we shall see that the diversification loss is close to negligible for a large part of its distribution. However, similar to the Swedish case, losses become rather high in the upper tail. For this reason, I do not only estimate Equation (4) by OLS, but also by means of quantile regressions. An additional benefit of quantile regression is that it provides a direct way to incorporate non-participants in the estimations, provided that the diversification measure is well-defined for non-participants. This is the case for the return loss (3): $\omega_h = 0$ or $RSL_h = 0$ imply $RL_h = 0$. Note that the quantile under consideration needs to be strictly positive for all population groups, otherwise the estimator is not well defined. The typical way of including non-participants in a least squares regression would be to model (4) as a two-part process of first deciding whether to invest in risky assets and then

[9] The value of $\frac{RSL_h}{1-RSL_h}$ becomes extremely high if the expected return on a household's portfolio μ_h is close to zero. I therefore winsorise $\frac{RSL_h}{1-RSL_h}$ in the decomposition exercises below.

how to invest in them (see Pohlmeier and Ulrich (1995) for such a model in another context and Calvet *et al.* (2007) for an application to portfolio choice). Given that the participation decision has been studied extensively, including with the very data used here (van Rooij *et al.*, forthcoming), such an approach seems to be an unnecessary complication.

My analysis identifies subgroups of the population who are at an increased risk of obtaining inferior investment outcomes. It does not without further assumption follow that changing a covariate would lead to a change in Y corresponding to b. Nevertheless, the analysis is an important improvement over the state of the art because it permits to identify conditional relationships. For example, in a related contribution Korniotis and Kumar (2009) first regress cognitive skills in an auxiliary dataset on a number of covariates. They then use the estimated coefficients to predict a smartness score in the dataset containing investment outcomes. Such a procedure only permits the estimation of the bivariate relationship between "smartness" and investment outcomes and does not allow for separate effects of covariates entering the index. The analysis of Grinblatt *et al.* (forthcoming) shows the relation between a measure of cognitive skills and some measures of diversification for Finnish males; but the authors can neither condition on education at the individual level nor disentangle whether part of the relationship is mediated through financial advice.

2.5 Inputs to investment production

As discussed in the introduction, there has been a huge upsurge in studies that aim to measure the individual skills that enter the right hand side of (4). One reason for using the 2004 and 2005 portfolio data is that at that point in time, Maarten van Rooij, Annamaria Lusardi, and Rob Alessie fielded a battery of questions aimed at estimating respondents' financial literacy. The data form the basis of van Rooij *et al.* (forthcoming) and the authors kindly provided me with data and code. The questions are similar to those in Lusardi and Mitchell (2007b), they are discussed in detail and compared to other measures in Hung *et al.* (2009). A first set of questions, coined basic financial literacy, contains 5 quiz-like simple math problems. A good example is the numeracy question: *Suppose you had €100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow? (i) More than €102; (ii) Exactly €102; (iii) Less than €102; (iv) Do not know; (v) Refusal.* Table A.2 in the Appendix shows that 95% of respondents correctly answer this question. The other questions have a similar structure of relatively simple math problems, but correct response rates are lower. Similar to van Rooij *et al.* (forthcoming), I assume that a single factor is underlying the five questions and normalise this factor to have zero mean and unit variance. The only difference is how I treat "Do not know" and "Refusal" answers. Instead of coding another variable, which will have a complicated correlation structure with the "substantive" answer, I assign those answers the probability of a random guess being correct, e.g. 1/3 for the question above. This can be rationalised by a linear probability model and would be exactly correct if all factor loadings were equal to each other. In the present case, it will be a reasonable approximation. All results are robust to using the exact procedure of van Rooij *et al.* (forthcoming). The basic financial literacy measures whether individuals possess the necessary cognitive abilities to perform simple numerical computations, which will be important for informed financial decision-making. Indeed, the survey instrument resembles to some extent the numeracy component of the cognitive ability score used in Christelis *et al.* (2010b) to explain stockholding. The first row of Table 2 shows that participants in risky asset markets have a significantly higher basic financial literacy score, confirming one of the main results of van Rooij *et al.* (forthcoming) and Christelis *et al.* (2010b). The distribution of the index is left-skewed since 45% of respondents in the entire sample answer all questions correctly, leading to a maximum index value of 0.63.

The second part of the financial literacy module, termed advanced financial literacy, consists of questions relating to knowledge of financial instruments and concepts. For example, the following question asks about the diversification properties of stocks and mutual funds: *Buying a company stock usually provides a safer return than a stock mutual fund. True or false? (i) True; (ii) False; (iii) Do not know; (iv) Refusal.* Only two thirds of all respondents give the correct answer, but three quarters of those who participate in risky asset markets. The advanced financial literacy index, constructed in the same way as the basic literacy index, also takes on higher values on average for holders of risky assets (see row 2 in Table 2) – again as in van Rooij *et al.* (forthcoming). Conceptually, these financial knowledge questions might be more problematic as inputs in (4) than the basic math skills in the basic module. The reason is that they may be largely shaped by investor experience – one would expect an increase in the probability of a correct answer to the example question for somebody who has monitored the evolution of stock and mutual fund returns in his or her portfolio for a while. This is less problematic for analysing efficiency of the risky portfolio than for studying the participation decision, but it remains a concern. The same comment applies to self-assessed financial knowledge, which is the third variable aimed at measuring financial literacy. I dichotomise the four-point rating into a dummy variable, which equals one for 19% (28%) of all households (participants in risky asset markets).

Abstracting from agency problems and potential costs, rational households who realise their lack of investment skills would seek external help.[10] The most important source of financial advice is directly asked for in the DHS questionnaire. The second part of Table 2 shows that about a quarter of respondents seek help from professional advisors and that another quarter rely upon the advice of family and friends. The remaining half is made up of a number of categories: Newspapers; financial magazines; guides; books; brochures from the bank or mortgage advisor; advertisements; financial computer programs; the Internet; other. I label the aggregate category "reliance on own financial judgement".[11] Their percentage rises among participants in risky asset markets, entirely at the expense of those who turn to their friends and family for financial advice. It is especially interesting to compare the level of financial literacy among the different groups of advice-seeking. Among all respondents, it is significantly lower among those who ask their friends and family compared to any of the two other groups. The same pattern remains for those with risky assets, although the sample size compromises statistical significance.

The remaining inputs to the production function are additional controls which serve to sharpen the interpretation of the financial literacy and advice variables. First, the financial literacy variables could merely be an approximation for education if it was not controlled for. Another angle to look at the relation is that to extent that education signals cognitive abilities, one would like to know whether specific (i.e. the basic literacy index) or general

[10] Hackethal *et al.* (2011) show that professional advice not necessarily leads to better outcomes, however.

[11] It is debatable whether those who cite brochures from financial institutions as their most important source of advice should rather be added to the category of professional financial advisors. Presumably, a financial institution's advisors and brochures would recommend similar investment strategies. I prefer the classification I chose because brochures seem to focus on advertising specific investments, while advisors would (hopefully) make recommendations based on the entire portfolio. In any case, all results survive a reclassification.

Table 2: Descriptive statistics on the covariates

Variable name	Entire sample		Portf. returns avail.	
	Mean	Std. dev.	Mean	Std. dev.
Basic fin. literacy index	0.000	1.000	0.249	0.765
Advanced fin. literacy index	0.000	1.000	0.563	0.756
High self-rated fin. knowledge	0.189	.	0.280	.
Financial advice: Professionals	0.230	.	0.247	.
Financial advice: Family/friends	0.238	.	0.131	.
Financial advice: Own judgement	0.533	.	0.622	.
Prof. advice * bas. literacy	0.046	0.985	0.201	0.943
Advice fam./friends * bas. literacy	-0.184	1.040	0.143	0.691
Own fin. jugdement * bas. literacy	0.071	0.978	0.297	0.712
No/elementary/secondary education	0.584	.	0.438	.
Higher vocal education	0.267	.	0.318	.
Academic education	0.149	.	0.244	.
Age 26-40	0.300	.	0.185	.
Age 41-64	0.481	.	0.538	.
Age 65+	0.219	.	0.277	.
Female	0.208	.	0.159	.
High tolerance for risky investm.	0.000	1.000	0.461	0.951
Household size	2.365	1.301	2.386	1.330
Degree of urbanisation	0.000	1.000	-0.002	0.987
Net annual household income	31,711	37,499	40,174	63,195
Log net household income	10.215	0.492	10.413	0.513
Value of total financial assets	41,513	76,736	84,845	114,436
Log financial assets	9.700	1.399	10.713	1.185
Value of total non-financial assets	173,053	198,047	238,496	228,589
Log total non-fin. assets	10.222	2.931	11.036	2.608
Value of total debt	55,242	82,810	66,270	91,551
Log total debt	7.458	3.969	7.986	3.993

Source: CentERpanel, own calculations. All statistics are adjusted for sampling weights, standard deviations of dummy variables are not shown. Variables relating to individuals rather than the household (i.e. all variables except for the last section of the table) are for the financial decider, as defined in Section 2.4. The number of observations where the covariates for the preferred specification (all covariates) are present is 958 (798) for the entire sample and 270 (238) for participants in risky asset markets with detailed portfolio information. For most covariates, the number is much closer to the relevant figures reported in column "# obs" of Table 1. The interaction terms in the third part of the table give averages of financial literacy within each category of financial advice.

abilities matter more. I include education in three categories and as expected, it is higher among those with risky assets in their portfolio. Second, cognitive functioning declines with age. However, age may have a positive effect on investor performance through experience (Korniotis and Kumar, forthcoming). Third, cognitive abilities have been shown to correlate with risk aversion (Dohmen *et al.*, 2010). On average, women are more risk averse than men (Croson and Gneezy, 2009), so I include a gender dummy. Furthermore, I use a measure of willingness to take financial risks derived from the degree of agreement with six different statements, each measured on a 7-point scale (e.g. *It is more important to have a safe investment with guaranteed returns than taking risk.* → *Totally disagree / Disagree / Partly disagree / Neither agree or disagree / Partly agree / Agree / Totally agree*). I add up the answers and standardise the resulting variable to have mean zero and unit variance. The bivariate correlations with risky asset holdings both go in the expected direction, see Table 2 once more. Including these variables in the regression has the drawback of reducing the sample size by about 15%. Hence, I do not include the risk aversion measures in my preferred specification and relegate the tables with added variables to the Appendix.

In order to compare my results to those of Calvet *et al.* (2007), I furthermore include measures of household size, the degree of urbanisation, household income, wealth, and liabilities in various additional specifications. The reason for not incorporating these variables in my standard specification is that their interpretation in the production function framework (4) is not obvious. Almost all explanations would go through abilities (e.g. smart individuals would have higher labour earnings and better investment outcomes) or financial advice (e.g. for rich households professional advice might be cheaper relative to asset volume). Again, I discuss the results in the text and all corresponding tables can be found in the Appendix.

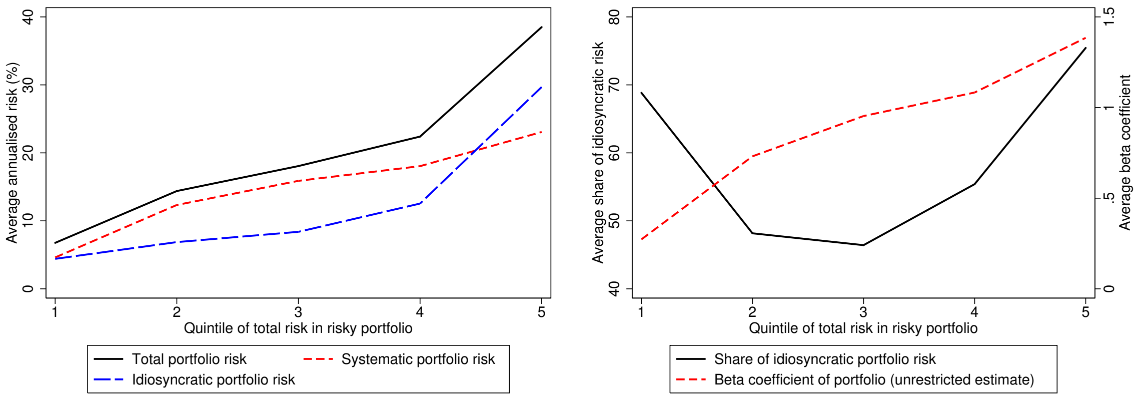
3 Results

3.1 The distribution of efficiency measures

For all participants in risky asset markets, Figure 2 contains plots of several measures for each quintile of the distribution of total portfolio risk, as inferred from equation (1). Shown in the left panel, total portfolio risk rises from less than 10% annually in the lowest quintile to almost 40% in the top quintile, with the most pronounced rise at the top. The systematic component moves almost in parallel for the first three quintiles, only then its slope becomes much flatter. Accordingly, the idiosyncratic component shows its steepest increase at the top of the portfolio risk distribution, suggesting that inefficient investing is by far the strongest there. The numbers are remarkably close to those found by Calvet *et al.* (2007) for Swedish households – they report 11% (19.5%, 36.4%) at the 10^{th} (50^{th}, 90^{th}) percentiles for total portfolio risk.[12] They also find the same U-shaped pattern for the idiosyncratic risk share displayed in the right panel, again with similar magnitudes. In both countries, the high values at the lower end of the distribution are driven by bond mutual funds, which display a low correlation with the benchmark index. This is also reflected in the average beta coefficient inferred from (1), which rises strongly in total portfolio risk. Again, the magnitudes are very close to those reported in Calvet *et al.* (2007).

[12]Given the huge sample size, Calvet *et al.* (2007) calculate averages around specific percentiles, which enables them to go much further into the tails of the distribution. I compare their reports for the midpoint of quintiles to the quintile-specific averages calculated in my analysis.

Figure 2: Portfolio risk components by quintile of total portfolio risk

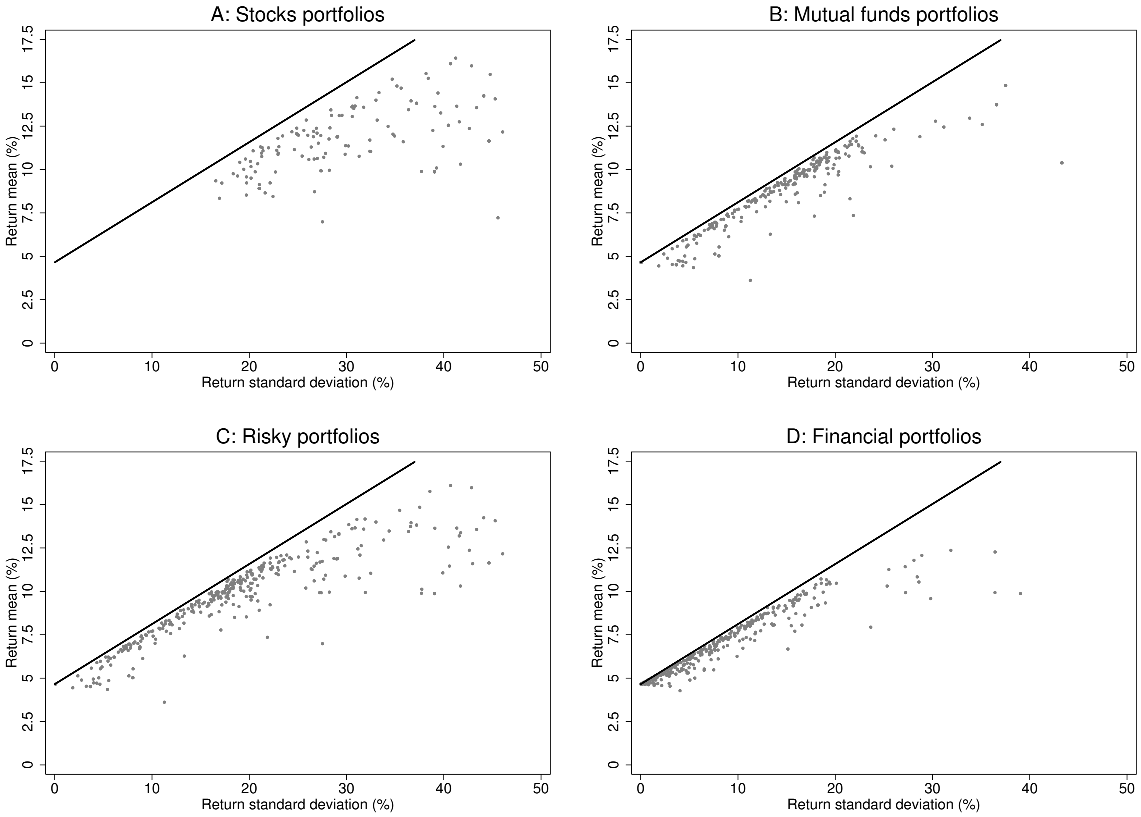


Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details.

In order to get an understanding of the basic characteristics of household portfolios, it is useful to plot them in the mean-variance plane. Panel A of Figure 3 does this for the pure stock portfolios and reveals a picture of strong underdiversification, which is similar to the findings of Calvet *et al.* (2007) for Sweden and Goetzmann and Kumar (2008) for the U.S.. The mutual fund component of households' portfolios appears much better diversified, even though the CAPM is applied after subtracting mutual fund fees. The major part of the distribution lines up right below the efficient frontier (Panel B of Figure 3). Nevertheless, a substantial fraction of mutual funds perform significantly worse than the market portfolio, conditional on the level of risk. Panel C contains the aggregate of stocks and mutual funds and shows that many households reduce the risk of their stock portfolios by additionally investing in mutual funds (of all risky asset owners, 55% only own mutual funds, 18% only own stocks, and 26% own both). The picture is yet more positive when holdings of safe assets are taken into account in Panel D of Figure 3. Relatively few outliers with severe losses are left at high levels of risk, but there is a number of households with a portfolio that is 1-2 percentage points below the efficient frontier at relatively low levels of risk. Diversification losses of this magnitude will be substantial when accumulated over the life-cycle (Calvet *et al.*, 2007; Tang *et al.*, 2010).

In order to allow for an easier interpretation and to perform quantitative analyses, it is useful to reduce the 2-dimensional information in Figure 3 to a single dimension. This is the purpose of the relative Sharpe ratio loss (2) and the return loss (3) presented in Section 2.4. Their quintile-specific values are plotted in Figure 4. The relative Sharpe ratio loss, shown in the left panel, is limited to far less than 20% in the bottom three quintiles, before reaching 27% and 64% in the upper quintiles. Again, this pattern mirrors the findings of Calvet *et al.* (2007) very closely: Most households largely avoid inefficient risk-taking, but almost two thirds of all the risk the average household in the top quintile takes remains uncompensated. The right panel of Figure 4 contains various measures of return loss. The solid black line just considers the risky portfolio, i.e. it is the vertical distance between the location of a household in Panel C of Figure 3 and the efficient frontier. Put differently, the risky asset share ω_h in (3) is set to one. The average return households lose on their risky portfolio compared to an efficient investment equals 180 basis points per year, which is just above the number reported by Calvet *et al.* (2007) for the unhedged world index as the benchmark. It

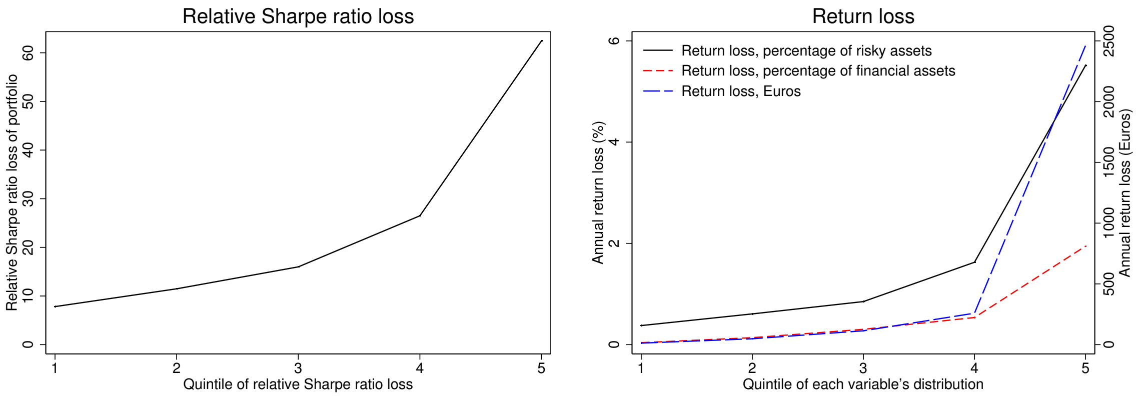
Figure 3: The mean-variance characteristics of household portfolios



Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details.

is below 100 basis points for the bottom three quintiles and 585 basis points in the highest. These return losses are far lower when the entire portfolio is taken as the basis – again, they are very limited for the first four quintiles (56 basis points in the fourth), but reach the substantial amount of more than 2% in the highest quintile. The average is about 59 basis points, substantially less than the 180 basis points for the risky portfolio multiplied with the risky asset share of .39 (see Figure 1). This implies a negative covariance between the risky asset share and the losses from underdiversification multiplied with the household portfolio's beta.[13] This illustrates the limited usefulness of the relative Sharpe ratio loss for assessing the diversification losses incurred on the entire portfolio – on average, those with higher values of RSL_h have a smaller share in risky assets, so the losses are less important for them. Finally, the third line in the right panel of Figure 4 demonstrates that, again as in Sweden, the losses are by no means negligible in monetary terms for substantial parts of the population.

Figure 4: Mean-variance measures of diversification losses.



Source: CentERpanel, Datastream, Euroinvestor, own calculations.

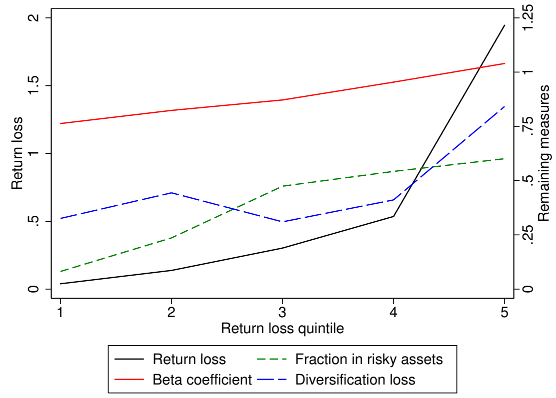
Finally, I plot the various components of the return loss as exemplified by the right-hand-side of (3) by its quintiles: The fraction in risky assets ω_h, the beta coefficient on the household portfolio, and the transformation of the relative Sharpe ratio loss $\frac{RSL_h}{1-RSL_h}$. Of course, the different components do not add up due to Jensen's inequality, so the graphical illustration does not qualify as a decomposition. Nevertheless, it remains useful to get a rough idea of the underlying mechanisms. The beta coefficient rises almost linearly over the quintiles, so inefficient and efficient risk-taking at least go hand in hand on average. The risky asset share increases quickly until the middle of the return loss distribution and modestly afterwards. The diversification loss is fairly constant in the lower quintiles (implying relative Sharpe ratio losses between 24% and 30%) and increases strongly in the top quintile (implied $RSL_h = .48$). Compared to the lower quintiles, the prime driving force behind the highest return losses thus seems to be uncompensated risk taking.

3.2 How do investment outcomes vary in the population?

The previous section has shown that the descriptive results of Calvet *et al.* (2007) can be replicated to a large extent for another country and, more relevantly, on a dataset that is

[13]This finding is confirmed by statistical analysis; and it is also true for the covariance between the relative Sharpe ratio loss and ω_h.

Figure 5: Return loss and its components by quintile



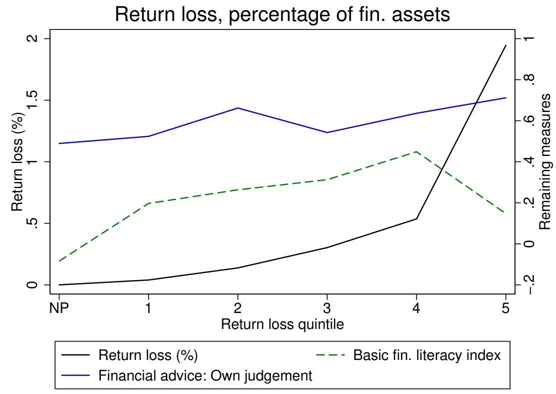
Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details.

fairly easy to collect anywhere. While this is an important step, the big advantage of the CentERpanel data is the availability of inputs to the investment production function (4) that have been of particular interest in the literature. Figure 6 once more plots the quintile-specific averages of the return loss, adding non-participants in risky asset markets (for whom $RL_h = 0$ since $w_h = 0$ and $\frac{RSL_h}{1-RSL_h} = 0$). For each return loss quintile, the figure furthermore adds the average values of the basic financial literacy index and the share of households relying on their own financial judgement. The financial literacy index shows an inverse U-shaped pattern – it is lowest for the non-participants (previously shown by van Rooij *et al.* (forthcoming) and Christelis *et al.* (2010b)), rises monotonously until the fourth return loss quintile, before dropping to its second-lowest value in the top quintile. In conjunction with the fact that diversification losses seem to be the driving force behind the largest return losses (Figure 5), this suggests that low investment skills may play an important role in determining the worst outcomes. The same inverse U-shaped pattern is found for the amount of idiosyncratic risk or the relative Sharpe ratio loss, see Figure B.1 in the Appendix. The fraction of individuals relying on their own financial judgements is generally rising in the return loss, although the high value in the second quintile is an exception to the rule. While the bivariate relations are suggestive, a more formal analysis is required to shed light on potential mechanisms.

The first column of Table 3 shows the results for an OLS regression of my preferred set of covariates on the sample of participants. The first three rows, containing the basic financial literacy index, the advice variable, and their interaction, already contain the basic result of my analysis.[14] Financial literacy does not have an effect for those who seek external advice – the coefficient in the first row is close to zero (2.4 basis points per year) and precisely estimated (the 95% confidence interval ranges from -7.2bp to 12bp). The dummy for deciding on the basis of self-collected information takes on a large and significantly positive value – those relying on their own judgement with a financial literacy score of zero on average incur a return loss that is 48 basis points higher than those who rely on external advice. The interaction term shows that this effect is much worse for those with negative values of the financial literacy index and that it almost exactly cancels out for those who achieve the highest financial literacy score. These households are estimated to incur an insignificant extra return

[14]The other financial literacy measures did not turn out to have an effect and results for the two left-out groups of financial advice were very similar. More extended specifications are discussed below.

Figure 6: Financial literacy, financial advice, and diversification losses



Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. The return loss quintile value "NP" stands for non-participants.

loss of 48.4bp − 73.7bp × .63 = 1.9bp on average, compared to those seeking advice from professionals or family/friends with the same level of financial literacy. The coefficients on all other covariates are insignificant and much smaller than those for financial literacy and advice.

The results of the quantile regressions shown in the remaining columns of Table 3 are yet more interesting because they show that the averages are entirely driven by effects in the top third of the return loss distribution. None of the percentiles varies much with the level of financial literacy among those seeking external advice. All else equal, the 90$^{\text{th}}$ percentile of the return loss is 148 basis points higher among those who rely on their own financial judgement and have a financial literacy index of zero. Again, the effect becomes much worse for negative values of the financial literacy index and reduces to 40bp for those with the maximum financial literacy score. The same pattern holds for the 70$^{\text{th}}$ percentile, although the magnitudes are substantially smaller. The variation of the coefficients across the quantiles is significant because it shows that (a) most households achieve reasonable investment outcomes regardless of their characteristics and (b) the worst outcomes are concentrated among those who neither seek advice from other individuals nor have a high level of numerical skills.

Finally, one should note that there appears to be an age effect that is hidden in the OLS estimates. Return losses are significantly higher around the middle of the distribution for the oldest age group (age 65+) and there is a large negative coefficient for the 90$^{\text{th}}$ percentile, although it is not significant. The same pattern prevails in the middle age group, but all coefficients are insignificant. The coefficients on education are either tiny or point in the expected direction with none of them being significant. Last, females incur larger return losses at the higher quintiles, although they are significant only at one quantile.

The estimates reported in Table 3 are valid for the sample of participants in risky asset markets, but they may be different in the general population. To see this, assume there are two groups in the population. One group's members mostly stay out of risky assets and the remaining members invest very inefficiently. The second group fully participates and invests quite efficiently. Conditioning on participation will lead to all quantiles being higher for the

Table 3: Contributors to return loss

	OLS	p_{10}	p_{30}	p_{50}	p_{70}	p_{90}
Basic fin. literacy index	0.0335	-0.0156*	0.0151	0.0522*	0.0631	0.216
	(0.67)	(-1.70)	(0.36)	(1.69)	(1.20)	(0.89)
Financial advice: Own judgement	0.507**	0.00600	0.0443	0.0482	0.389***	1.541***
	(2.35)	(0.38)	(0.94)	(1.15)	(4.67)	(2.72)
Own fin. jugdement * bas. literacy	-0.733*	0.0148	-0.0146	-0.0154	-0.496***	-1.789***
	(-1.91)	(1.04)	(-0.26)	(-0.36)	(-5.38)	(-2.95)
Higher vocal education	0.0716	0.00293	-0.0144	-0.00195	-0.0949	-0.330
	(0.48)	(0.16)	(-0.30)	(-0.05)	(-1.12)	(-0.58)
Academic education	-0.0244	0.000249	-0.00647	-0.0886*	-0.0305	-0.248
	(-0.18)	(0.01)	(-0.11)	(-1.71)	(-0.30)	(-0.37)
Age 41-64	-0.156	0.00470	0.0375	0.0656	-0.0556	-0.519
	(-0.77)	(0.20)	(0.63)	(1.20)	(-0.53)	(-0.72)
Age 65+	-0.0640	0.0341	0.131**	0.262***	0.158	-0.413
	(-0.29)	(1.38)	(2.03)	(4.33)	(1.40)	(-0.53)
Female	0.230	0.0304	0.0214	0.158***	0.180	0.705
	(1.21)	(1.39)	(0.35)	(2.75)	(1.54)	(0.82)
Constant	0.453**	0.0147	0.0734	0.141**	0.382***	1.252*
	(2.21)	(0.73)	(1.13)	(2.35)	(3.42)	(1.94)
Observations	441	270	270	270	270	270
Adjusted R^2	0.114					

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. The OLS regression has been estimated on both waves of data and standard errors are clustered at the household level; the quantile estimates are based on a pure cross-section. All regressions use sampling weights.

first group compared to the second group. Not doing so will reverse the order except for the highest quantiles. Table 4 presents the results of including the non-participants in the estimation sample, who have a return loss of zero. Remember from Section 2.4 that the quantiles under consideration need to be strictly positive for the estimator to be well-defined, so results of this exercise are presented in for every fifth percentile starting with the 75$^{\text{th}}$.

Except for the female dummy, the estimates in the lower quantiles are all positive, albeit relatively small. This is an almost mechanical consequence of the different characteristics of participants and non-participants (compare the first and third column of Table 2). Unless someone invested in the efficient market portfolio directly, the return loss will be positive for participants. There does not seem to be any effect of financial literacy for those seeking others' advice. The "self-deciders" with a financial literacy score of zero have consistently higher return losses at every quantile considered, reaching magnitudes of more than 100 basis points at the 95$^{\text{th}}$ percentile. The interaction effect is small and mostly insignificant over the first 4 quantiles considered. It is back to the previous interpretation for the highest quintile. Interestingly, the education variables are positive over the entire distribution and significantly so in the lower part. The more educated invest more aggressively, but education does not lead to very efficient risk-taking. The same comment applies to the highest age group – at each quantile under consideration, their return loss is substantially higher than that of the

Table 4: Contributors to return loss, including non-participants

	p_{75}	p_{80}	p_{85}	p_{90}	p_{95}
Basic fin. literacy index	0.00900	0.0134	0.0223	0.0522	0.106
	(0.70)	(0.86)	(0.76)	(0.88)	(0.56)
Financial advice: Own judgement	0.103^{***}	0.116^{***}	0.180^{***}	0.199^{**}	1.045^{***}
	(6.04)	(4.48)	(3.52)	(2.06)	(2.96)
Own fin. jugdement * bas. literacy	0.0336^{*}	0.0380	0.0603	0.0459	-1.210^{***}
	(1.86)	(1.55)	(1.30)	(0.55)	(-3.60)
Higher vocal education	0.116^{***}	0.153^{***}	0.127^{**}	0.188^{*}	0.336
	(6.02)	(5.33)	(2.25)	(1.73)	(0.86)
Academic education	0.0980^{***}	0.200^{***}	0.173^{**}	0.158	0.438
	(4.00)	(5.44)	(2.35)	(1.13)	(0.84)
Age 41-64	0.0247	0.0298	0.0750	0.0984	0.0609
	(1.15)	(0.93)	(1.20)	(0.88)	(0.14)
Age 65+	0.183^{***}	0.210^{***}	0.335^{***}	0.427^{***}	0.405
	(7.38)	(5.63)	(4.60)	(3.27)	(0.85)
Female	-0.0183	-0.0381	-0.0428	-0.0800	-0.0356
	(-0.87)	(-1.17)	(-0.67)	(-0.67)	(-0.08)
Constant	0.0146	0.0400	0.0664	0.156	0.355
	(0.69)	(1.26)	(1.05)	(1.39)	(0.92)
Observations	875	875	875	875	875

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. All estimates are based on a cross-section of households, including those who do not own any risky assets. For this reason, the coefficients are only estimated at higher quantiles where the return losses are strictly positive for all groups. All regressions use sampling weights.

youngest age group. Females are estimated to have lower return losses than males at each quantile of the distribution, although the coefficients are small and none of them is significant. Nevertheless, it nicely illustrates the differences that may arise from estimations on the two samples. Taken at face value, the coefficients in Table 3 imply that participating females incur a higher return loss conditional on the other covariates. But the results in Table 4 show that in the entire population another effect dominates: Many stay out of risky asset markets, potentially because they realise their lack of skills. This is well in line with overconfidence being stronger among men (Barber and Odean, 2001).

3.3 Robustness of the previous results

In appendices A.2 (regressions on the sample of participants) and A.3 (regressions on all households), I consider robustness to including a number of additional covariates in regressions similar to those underlying Tables 3 and Tables 4. In Tables A.3 and A.7, I add the advanced financial literacy index, self-rated financial knowledge, and the willingness to take financial risk. None of the extra financial literacy variables has an effect for the selected sample, which is also true for their interactions with financial advice in an unreported regression. In the entire sample (Table A.7), the previously mentioned "mechanical" effect occurs for the advanced literacy index as well. In any case, as mentioned in Section 2.5, the coefficients are difficult to interpret when the participation decision is involved because knowledge about assets' characteristics likely comes from experience with these assets.

As expected, higher risk appetite goes hand in hand with higher return losses across the distribution in both samples. The variable is strongly significant and leads to a doubling of the adjusted R^2 in OLS the specification of Table A.3 compared to Table 3 (leaving out the other extra variables does not seriously change the adjusted R^2). It is insignificant at the highest quintiles, which may be a sign that the most inferior outcomes are driven to a larger extent by mistakes than by more aggressive and relatively efficient investing. The magnitude of the coefficients remains non-negligible, however, so one should not interpret too much into this. Including the additional variables sharpens the age effect in both samples and does the same for the female dummy in the sample of participants. The coefficients on education do not change in the selected sample. Interestingly, they become much smaller and even reverse for the higher quantiles of the overall population – much of the added return loss in Table 4 appears to be driven by increased risk tolerance of the more educated. Most importantly, though, none of the main results regarding financial literacy and advice is altered by adding the covariates.

The same is true for the specification in Tables A.4 and A.8, which adds only the question on diversification from the advanced financial literacy battery to the variables used in Tables 3 and 4, since I expect this to have the strongest relation with my outcome. As expected, all coefficients are negative, they are significant for the OLS regression and at the 70^{th} percentile. The picture reverses for the general population, where all coefficients are positive and significantly so up to the 85^{th} percentile. This suggest that staying out of risky asset markets may largely be a rational decision: Little knowledge is associated with worse outcomes conditional on participation; more of the uninformed households staying out leads to smaller return losses for that group in the lower part of the unconditional distribution.

Tables A.5 and A.9 show that separating the groups relying on advice from professionals and family/friends does not make a difference. The basic financial literacy index matters only for for the "self-deciders" (left-out category) in the first row – for the other two groups it cancels out with the interaction terms. Both dummies are negative with comparable magnitudes. My main results also continue to hold when adding household size, urbanisation, and financial variables to the preferred specification in Tables A.6 and A.10. The only noteworthy feature is that the wealth-gradient in risky asset ownership (Figure 1) is reflected in higher return losses across the distribution in Table A.10. Conditional on participation, nothing remains, except for a tiny effect at the lowest decile (Table A.6).

As in Calvet *et al.* (2007), the treatment of mutual fund fees does not lead to different conclusions. Tables A.11 and A.12 in Appendix A.4 show that only minor numerical differences arise when re-doing the analysis without subtracting those fees.

Calvet *et al.* (2007) take the logarithm of (3) as the dependent variable in their regression analyses, which allows them to decompose the OLS regression coefficients into the average contributions of risky asset share, portfolio beta, and diversification loss. So far, I have not used the same approach. The first reason is that the decomposition is only valid for OLS and does not carry over to the quantile regressions. Arguably, the latter yielded the more interesting results: We have seen that the most important effects happen at the top of the distribution of the return loss, whereas the outcomes in the lower-middle parts of the distribution do not vary much with observable characteristics. Secondly, I believe that the scaling of the logarithmic regressions is less useful than a direct estimate. As is clear from Figure 4, the return loss on the financial portfolio is limited to very low numbers for large parts of the population, which are of little concern in welfare terms. On the logarithmic scale, however, the return losses of 4bp at the 10$^{\text{th}}$ percentile and of 10bp at the 25$^{\text{th}}$ percentile are almost as far apart from each other as the 61bp at the 75$^{\text{th}}$ percentile and the 149bp at the 90$^{\text{th}}$ percentile. The latter are much more relevant and should have a greater impact on the regression results. Nevertheless, the logarithmic decompositions are interesting for a comparison of results and the estimates are presented in Appendix A.5.

Table A.13 contains the results from a regression model that mirrors Table 5 in Calvet *et al.* (2007) as closely as possible. Only the age coefficient in the first column is significant and positive. The lack of significance of most coefficients does not come as a surprise when comparing the sample sizes and the fact that the R^2 is the same. The maximum t-statistic reported in Table 5 of Calvet *et al.* (2007) is 29.3. For a sample size of 373 (the reported 589 adjusted for clustering), this would reduce to 1.79. The remaining coefficients on age show that the increased return loss seems to be a combination of a larger risky share and inefficient investing. Finally, the significantly negative coefficient on the log of financial assets in column 4 confirms the result of Calvet *et al.* (2007) that wealthier households invest more efficiently on average – although in the Netherlands, there is no hard evidence that they invest more aggressively conditional on participation. Given the plots in Figure 1, this does not come as a surprise. Table A.14 contains the results for the preferred set of covariates. The basic message from the regression in levels holds up, although financial literacy now appears to have an effect also for those who turn to other persons for financial advice.

As a last robustness check, I consider a number of different measures of diversification loss as dependent variables in Appendix A.6. All results are broadly similar to those from Table 3. In particular, reliance on one's own financial judgement clearly leads to more uncompensated risk-taking, regardless of the diversification measure. When using the share of idiosyncratic risk as the dependent variable (Table A.15), the effect of financial literacy appears to be present for all households at the 90$^{\text{th}}$ percentile. At the 70$^{\text{th}}$ percentile, the coefficients are the same as before, although only borderline significant. Tables A.16 and A.17 contain the results for the return loss measured as the fraction of risky financial assets (i.e. setting $w_h = 1$) and the relative Sharpe ratio loss. As mentioned before, the right tails of both distributions are strongly affected by portfolios that are almost riskless but have a low correlation with the index on the risky part. I therefore exclude households whose portfolios have an annual standard deviation below 2% from the analysis. The results show the same pattern as before, although some coefficients are not significant.

Finally, Calvet *et al.* (2009b, see particularly the Online Appendix) consider a number of different measures of diversification which can be constructed using typically available survey data. The fraction of risky assets invested in shares as compared to mutual funds emerges as their preferred measure. Table A.18 contains the results of a set of regressions with this dependent variable. The OLS estimates only pick up a larger share of direct stock holdings among the "self-deciders", basic financial literacy is insignificant regardless of the main source of financial advice. The quantile regressions are meaningless except for the 7$^{\text{th}}$ decile, again only the advice dummy emerges as significant. In my data, the correlation between the relative Sharpe ratio loss and the share of direct stockholdings is only .22 compared to .49 reported by (Calvet *et al.*, 2009b).[15] Less detailed diversification measures might thus be

[15] Calvet *et al.* (2009a) consider the share of funds and the Sharpe ratio, which leads to exactly the same

helpful proxy variables in some cases, but when portfolio diversification is the outcome of interest, they do not permit to go into nearly as much detail as the measures employed by Calvet *et al.* (2007, 2009b) and in this paper.

A very similar index is constructed by Guiso and Jappelli (2009). In order to keep the dependent variable increasing in the degree of underdiversification, I take the inverse of their index and define it as $\frac{1-\alpha_h}{\min\{1, N_{h,\text{shares}}\}}$, where α_h is the fraction of the risky portfolio invested in mutual funds; and $N_{h,\text{shares}}$ is the number of shares held directly by the household. This index has a correlation of .34 with the relative Sharpe ratio loss and the results are displayed in Table A.19. Again, the OLS coefficient only picks up the higher diversification losses of those relying on their own judgement. However, the results at the 70$^{\text{th}}$ percentile now show the same effects as for the more general measures: Higher diversification losses for the "self-deciders", which get much worse with below-average financial literacy and almost disappear for those with the highest level of financial literacy.

4 Discussion and conclusions

My analysis has shown that detailed portfolio information can be obtained fairly easily from survey respondents. Although the merging of names to ISIN numbers and the corresponding return series is rather tedious, the resulting information is well worth the effort. The graphs in Section 3.1 show that the distributions of various diversification measures track those in Sweden very closely – which were obtained by Calvet *et al.* (2007) based on a dataset of unprecedented quality. As Sweden has abolished its wealth tax in the meantime, the data will not be updated anymore and the need for alternatives has become even more important. I have shown such an alternative, which is easily replicated anywhere in the world.

On the substantive side, my results show that the largest losses resulting from underdiversification are incurred by those who neither turn to external help with their investments nor have good skills in basic numerical operations and concepts. These effects are strong enough to drive average coefficients and they are robust to controlling for a number of covariates, including education level, age, financial knowledge in various forms, attitude to financial risk-taking, measures of wealth, and household income. These results are consistent with and refine those of Guiso and Jappelli (2009), who consider a financial literacy index that is a mixture of financial knowledge and numeracy skills and find a positive impact on portfolio diversification. Quantile regression analyses of the entire sample – including non-participants in risky asset markets – show that the main results hold up and suggest that non-participation is a response to a perceived lack of investment skill. This is consistent with Christelis *et al.* (2010b) and Grinblatt *et al.* (forthcoming), who find that cognitive abilities are an important contributor to participation in risky asset markets.

My results help to inform two related literatures. First, they provide a new angle to look at the question posed by Korniotis and Kumar (2009), whether *portfolio distortions reflect superior information or psychological biases.* Recently, van Nieuwerburgh and Veldkamp (2010) provided a theoretical rationale for the former in the sense that if information about returns is specific to a stock (industry), investors will hold a less than perfectly diversified portfolio. My results are consistent with such an explanation for portfolios in the middle-upper region of the return loss distribution. However, the fact that the highest return losses are

figures as using the inverse share and the relative Sharpe ratio loss.

incurred by households without high financial skills relying on their own financial judgement makes it unlikely that superior information was the driving force behind the choice of portfolios. Rather, as suggested by Kimball and Shumway (2010), they are more likely to reflect investment mistakes.

This directly leads to the issue of financial literacy, which has received a lot of attention recently, especially from the policy side (e.g Lusardi, 2010). My results suggest that the majority of Dutch household reach reasonably effective investment outcomes in terms of the risk-return trade-off, regardless of their level of financial literacy. Many of them achieve this by choosing a very low level of risk, others by turning to external help. Both strategies are consistent with a rational response to poor self-perceived investment skill. Corroborative evidence comes from Choi *et al.* (2010), who show that the (self-reported) likelihood of changing one's mind after consulting an investment advisor decreases in the quality of an experimental investment decision. The one group where the most severe investment mistakes occur are those individuals who neither seek external advice nor have a high level of financial-numerical skills. In other words, the overconfident: Consistent with Guiso and Jappelli (2006), these individuals trust their own capabilities more than those of others and seem to overestimate the former. The fact that two factors (low level of skills and deciding by oneself) seem to be present also suggests two potential starting points for policy interventions aiming to prevent the most inferior investment outcomes.

That the factor measuring financial-numerical skill turned out to be much more important than financial knowledge suggests that increasing the latter would not help much for portfolio outcomes. The nature of the questions in in van Rooij *et al.*'s (forthcoming) basic financial literacy index – very simple math quizzes worded in financial terms – suggests an interpretation as a subcomponent of cognitive functioning, which has also been shown to correlate strongly with the stock market participation decision (Christelis *et al.*, 2010b). It becomes increasingly difficult to compensate for low levels of cognitive skills after reaching adolescence (e.g. Cunha *et al.*, 2010, and the references therein), so it seems difficult to influence current generations through this channel. However, it suggests another reason why early interventions to foster the skills of disadvantaged children may be hugely beneficial. This assessment is also in line with the findings of Agarwal and Mazumder (2010); Grinblatt *et al.* (2011, forthcoming), who show that various financial mistakes are correlated with broad measures of cognitive functioning drawn from military qualifications tests.

This leaves the second channel, namely helping individuals get competent financial advice. This is more difficult than it seems at first sight because my estimates do not necessarily yield the causal effect of mandating advice. Indeed, in the experiments of Hung and Yoong (2010), only solicited advice had an effect on portfolio performance – unsolicited did not. Nevertheless, expanding the availability of external guidance seems to be the most promising route. Academic economist's typical advice of investing in low-fee index funds competes with many attempts to guide household's behaviour where the form of the message is designed by professional marketing forces, but the content is likely to suit to the needs of consumers less than optimally (Inderst and Ottaviani, 2009). Further research in that direction – how regulation can help shape correctly incentivised marketing forces – seems very promising in this light (also see Campbell *et al.* (2011) for an elaboration of this point and Suvorov and Tsybuleva (2010) for a first theoretical contribution in this direction). The point is reinforced by Hackethal *et al.*'s 2011 finding that reliance on professional advisors leads to higher portfolio turnover, consistent with typical incentive structures. While my analysis controls for annual fees charged by mutual funds, my cross-sectional data does not permit me to estimate the costs associated with portfolio rebalancing. Analogously to the interaction effects of financial advice and numerical-financial skills found for cross-sectional diversification measures in this paper, an important step for future research would be to collect a dataset that contains both inputs and detailed trading behaviour.

References

Agarwal, S. and Mazumder, B. (2010). Cognitive abilities and household financial decision making, available at SSRN, `http://ssrn.com/paper=1651312`.

Alessie, R., Hochgürtel, S. and van Soest, A. (2002). Household portfolios in the Netherlands. In L. Guiso, M. Haliassos and T. Jappelli (eds.), *Household Portfolios*, MIT Press, pp. 340–388.

—, — and van Soest, A. (2004). Ownership of stocks and mutual funds: A panel data analysis. *Review of Economics and Statistics*, **86** (3), 783–796.

—, — and van Soest, A. (2006). Non-take-up of tax-favored savings plans: Evidence from Dutch employees. *Journal of Economic Psychology*, **27** (4), 483–501.

Barber, B. M. and Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *The Journal of Finance*, **55** (2), 773–806.

— and — (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, **116** (1), 261–292.

Barberis, N., Huang, M. and Thaler, R. H. (2006). Individual preferences, monetary gambles, and stock market participation: A case for narrow framing. *American Economic Review*, **96** (4), 1069–1090.

Bayer, P. J., Bernheim, B. D. and Scholz, J. K. (2009). The effects of financial education in the workplace: Evidence from a survey of employers. *Economic Inquiry*, **47** (4), 605–624.

Bergstresser, D. and Poterba, J. (2004). Asset allocation and asset location: Household evidence from the survey of consumer finances. *Journal of Public Economics*, **88** (9-10), 1893–1915.

Bilias, Y., Georgarakos, D. and Haliassos, M. (2009). Equity culture and the distribution of wealth, available at `http://ssrn.com/abstract=731805`.

—, — and — (2010). Portfolio inertia and stock market fluctuations. *Journal of Money, Credit and Banking*, **42** (4), 715–742.

Blume, M. E. and Friend, I. (1975). The asset structure of individual portfolios and some implications for utility functions. *The Journal of Finance*, **30** (2), 585–603.

Calvet, L. E., Campbell, J. Y. and Sodini, P. (2007). Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, **115** (5), 707–747.

—, — and — (2009a). Fight or flight? Portfolio rebalancing by individual investors. *Quarterly Journal of Economics*, **124** (1), 301–348.

—, — and — (2009b). Measuring the financial sophistication of households. *American Economic Review*, **99** (2), 393–398.

Campbell, J. Y. (2006). Household finance. *Journal of Finance*, **61**, 1553–1604.

—, Jackson, H. E., Madrian, B. C. and Tufano, P. (2011). Consumer financial protection. *Journal of Economic Perspectives*, **25** (1), 91–114.

Choi, J. J., Laibson, D. and Madrian, B. C. (2010). Why does the law of one price fail? An experiment on index mutual funds. *Review of Financial Studies*, **23** (4), 1405–1432.

Christelis, D., Georgarakos, D. and Haliassos, M. (2010a). Differences in portfolios across countries: Economic environment versus household characteristics, available at `http://ssrn.com/abstract=1089802`.

—, — and — (forthcoming). Stockholding: Participation, location, and spillovers. *Journal of Banking & Finance*.

—, Jappelli, T. and Padula, M. (2010b). Cognitive abilities and portfolio choice. *European Economic Review*, **54** (1), 18–38.

Cole, S. and Shastry, G. K. (2009). Smart money: The effect of education, cognitive ability, and financial literacy on financial market participation, mimeo, Harvard Business School and University of Virginia.

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, **47** (2), 448–474.

Cunha, F., Heckman, J. J. and Schennach, S. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, **78** (3), 883–931.

Dimmock, S. G. and Kouwenberg, R. R. (2010). Loss-aversion and household portfolio choice. *Journal of Empirical Finance*, **17** (3), 441–459.

Dohmen, T., Falk, A., Huffman, D. and Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, **100** (3), 1238–1260.

Gerardi, K., Goette, L. and Meier, S. (2010). Financial literacy and subprime mortgage delinquency: Evidence from a survey matched to administrative data, Federal Reserve Bank of Atlanta Working Paper 2010-10.

Goetzmann, W. N. and Kumar, A. (2008). Equity portfolio diversification. *Review of Finance*, **12** (3), 433–463.

Graham, J. R., Harvey, C. R. and Huang, H. (2009). Investor competence, trading frequency, and home bias. *Management Science*, **55** (7), 1094–1106.

Grinblatt, M., Keloharju, M. and Linnainmaa, J. (2011). IQ and stock market trading, The University of Chicago Booth School of Business, Working Paper No. 09-33.

—, — and — (forthcoming). IQ and stock market participation. *Journal of Finance*.

Guiso, L., Haliassos, M. and Jappelli, T. (eds.) (2002). *Household Portfolios*. Cambridge, MA: MIT Press.

— and Jappelli, T. (2006). Information acquisition and portfolio performance, University of Naples, CSEF Working Paper No. 167.

— and — (2009). Financial literacy and portfolio diversification, University of Naples, CSEF Working Paper no. 212.

Hackethal, A., Haliassos, M. and Jappelli, T. (2011). Financial advisors: A case of babysitters?, mimeo, Goethe Universität Frankfurt. Available at `http://ssrn.com/abstract=1360440`.

Haliassos, M. and Bertaut, C. C. (1995). Why do so few hold stocks? *Economic Journal*, **105** (432), 1110–1129.

Heaton, J. and Lucas, D. (2000). Portfolio choice in the presence of background risk. *Economic Journal*, **110** (460), 1–26.

Hilgert, M. A., Hogarth, J. M. and Beverly, S. G. (2003). Household financial management: The connection between knowledge and behavior. *Federal Reserve Bulletin*, **89** (7), 309–322.

Hung, A. A., Parker, A. M. and Yoong, J. (2009). Defining and measuring financial literacy, RAND Labor and Population Working Paper 708.

— and Yoong, J. K. (2010). Asking for help: Survey and experimental evidence on financial advice and behavior change, RAND Labor and Population Working Paper 714.

Inderst, R. and Ottaviani, M. (2009). Misselling through agents. *American Economic Review*, **99** (3), 883–908.

Ivković, Z., Poterba, J. and Weisbenner, S. (2005). Tax-motivated trading by individual investors. *American Economic Review*, **95** (5), 1605–1630.

—, Sialm, C. and Weisbenner, S. (2008). Portfolio concentration and the performance of individual investors. *Journal of Financial and Quantitative Analysis*, **43** (03), 613–655.

Juster, F. T., Smith, J. P. and Stafford, F. (1999). The measurement and structure of household wealth. *Labour Economics*, **6** (2), 253–275.

Kimball, M. S. and Shumway, T. (2010). Investor sophistication and the home bias, diversification, and employer stock puzzles, mimeo, University of Michigan, available at `http://ssrn.com/abstract=1572866`.

Korniotis, G. M. and Kumar, A. (2009). Do portfolio distortions reflect superior information or psychological biases?, mimeo, University of Texas at Austin. Available at `http://ssrn.com/abstract=1018668`.

— and — (forthcoming). Do older investors make better investment decisions? *Review of Economics and Statistics*.

Lusardi, A. (2010). Americans' financial capability, Report Prepared for the Financial Crisis Inquiry Commission, available at `http://www.dartmouth.edu/~alusardi/policy.html`.

— and Mitchell, O. S. (2007a). Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics*, **54** (1), 205–224.

— and — (2007b). Financial literacy and retirement planning: New evidence from the RAND American Life Panel, University of Pennsylvania, The Wharton School, Pension Research Council Working Paper WP2007-33.

— and Tufano, P. (2009). Debt literacy, financial experiences, and overindebtedness, NBER Working Paper 14808.

Mankiw, N. G. and Zeldes, S. P. (1991). The consumption of stockholders and nonstockholders. *Journal of Financial Economics*, **29** (1), 97–112.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, **7** (1), 77–91.

Massa, M. and Simonov, A. (2006). Hedging, familiarity and portfolio choice. *Review of Financial Studies*, **19** (2), 633–685.

Odean, T. (1998). Are investors reluctant to realize their losses? *The Journal of Finance*, **53** (5), 1775–1798.

Pohlmeier, W. and Ulrich, V. (1995). An econometric model of the two-part decision-making process in the demand for health care. *The Journal of Human Resources*, **30** (2), 339–361.

Polkovnichenko, V. (2005). Household portfolio diversification: A case for rank-dependent preferences. *Review of Financial Studies*, **18** (4), 1467–1502.

Roussanov, N. (2010). Diversification and its discontents: Idiosyncratic and entrepreneurial risk in the quest for social status. *Journal of Finance*, **65** (5), 1755–1788.

Smith, J. P., McArdle, J. J. and Willis, R. (2010). Financial decision making and cognition in a family context. *The Economic Journal*, **120** (548), F363–F380.

Suvorov, A. and Tsybuleva, N. (2010). Advice by an informed intermediary: Can you trust your broker? *The B.E. Journal of Theoretical Economics*, **10** (1, Topics), Article 50.

Tang, N., Mitchell, O. S., Mottola, G. R. and Utkus, S. P. (2010). The efficiency of sponsor and participant portfolio choices in 401(k) plans. *Journal of Public Economics*, **94** (11-12), 1073–1085.

van Nieuwerburgh, S. and Veldkamp, L. (2010). Information acquisition and under-diversification. *Review of Economic Studies*, **77** (2), 779–805.

van Rooij, M. C. J., Lusardi, A. and Alessie, R. (forthcoming). Financial literacy and stock market participation. *Journal of Financial Economics*.

Vissing-Jørgensen, A. (2002). Towards an explanation of household portfolio choice heterogeneity: Nonfinancial income and participation cost structures, NBER Working Paper 8884.

A Additional tables

A.1 Detailed data description

Table A.1: Portfolio variables, descriptions

Level 0	Level 1	Level 2	Level 3	Level 4
Mutual funds	Mutual and growth funds	Risky financial assets	Total financial assets	Total assets
Growth funds	Mutual and growth funds	Risky financial assets	Total financial assets	Total assets
Shares	Shares	Risky financial assets	Total financial assets	Total assets
Company and mortgage bonds	Bonds and options	Risky financial assets	Total financial assets	Total assets
Options	Bonds and options	Risky financial assets	Total financial assets	Total assets
Checking account with positive balance	Checking and savings accounts	Safe financial assets	Total financial assets	Total assets
Savings and deposit accounts	Checking and savings accounts	Safe financial assets	Total financial assets	Total assets
Bank certificates and deposits	Checking and savings accounts	Safe financial assets	Total financial assets	Total assets
Saving certificates	Checking and savings accounts	Safe financial assets	Total financial assets	Total assets
Saving or endowment insurance policy	Cash value of insurances	Safe financial assets	Total financial assets	Total assets
Mortgage-related life insurance	Cash value of insurances	Safe financial assets	Total financial assets	Total assets
Life-cycle savings plan	Cash value of insurances	Safe financial assets	Total financial assets	Total assets
Single premium annuity insurance policy	Cash value of insurances	Safe financial assets	Total financial assets	Total assets
Boat	Durables	Non-financial assets (excl. real estate)	Total non-financial assets	Total assets
Car	Durables	Non-financial assets (excl. real estate)	Total non-financial assets	Total assets
Caravan	Durables	Non-financial assets (excl. real estate)	Total non-financial assets	Total assets
Motorbike	Durables	Non-financial assets (excl. real estate)	Total non-financial assets	Total assets
Business equity	Other non-financial assets	Non-financial assets (excl. real estate)	Total non-financial assets	Total assets
Money lent to family/friends	Other non-financial assets	Non-financial assets (excl. real estate)	Total non-financial assets	Total assets
Assets not mentioned in other categories	Other non-financial assets	Non-financial assets (excl. real estate)	Total non-financial assets	Total assets
Primary housing	Primary housing	Total real estate	Total non-financial assets	Total assets
Secondary housing	Total secondary real estate	Total real estate	Total non-financial assets	Total assets
Other real estate	Total secondary real estate	Total real estate	Total non-financial assets	Total assets
Credit-card debt	Total consumer credit	Total non-mortgage debt	Total debt	Total debt
Extended lines of credit	Total consumer credit	Total non-mortgage debt	Total debt	Total debt
Other consumer credit	Total consumer credit	Total non-mortgage debt	Total debt	Total debt
Private loan	Total consumer credit	Total non-mortgage debt	Total debt	Total debt
Hire purchase contracts	Total consumer credit	Total non-mortgage debt	Total debt	Total debt
Checking account with negative balance	Total consumer credit	Total non-mortgage debt	Total debt	Total debt
Student loan	Total consumer credit	Total non-mortgage debt	Total debt	Total debt
Borrowed from friends/family	Other debt	Total non-mortgage debt	Total debt	Total debt
Debts not mentioned in other categories	Other debt	Total non-mortgage debt	Total debt	Total debt
Mortgage on primary housing	Mortgage on primary housing	Total mortgage debt	Total debt	Total debt
Mortgage on secondary housing	Total mortgages on secondary real estate	Total mortgage debt	Total debt	Total debt
Mortgage on other real estate	Total mortgages on secondary real estate	Total mortgage debt	Total debt	Total debt

Table A.2: Detailed documentation of the financial literacy modules

Variable name	Entire sample		Portf. returns avail.	
	Mean	Std. dev.	Mean	Std. dev.
Simplest Numeracy (Dk = 1/3)	0.943	0.220	0.969	0.163
Interest Compounding (Dk = 1/3)	0.814	0.383	0.863	0.340
Inflation (Dk = 1/3)	0.889	0.291	0.948	0.210
Time Value Of Money (Dk = 1/3)	0.783	0.404	0.884	0.315
Money Illusion (Dk = 1/3)	0.721	0.441	0.745	0.432
What Does The Stock Market? (Dk = 1/4)	0.753	0.397	0.865	0.329
What Does Stock Ownerwhip Mean? (Dk = 1/4)	0.697	0.441	0.795	0.401
What Do Mutual Funds Do? (Dk = 1/4)	0.781	0.376	0.901	0.284
What Does A Company Bond Do? (Dk = 1/4)	0.681	0.421	0.809	0.371
Bond Prices After Int. Rate Change? (Dk = 1/4)	0.375	0.420	0.530	0.473
Diversification Stock Vs. Mut. Fund (Dk = 1/2)	0.658	0.414	0.767	0.400
Bonds Or Stocks More Risky? (Dk = 1/2)	0.771	0.355	0.884	0.295
Equity Premium (Dk = 1/3)	0.576	0.450	0.749	0.421
Volatility Different Assets (Dk = 1/3)	0.792	0.362	0.917	0.268
Diversification, Direct Question (Dk = 1/3)	0.746	0.396	0.870	0.324
Bond Liquidity (Dk = 1/2)	0.552	0.401	0.674	0.416

Source: CentERpanel, own calculations. All statistics are adjusted for sampling weights. The number of observations where the covariates for the preferred specification are present is 958 for the entire sample and 270 for participants in risky asset markets with detailed portfolio information. The exact wording of the questions can be found in **?**.

A.2 Alternative specifications for contributors to return loss

Table A.3: Contributors to return loss, adding advanced financial literacy and risk tolerance

	OLS	p_{10}	p_{30}	p_{50}	p_{70}	p_{90}
Basic fin. literacy index	0.00669	-0.0166	0.0176	0.0535	0.105**	0.207
	(0.10)	(-1.61)	(0.53)	(1.28)	(2.07)	(0.74)
Advanced fin. literacy index	0.0597	0.0219*	0.0190	0.0424	-0.0652	0.0812
	(0.39)	(1.95)	(0.76)	(0.95)	(-1.05)	(0.17)
High self-rated fin. knowledge	0.0205	-0.0178	-0.0157	-0.0236	0.0558	-0.153
	(0.18)	(-0.66)	(-0.42)	(-0.38)	(0.68)	(-0.27)
Financial advice: Own judgement	0.539**	-0.00705	-0.0123	0.155**	0.681***	1.709***
	(2.31)	(-0.32)	(-0.33)	(2.55)	(8.07)	(2.68)
Own fin. jugdement * bas. literacy	-1.004**	0.0148	0.00243	-0.252***	-1.201***	-2.337***
	(-2.35)	(0.88)	(0.06)	(-3.72)	(-11.93)	(-2.94)
Higher vocal education	-0.0542	-0.0300	-0.0271	0.00832	-0.0719	-0.457
	(-0.42)	(-1.28)	(-0.73)	(0.13)	(-0.87)	(-0.83)
Academic education	-0.0936	-0.00607	-0.0209	-0.0280	0.00210	-0.402
	(-0.72)	(-0.22)	(-0.51)	(-0.38)	(0.02)	(-0.62)
Age 41-64	-0.165	0.00648	0.0606	0.110	-0.0574	-0.411
	(-0.80)	(0.25)	(1.38)	(1.43)	(-0.56)	(-0.55)
Age 65+	0.0390	0.0381	0.164***	0.292***	0.244**	-0.189
	(0.18)	(1.25)	(3.35)	(3.37)	(2.12)	(-0.21)
Female	0.382*	0.0531*	0.0756	0.240***	0.219*	0.809
	(1.73)	(1.73)	(1.47)	(2.75)	(1.92)	(0.94)
High tolerance for risky investm.	0.141**	0.0291***	0.0461***	0.0602**	0.109***	0.101
	(2.23)	(3.42)	(2.92)	(2.11)	(2.79)	(0.35)
Constant	0.398*	0.0269	0.0493	0.0231	0.382***	1.160
	(1.77)	(1.19)	(1.02)	(0.27)	(3.12)	(1.49)
Observations	405	239	239	239	239	239
Adjusted R^2	0.200					

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. The OLS regression has been estimated on both waves of data and standard errors are clustered at the household level; the quantile estimates are based on a pure cross-section. All regressions use sampling weights.

Table A.4: Contributors to return loss, adding diversification questions

	OLS	p_{10}	p_{30}	p_{50}	p_{70}	p_{90}
Basic fin. literacy index	0.0761	-0.00972	0.0264	0.0667*	0.142***	0.131
	(1.30)	(-0.88)	(0.54)	(1.89)	(3.25)	(0.43)
Fin. lit. - diversification, direct question	-0.526**	-0.0270	-0.0817	-0.0864	-0.433***	-0.882
	(-2.20)	(-1.10)	(-1.08)	(-1.26)	(-4.36)	(-1.39)
Financial advice: Own judgement	0.524**	0.00813	0.0158	0.0432	0.452***	1.410***
	(2.46)	(0.46)	(0.30)	(0.93)	(6.96)	(2.63)
Own fin. jugdement * bas. literacy	-0.742*	0.0145	-0.00632	-0.0318	-0.578***	-1.524**
	(-1.97)	(0.88)	(-0.10)	(-0.67)	(-8.08)	(-2.49)
Higher vocal education	0.0803	0.00506	0.00491	-0.00772	-0.0948	0.0709
	(0.55)	(0.23)	(0.09)	(-0.16)	(-1.44)	(0.14)
Academic education	0.0388	-0.00110	-0.00711	-0.0822	-0.0168	-0.0341
	(0.29)	(-0.04)	(-0.11)	(-1.42)	(-0.21)	(-0.05)
Age 41-64	-0.148	0.00985	0.0203	0.0607	-0.0604	-0.0694
	(-0.76)	(0.35)	(0.31)	(0.98)	(-0.72)	(-0.12)
Age 65+	-0.0636	0.0365	0.149**	0.262***	0.154*	-0.0215
	(-0.29)	(1.26)	(2.07)	(3.87)	(1.69)	(-0.03)
Female	0.207	0.0285	0.0105	0.149**	0.171*	0.556
	(1.12)	(1.11)	(0.15)	(2.21)	(1.84)	(0.95)
Constant	0.876***	0.0360	0.152*	0.226***	0.756***	1.485*
	(2.98)	(1.01)	(1.69)	(2.60)	(5.88)	(1.86)
Observations	441	270	270	270	270	270
Adjusted R^2	0.137					

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. The OLS regression has been estimated on both waves of data and standard errors are clustered at the household level; the quantile estimates are based on a pure cross-section. All regressions use sampling weights.

Table A.5: Contributors to return loss, separate measures of financial advice

	OLS	p_{10}	p_{30}	p_{50}	p_{70}	p_{90}
Basic fin. literacy index	-0.699*	0.00195	-0.0126	0.0421*	-0.452***	-1.574***
	(-1.78)	(0.15)	(-0.46)	(1.79)	(-5.48)	(-3.14)
Financial advice: Professionals	-0.529**	-0.00532	-0.0510	-0.0302	-0.417***	-1.573***
	(-2.39)	(-0.26)	(-1.29)	(-0.85)	(-4.21)	(-2.90)
Prof. advice * bas. literacy	0.710*	-0.0159	0.0130	0.0103	0.512***	1.783***
	(1.82)	(-1.02)	(0.34)	(0.34)	(5.25)	(3.28)
Financial advice: Family/friends	-0.465*	0.00531	-0.0350	-0.0837*	-0.372***	-1.264*
	(-1.89)	(0.18)	(-0.80)	(-1.94)	(-3.07)	(-1.88)
Advice fam./friends * bas. literacy	0.802**	0.0562*	0.0810	0.0919*	0.667***	2.063**
	(2.07)	(1.67)	(1.65)	(1.93)	(4.39)	(2.59)
Higher vocal education	0.0720	0.0127	0.00724	-0.0159	-0.0725	-0.327
	(0.48)	(0.62)	(0.21)	(-0.48)	(-0.80)	(-0.65)
Academic education	-0.0219	0.00151	-0.00478	-0.0949**	-0.00814	-0.248
	(-0.16)	(0.06)	(-0.11)	(-2.38)	(-0.08)	(-0.41)
Age 41-64	-0.149	0.00532	0.0160	0.0552	-0.0381	-0.249
	(-0.71)	(0.22)	(0.37)	(1.30)	(-0.33)	(-0.39)
Age 65+	-0.0495	0.0353	0.178***	0.269***	0.158	-0.147
	(-0.21)	(1.29)	(3.65)	(5.72)	(1.26)	(-0.21)
Female	0.224	0.0285	0.0490	0.157***	0.162	0.705
	(1.16)	(1.15)	(0.99)	(3.59)	(1.26)	(0.92)
Constant	0.952***	0.0190	0.105**	0.197***	0.762***	2.523***
	(3.22)	(0.92)	(2.30)	(4.37)	(6.37)	(3.40)
Observations	441	270	270	270	270	270
Adjusted R^2	0.111					

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. The OLS regression has been estimated on both waves of data and standard errors are clustered at the household level; the quantile estimates are based on a pure cross-section. All regressions use sampling weights.

Table A.6: Contributors to return loss, adding household size, urbanisation, and financial variables to the preferred specification

	OLS	p10	p30	p50	p70	p90
Basic fin. literacy index	0.0511	-0.0120	0.0197	0.0470	0.0877	0.189
	(0.75)	(-1.08)	(0.58)	(1.00)	(1.42)	(0.43)
Financial advice: Own judgement	0.489**	0.00903	0.0400	0.0297	0.318***	1.454*
	(2.21)	(0.45)	(1.06)	(0.50)	(3.60)	(1.96)
Own fin. jugdement * bas. literacy	-0.736*	0.0186	0.0148	-0.0200	-0.416***	-1.720**
	(-1.91)	(1.02)	(0.33)	(-0.32)	(-4.14)	(-2.03)
Higher vocal education	0.0935	-0.00437	-0.00887	-0.0196	-0.0497	-0.133
	(0.68)	(-0.18)	(-0.24)	(-0.31)	(-0.54)	(-0.17)
Academic education	0.0161	-0.0270	-0.0473	-0.0996	0.0657	-0.0737
	(0.11)	(-0.98)	(-1.01)	(-1.34)	(0.57)	(-0.08)
Age 41-64	-0.104	-0.00963	0.0168	0.0681	0.0139	-0.169
	(-0.44)	(-0.33)	(0.33)	(0.83)	(0.12)	(-0.18)
Age 65+	-0.0943	0.0172	0.118**	0.226**	0.235*	-0.0584
	(-0.38)	(0.50)	(1.99)	(2.32)	(1.75)	(-0.06)
Female	0.131	0.0545**	0.0230	0.0763	0.0719	0.353
	(0.50)	(2.10)	(0.45)	(0.86)	(0.52)	(0.29)
Household size	-0.0269	-0.00606	-0.0166	-0.0209	0.00560	0.0213
	(-0.52)	(-0.68)	(-1.20)	(-0.90)	(0.16)	(0.07)
Degree of urbanisation	0.0165	-0.00298	0.00214	0.0221	-0.00532	-0.0366
	(0.21)	(-0.25)	(0.11)	(0.73)	(-0.12)	(-0.10)
Log net household income	-0.221	-0.0137	0.0296	0.0216	-0.0589	-0.220
	(-1.02)	(-0.49)	(0.76)	(0.34)	(-0.58)	(-0.22)
Log financial assets	0.00443	0.000489	-0.00661	-0.0463*	-0.0148	0.00366
	(0.05)	(0.05)	(-0.43)	(-1.68)	(-0.33)	(0.01)
Log total non-fin. assets	-0.000841	0.0137***	-0.00479	0.0138	-0.0167	-0.0968
	(-0.02)	(2.78)	(-0.56)	(0.87)	(-0.62)	(-0.42)
Log total debt	0.000556	-0.00336	0.000710	-0.00751	0.00125	-0.00640
	(0.03)	(-1.15)	(0.13)	(-0.81)	(0.09)	(-0.05)
Constant	2.778	0.0700	-0.0546	0.403	1.277	4.259
	(1.48)	(0.31)	(-0.15)	(0.67)	(1.29)	(0.48)
Observations	439	269	269	269	269	269
Adjusted R^2	0.112					

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. The OLS regression has been estimated on both waves of data and standard errors are clustered at the household level; the quantile estimates are based on a pure cross-section. All regressions use sampling weights.

A.3 Alternative specifications for contributors to return loss, including non-participants

Table A.7: Contributors to return loss, adding advanced financial literacy and risk tolerance, including non-participants

	p_{75}	p_{80}	p_{85}	p_{90}	p_{95}
Basic fin. literacy index	0.00717	0.0140	0.0149	0.0231	0.0656
	(0.39)	(0.77)	(0.58)	(0.42)	(0.40)
Advanced fin. literacy index	0.0524***	0.0594***	0.0815***	0.101**	0.162
	(3.56)	(4.33)	(3.53)	(2.39)	(0.93)
High self-rated fin. knowledge	0.0282	0.0219	-0.0150	-0.00427	0.101
	(1.01)	(0.82)	(-0.35)	(-0.06)	(0.30)
Financial advice: Own judgement	0.00761	0.0558**	0.0806**	0.230***	1.008***
	(0.29)	(2.25)	(2.04)	(3.25)	(2.68)
Own fin. jugdement * bas. literacy	0.0210	0.0247	-0.00406	-0.352***	-1.760***
	(0.83)	(1.02)	(-0.11)	(-4.71)	(-6.42)
Higher vocal education	0.0185	0.0273	-0.0312	0.0323	-0.0940
	(0.63)	(0.98)	(-0.71)	(0.42)	(-0.29)
Academic education	0.0355	0.0753**	-0.0208	-0.0439	-0.234
	(0.95)	(2.15)	(-0.36)	(-0.44)	(-0.52)
Age 41-64	0.0383	0.0617**	0.0493	0.0273	-0.131
	(1.17)	(2.01)	(1.00)	(0.31)	(-0.31)
Age 65+	0.137***	0.178***	0.228***	0.333***	0.0830
	(3.60)	(4.90)	(3.83)	(3.19)	(0.17)
Female	0.0311	0.0422	0.0635	0.150	0.126
	(0.88)	(1.20)	(1.17)	(1.46)	(0.27)
High tolerance for risky investm.	0.0758***	0.0984***	0.123***	0.181***	0.238
	(5.80)	(7.78)	(6.03)	(4.72)	(1.46)
Constant	0.0848**	0.108***	0.209***	0.322***	0.779*
	(2.48)	(3.23)	(3.86)	(3.34)	(1.92)
Observations	737	737	737	737	737

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. All estimates are based on a cross-section of households, including those who do not own any risky assets. For this reason, the coefficients are only estimated at higher quantiles where the return losses are strictly positive for all groups. All regressions use sampling weights.

Table A.8: Contributors to return loss, adding diversification questions, including non-participants

	p_{75}	p_{80}	p_{85}	p_{90}	p_{95}
Basic fin. literacy index	0.00591	0.00782	0.0100	0.0406	0.101
	(0.44)	(0.43)	(0.37)	(0.66)	(0.47)
Fin. lit. - diversification, direct question	0.0644**	0.0702**	0.137***	0.0925	0.0198
	(2.58)	(2.12)	(2.87)	(0.68)	(0.04)
Financial advice: Own judgement	0.0670***	0.126***	0.151***	0.171*	1.042***
	(3.64)	(5.31)	(4.20)	(1.81)	(2.84)
Own fin. jugdement * bas. literacy	0.0332*	0.0507**	0.0705*	0.0574	-1.206***
	(1.78)	(2.00)	(1.92)	(0.70)	(-3.48)
Higher vocal education	0.122***	0.139***	0.0872**	0.194*	0.336
	(5.77)	(5.06)	(2.06)	(1.81)	(0.85)
Academic education	0.0852***	0.171***	0.143***	0.158	0.434
	(3.11)	(4.87)	(2.69)	(1.13)	(0.81)
Age 41-64	0.0321	0.0321	0.0475	0.0867	0.0609
	(1.40)	(1.09)	(1.09)	(0.79)	(0.14)
Age 65+	0.148***	0.179***	0.309***	0.421***	0.405
	(5.61)	(5.19)	(6.08)	(3.23)	(0.83)
Female	-0.0421*	-0.0353	0.00270	-0.0620	-0.0356
	(-1.80)	(-1.17)	(0.06)	(-0.53)	(-0.07)
Constant	-0.00382	-0.0000749	0.00272	0.0904	0.338
	(-0.13)	(-0.00)	(0.05)	(0.57)	(0.56)
Observations	875	875	875	875	875

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. All estimates are based on a cross-section of households, including those who do not own any risky assets. For this reason, the coefficients are only estimated at higher quantiles where the return losses are strictly positive for all groups. All regressions use sampling weights.

Table A.9: Contributors to return loss, separate measures of financial advice, including non-participants

	p75	p80	p85	p90	p95
Basic fin. literacy index	0.0431***	0.0526**	0.0854**	0.0981	-1.104***
	(2.75)	(2.55)	(2.47)	(1.62)	(-3.62)
Financial advice: Professionals	-0.0935***	-0.0867***	-0.170***	-0.152	-1.031***
	(-3.83)	(-2.67)	(-3.10)	(-1.40)	(-2.63)
Prof. advice * bas. literacy	-0.0347	-0.0401	-0.0634	-0.0381	1.189***
	(-1.22)	(-1.42)	(-1.37)	(-0.40)	(3.33)
Financial advice: Family/friends	-0.106***	-0.115***	-0.196***	-0.192	-1.066**
	(-4.10)	(-3.30)	(-3.18)	(-1.54)	(-2.10)
Advice fam./friends * bas. literacy	-0.0322	-0.0336	-0.0514	-0.0377	1.216**
	(-1.25)	(-1.03)	(-0.90)	(-0.32)	(2.11)
Higher vocal education	0.106***	0.120***	0.0983*	0.177*	0.336
	(4.57)	(3.91)	(1.85)	(1.65)	(0.82)
Academic education	0.0871***	0.167***	0.163**	0.158	0.419
	(2.96)	(4.24)	(2.47)	(1.15)	(0.77)
Age 41-64	0.0253	0.0322	0.0639	0.0951	0.0609
	(0.99)	(0.94)	(1.09)	(0.86)	(0.14)
Age 65+	0.182***	0.225***	0.323***	0.438***	0.405
	(6.17)	(5.60)	(4.72)	(3.33)	(0.82)
Female	-0.0197	-0.0425	-0.0337	-0.0767	-0.0356
	(-0.78)	(-1.21)	(-0.55)	(-0.65)	(-0.07)
Constant	0.119***	0.158***	0.266***	0.354***	1.400***
	(4.25)	(4.28)	(4.15)	(3.15)	(2.98)
Observations	875	875	875	875	875

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. All estimates are based on a cross-section of households, including those who do not own any risky assets. For this reason, the coefficients are only estimated at higher quantiles where the return losses are strictly positive for all groups. All regressions use sampling weights.

Table A.10: Contributors to return loss, adding household size, urbanisation, and financial variables to the preferred specification. Non-participants are included in the estimation sample.

	p75	p80	p85	p90	p95
Basic fin. literacy index	0.00178	-0.000384	-0.00156	0.0209	0.0916
	(0.16)	(-0.02)	(-0.06)	(0.39)	(0.43)
Financial advice: Own judgement	0.0412**	0.104***	0.118***	0.200**	0.657
	(2.49)	(4.68)	(3.08)	(2.26)	(1.65)
Own fin. jugdement * bas. literacy	0.0213	0.0343	0.0487	0.0543	-0.663*
	(1.32)	(1.52)	(1.36)	(0.69)	(-1.73)
Higher vocal education	0.0493***	0.0498**	0.0581	0.0371	0.215
	(2.70)	(2.00)	(1.39)	(0.38)	(0.49)
Academic education	-0.0122	0.00784	0.0120	-0.0800	0.00247
	(-0.51)	(0.24)	(0.22)	(-0.60)	(0.00)
Age 41-64	-0.0102	-0.0108	-0.0246	-0.00719	-0.0751
	(-0.49)	(-0.39)	(-0.51)	(-0.07)	(-0.15)
Age 65+	0.0923***	0.103***	0.142**	0.303**	0.330
	(3.80)	(3.12)	(2.45)	(2.30)	(0.59)
Female	-0.000979	-0.00744	0.00394	-0.0361	0.0115
	(-0.04)	(-0.24)	(0.07)	(-0.30)	(0.02)
Household size	-0.0000545	0.000641	0.00991	0.0162	0.00430
	(-0.01)	(0.06)	(0.55)	(0.39)	(0.02)
Degree of urbanisation	0.0132	0.0218*	0.0189	0.0208	-0.0743
	(1.52)	(1.82)	(0.95)	(0.46)	(-0.37)
Log net household income	0.00882	0.00131	0.00883	-0.00996	-0.157
	(0.47)	(0.05)	(0.21)	(-0.11)	(-0.35)
Log financial assets	0.0597***	0.0800***	0.0907***	0.121***	0.221
	(8.29)	(8.25)	(5.34)	(2.93)	(0.97)
Log total non-fin. assets	-0.00237	-0.00299	-0.00377	-0.0108	-0.0297
	(-0.55)	(-0.50)	(-0.36)	(-0.45)	(-0.26)
Log total debt	0.00307	0.00246	0.00309	0.000588	-0.000964
	(1.12)	(0.68)	(0.50)	(0.04)	(-0.02)
Constant	-0.557***	-0.633***	-0.790*	-0.696	0.393
	(-3.01)	(-2.62)	(-1.95)	(-0.85)	(0.10)
Observations	860	860	860	860	860

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. All estimates are based on a cross-section of households, including those who do not own any risky assets. For this reason, the coefficients are only estimated at higher quantiles where the return losses are strictly positive for all groups. All regressions use sampling weights.

A.4 Contributors to return loss, excluding mutual fund fees

Table A.11: Contributors to return loss, excluding mutual fund fees.

	OLS	p_{10}	p_{30}	p_{50}	p_{70}	p_{90}
Basic fin. literacy index	0.0354	-0.0165*	0.0157	0.0554*	0.0668	0.228
	(0.67)	(-1.71)	(0.34)	(1.69)	(1.24)	(0.90)
Financial advice: Own judgement	0.534**	0.00628	0.0474	0.0503	0.411***	1.619***
	(2.35)	(0.38)	(0.92)	(1.14)	(4.79)	(2.73)
Own fin. jugdement * bas. literacy	-0.773*	0.0158	-0.0177	-0.0164	-0.524***	-1.888***
	(-1.91)	(1.05)	(-0.29)	(-0.36)	(-5.51)	(-2.98)
Higher vocal education	0.0751	0.00301	-0.0139	-0.00244	-0.1000	-0.347
	(0.48)	(0.16)	(-0.27)	(-0.05)	(-1.15)	(-0.58)
Academic education	-0.0257	0.00141	-0.00564	-0.0943*	-0.0331	-0.256
	(-0.17)	(0.06)	(-0.09)	(-1.72)	(-0.32)	(-0.36)
Age 41-64	-0.164	0.00487	0.0395	0.0706	-0.0580	-0.548
	(-0.77)	(0.20)	(0.61)	(1.22)	(-0.54)	(-0.72)
Age 65+	-0.0666	0.0361	0.139**	0.276***	0.166	-0.433
	(-0.28)	(1.39)	(1.97)	(4.32)	(1.43)	(-0.53)
Female	0.242	0.0312	0.0214	0.165***	0.190	0.747
	(1.21)	(1.35)	(0.32)	(2.73)	(1.59)	(0.83)
Constant	0.478**	0.0157	0.0773	0.150**	0.403***	1.320*
	(2.21)	(0.73)	(1.09)	(2.35)	(3.50)	(1.96)
Observations	441	270	270	270	270	270
Adjusted R^2	0.115					

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. The OLS regression has been estimated on both waves of data and standard errors are clustered at the household level; the quantile estimates are based on a pure cross-section. All regressions use sampling weights.

Table A.12: Contributors to return loss, excluding mutual fund fees.

	p_{75}	p_{80}	p_{85}	p_{90}	p_{95}
Basic fin. literacy index	0.00943	0.0144	0.0236	0.0554	0.111
	(0.71)	(0.88)	(0.78)	(0.90)	(0.56)
Financial advice: Own judgement	0.109^{***}	0.123^{***}	0.190^{***}	0.208^{**}	1.105^{***}
	(6.19)	(4.52)	(3.58)	(2.06)	(2.97)
Own fin. jugdement * bas. literacy	0.0354^{*}	0.0400	0.0636	0.0479	-1.276^{***}
	(1.90)	(1.55)	(1.32)	(0.55)	(-3.60)
Higher vocal education	0.124^{***}	0.162^{***}	0.133^{**}	0.197^{*}	0.355
	(6.25)	(5.38)	(2.27)	(1.72)	(0.86)
Academic education	0.104^{***}	0.210^{***}	0.178^{**}	0.168	0.465
	(4.13)	(5.44)	(2.33)	(1.15)	(0.84)
Age 41-64	0.0258	0.0309	0.0799	0.106	0.0618
	(1.17)	(0.92)	(1.23)	(0.91)	(0.14)
Age 65+	0.192^{***}	0.221^{***}	0.354^{***}	0.450^{***}	0.424
	(7.53)	(5.66)	(4.68)	(3.29)	(0.85)
Female	-0.0192	-0.0403	-0.0454	-0.0845	-0.0325
	(-0.88)	(-1.18)	(-0.68)	(-0.68)	(-0.07)
Constant	0.0153	0.0429	0.0705	0.165	0.374
	(0.70)	(1.29)	(1.07)	(1.41)	(0.92)
Observations	875	875	875	875	875

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. The OLS regression has been estimated on both waves of data and standard errors are clustered at the household level; the quantile estimates are based on a pure cross-section. All regressions use sampling weights.

A.5 Contributors to return loss, log-decomposition of Calvet *et al.* (2007)

Table A.13: Contributors to return loss and its components in logarithms, covariates similar to Calvet *et al.* (2007)

	RL_h	ω_h	β_h	$\frac{RSL_\mathrm{h}}{1-RSL_\mathrm{h}}$
Higher vocal education	0.0367	0.0564	0.0170	-0.0439
	(0.22)	(0.39)	(0.14)	(-0.34)
Academic education	-0.0468	0.133	-0.0922	-0.0738
	(-0.21)	(0.71)	(-0.52)	(-0.44)
Age 41-64	0.188	0.212	-0.151	0.0935
	(0.80)	(1.14)	(-1.24)	(0.61)
Age 65+	0.599^{**}	0.614^{***}	-0.274^{*}	0.172
	(2.29)	(2.81)	(-1.72)	(0.86)
Female	0.0601	0.212	-0.0691	-0.0617
	(0.26)	(1.17)	(-0.53)	(-0.36)
Household size	-0.0702	-0.00642	-0.00857	-0.0452
	(-1.12)	(-0.11)	(-0.17)	(-0.92)
Degree of urbanisation	0.0432	0.00668	-0.0804	0.0870
	(0.51)	(0.10)	(-1.13)	(1.35)
Log net household income	-0.187	-0.0860	-0.0613	-0.0477
	(-1.08)	(-0.56)	(-0.57)	(-0.44)
Log financial assets	-0.0700	0.0641	0.0228	-0.134^{**}
	(-0.81)	(0.99)	(0.44)	(-2.19)
Log total non-fin. assets	0.0107	-0.00588	-0.00232	0.0101
	(0.22)	(-0.16)	(-0.07)	(0.28)
Log total debt	-0.000265	0.00380	0.0120	-0.0145
	(-0.01)	(0.19)	(0.66)	(-0.84)
Constant	-3.558^{**}	-1.596	0.161	0.612
	(-2.18)	(-1.04)	(0.20)	(0.58)
Observations	589	589	589	589
Adjusted R^2	0.027	0.025	0.002	0.026

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details and for the definitions of the dependent variables. The regressions have been estimated on both waves of data and standard errors are clustered at the household level. All regressions use sampling weights.

Table A.14: Contributors to return loss and its components in logarithms

	RL_{h}	ω_{h}	β_{h}	$\frac{RSL_{\mathrm{h}}}{1-RSL_{\mathrm{h}}}$
Basic fin. literacy index	0.223**	0.329	-0.0634	-0.0539
	(2.30)	(1.62)	(-0.71)	(-0.35)
Financial advice: Own judgement	0.339*	0.0992	0.127	0.130
	(1.88)	(0.59)	(0.78)	(0.85)
Own fin. jugdement * bas. literacy	-0.456**	-0.530**	0.239	-0.106
	(-1.97)	(-2.17)	(1.16)	(-0.50)
Higher vocal education	-0.0362	-0.0285	0.0789	-0.0804
	(-0.18)	(-0.18)	(0.59)	(-0.56)
Academic education	-0.190	0.172	-0.209	-0.127
	(-0.82)	(0.89)	(-0.86)	(-0.68)
Age 41-64	0.0585	0.229	-0.184	-0.0159
	(0.23)	(1.14)	(-1.16)	(-0.09)
Age 65+	0.600**	0.686***	-0.263	0.101
	(2.28)	(3.18)	(-1.46)	(0.49)
Female	0.473*	0.319	-0.00710	0.171
	(1.81)	(1.57)	(-0.04)	(0.80)
Constant	-6.435***	-1.953***	-0.282*	-1.396***
	(-24.20)	(-8.65)	(-1.73)	(-7.01)
Observations	441	441	441	441
Adjusted R^2	0.047	0.057	0.013	-0.000

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details and for the definitions of the dependent variables. The regressions have been estimated on both waves of data and standard errors are clustered at the household level. All regressions use sampling weights.

A.6 Contributors to alternative diversification measures

Table A.15: Contributors to idiosyncratic portfolio risk.

	OLS	p_{10}	p_{30}	p_{50}	p_{70}	p_{90}
Basic fin. literacy index	-1.323	-0.227	-0.927	-0.491	0.231	-7.873***
	(-0.72)	(-0.45)	(-1.20)	(-1.01)	(0.11)	(-3.30)
Financial advice: Own judgement	3.917**	0.806*	0.748	2.225***	7.986***	11.76***
	(2.37)	(1.82)	(0.84)	(3.28)	(3.07)	(2.77)
Own fin. jugdement * bas. literacy	-0.377	-0.304	0.533	-0.929	-4.984	-0.735
	(-0.14)	(-0.53)	(0.56)	(-1.23)	(-1.65)	(-0.21)
Higher vocal education	0.0219	0.390	1.412	0.306	0.801	0.165
	(0.01)	(0.79)	(1.46)	(0.44)	(0.31)	(0.03)
Academic education	-3.346**	-0.0129	0.760	-1.129	-2.933	-7.223
	(-2.22)	(-0.03)	(0.67)	(-1.35)	(-0.96)	(-1.33)
Age 41-64	-2.619	-0.523	0.729	0.882	-3.133	-3.264
	(-1.26)	(-0.86)	(0.60)	(1.01)	(-0.95)	(-0.61)
Age 65+	-3.199	-0.301	0.788	0.645	-3.586	-9.155
	(-1.46)	(-0.45)	(0.60)	(0.67)	(-0.99)	(-1.52)
Female	0.724	-0.179	-1.389	0.354	1.541	-3.683
	(0.45)	(-0.28)	(-1.05)	(0.37)	(0.43)	(-0.66)
Constant	13.63***	4.132***	5.170***	7.294***	12.92***	28.96***
	(5.79)	(6.51)	(4.01)	(7.69)	(3.52)	(4.26)
Observations	441	270	270	270	270	270
Adjusted R^2	0.042					

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. The OLS regression has been estimated on both waves of data and standard errors are clustered at the household level; the quantile estimates are based on a pure cross-section. All regressions use sampling weights.

Table A.16: Contributors to return loss as a fraction of risky financial assets.

	OLS	p_{10}	p_{30}	p_{50}	p_{70}	p_{90}
Basic fin. literacy index	0.0857	-0.121***	-0.0898***	-0.00550	0.0518	0.423
	(0.62)	(-2.86)	(-2.80)	(-0.08)	(0.25)	(0.80)
Financial advice: Own judgement	1.260***	0.0151	0.0872	0.318***	1.484***	3.724***
	(3.68)	(0.19)	(1.65)	(2.63)	(3.52)	(3.25)
Own fin. jugdement * bas. literacy	-0.799	0.0750	0.143***	0.0326	-1.596***	-2.059**
	(-1.54)	(1.11)	(3.18)	(0.26)	(-4.10)	(-2.18)
Higher vocal education	0.114	0.0227	0.00505	-0.0239	-0.0615	0.343
	(0.29)	(0.25)	(0.09)	(-0.19)	(-0.14)	(0.24)
Academic education	-0.538	-0.00718	-0.129**	-0.246*	-0.382	0.420
	(-1.52)	(-0.06)	(-1.98)	(-1.68)	(-0.74)	(0.27)
Age 41-64	-0.815	-0.0620	0.0826	-0.0135	-1.558**	-1.055
	(-1.52)	(-0.44)	(0.99)	(-0.08)	(-2.55)	(-0.60)
Age 65+	-1.085**	-0.0605	0.148*	0.132	-1.589**	-2.315
	(-2.00)	(-0.39)	(1.67)	(0.69)	(-2.38)	(-1.28)
Female	0.317	-0.00392	-0.0314	0.283	0.107	-0.204
	(0.86)	(-0.03)	(-0.43)	(1.60)	(0.19)	(-0.15)
Constant	2.057***	0.473***	0.516***	0.711***	2.731***	3.829*
	(3.70)	(3.21)	(6.19)	(3.87)	(4.09)	(1.67)
Observations	316	194	194	194	194	194
Adjusted R^2	0.062					

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. The OLS regression has been estimated on both waves of data and standard errors are clustered at the household level; the quantile estimates are based on a pure cross-section. Households with an annual standard deviation of the financial portfolio below 2% have been excluded. All regressions use sampling weights.

Table A.17: Contributors to relative Sharpe ratio loss

	OLS	p_{10}	p_{30}	p_{50}	p_{70}	p_{90}
Basic fin. literacy index	2.128	-1.093	0.0596	1.435	1.950	6.455
	(1.09)	(-1.37)	(0.07)	(1.53)	(0.47)	(1.37)
Financial advice: Own judgement	4.151	0.211	1.833	2.848^{*}	11.07^{**}	4.141
	(1.29)	(0.15)	(1.29)	(1.83)	(2.10)	(0.46)
Own fin. jugdement * bas. literacy	-6.525^{*}	0.778	1.009	-1.757	-9.394^{*}	-14.88^{**}
	(-1.73)	(0.66)	(0.81)	(-1.08)	(-1.72)	(-2.44)
Higher vocal education	-2.205	-0.322	-0.484	-0.397	-6.249	-4.473
	(-0.72)	(-0.22)	(-0.33)	(-0.25)	(-1.08)	(-0.49)
Academic education	-3.797	-0.0403	-1.728	-0.829	-4.399	-10.59
	(-0.99)	(-0.02)	(-0.94)	(-0.44)	(-0.69)	(-0.95)
Age 41-64	-1.309	0.428	-1.548	-4.302^{*}	-5.154	16.07
	(-0.36)	(0.20)	(-0.74)	(-1.96)	(-0.66)	(1.30)
Age 65+	-0.490	0.831	-0.737	-1.212	-5.755	4.135
	(-0.12)	(0.36)	(-0.33)	(-0.51)	(-0.70)	(0.31)
Female	2.215	0.161	3.693^{*}	7.738^{***}	1.887	-11.31
	(0.71)	(0.07)	(1.81)	(3.70)	(0.26)	(-1.25)
Constant	23.32^{***}	7.460^{***}	11.34^{***}	16.25^{***}	31.20^{***}	41.92^{**}
	(6.25)	(3.79)	(5.29)	(7.02)	(3.51)	(2.58)
Observations	316	194	194	194	194	194
Adjusted R^2	0.005					

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. The OLS regression has been estimated on both waves of data and standard errors are clustered at the household level; the quantile estimates are based on a pure cross-section. Households with an annual standard deviation of the financial portfolio below 2% have been excluded. All regressions use sampling weights.

Table A.18: Contributors to fraction in shares.

	OLS	p_{10}	p_{30}	p_{50}	p_{70}	p_{90}
Basic fin. literacy index	-0.00361	0	0	2.73e-11	0.0682	2.21e-18
	(-0.06)	(0.00)	(0.00)	(0.00)	(0.83)	(0.00)
Financial advice: Own judgement	0.158***	0	0	0.156***	0.619***	1.62e-18
	(2.75)	(0.00)	(0.00)	(11.40)	(6.03)	(0.00)
Own fin. jugdement * bas. literacy	0.00355	0	0	-0.00303	-0.127	-2.51e-18
	(0.04)	(0.00)	(0.00)	(-0.18)	(-1.17)	(-0.00)
Higher vocal education	0.0256	0	0	1.10e-10	0.00939	-5.40e-18
	(0.41)	(0.00)	(0.00)	(0.00)	(0.09)	(-0.00)
Academic education	-0.0941	0	0	-3.66e-10	-0.177	-0.0459***
	(-1.40)	(0.00)	(0.00)	(-0.00)	(-1.46)	(-2270126.15)
Age 41-64	-0.122	0	0	2.81e-10	0.0230	1.19e-18
	(-1.47)	(0.00)	(0.00)	(0.00)	(0.17)	(0.00)
Age 65+	-0.0727	0	0	6.81e-11	0.130	1.29e-18
	(-0.82)	(0.00)	(0.00)	(0.00)	(0.87)	(0.00)
Female	0.0495	0	0	2.09e-10	0.0832	-1.77e-19
	(0.63)	(0.00)	(0.00)	(0.00)	(0.56)	(-0.00)
Constant	0.329***	0	0	-1.66e-10	0.196	1***
	(3.62)	(0.00)	(0.00)	(-0.00)	(1.30)	(88079974.66)
Observations	441	270	270	270	270	270
Adjusted R^2	0.034					

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. The OLS regression has been estimated on both waves of data and standard errors are clustered at the household level; the quantile estimates are based on a pure cross-section. All regressions use sampling weights.

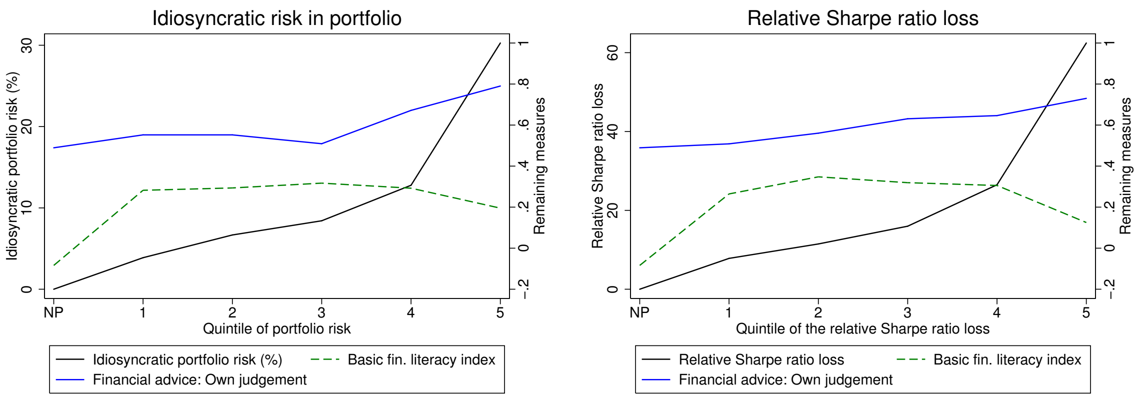
Table A.19: Contributors to the inverse of the diversification index D_1 from Guiso and Jappelli (2009)

	OLS	p10	p30	p50	p70	p90
Basic fin. literacy index	-0.0367	0	0	3.20e-11	0.0211	-0.250**
	(-0.54)	(0.00)	(0.00)	(0.00)	(0.46)	(-2.39)
Financial advice: Own judgement	0.0875*	0	0	0.0779***	0.227***	0.281*
	(1.74)	(0.00)	(0.00)	(11.73)	(4.21)	(1.91)
Own fin. jugdement * bas. literacy	0.00551	0	0	-0.0242***	-0.163**	0.193
	(0.07)	(0.00)	(0.00)	(-2.91)	(-2.51)	(1.55)
Higher vocal education	0.0277	0	0	-4.01e-10	0.0274	0.218
	(0.55)	(0.00)	(0.00)	(-0.00)	(0.50)	(1.18)
Academic education	-0.0499	0	0	-1.46e-10	-0.0378	0.0115
	(-0.93)	(0.00)	(0.00)	(-0.00)	(-0.60)	(0.06)
Age 41-64	-0.126*	0	0	3.14e-10	-0.0757	-0.330
	(-1.82)	(0.00)	(0.00)	(0.00)	(-1.11)	(-1.61)
Age 65+	-0.0769	0	0	2.40e-10	-0.0868	-0.270
	(-1.06)	(0.00)	(0.00)	(0.00)	(-1.17)	(-1.36)
Female	0.0711	0	0	-1.72e-11	0.0339	-3.93e-09
	(0.99)	(0.00)	(0.00)	(-0.00)	(0.45)	(-0.00)
Constant	0.231***	0	0	-5.85e-11	0.111	0.756***
	(3.08)	(0.00)	(0.00)	(-0.00)	(1.46)	(3.30)
Observations	433	266	266	266	266	266
Adjusted R^2	0.033					

Source: CentERpanel, Datastream, Euroinvestor, own calculations. See Section 2.3 for computational details. The OLS regression has been estimated on both waves of data and standard errors are clustered at the household level; the quantile estimates are based on a pure cross-section. All regressions use sampling weights.

B Additional graphs

Figure B.1: Financial literacy, financial advice, and diversification losses – alternative measures of diversification loss



Source: CentERpanel, Datastream, Euroinvestor, own calculations. The return loss quintile value "NP" stands for non-participants.