Netspar
Network for Studies on Pensions, Aging and Retirement

Netspar DISCUSSION PAPERS

Hans-Martin von Gaudecker

How does Household Portfolio Diversification Vary with Financial Literacy and Financial Advice?

DP 03/2011-026 (revised version December 7, 2013)

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

Journal of Finance, forthcoming

Hans-Martin von Gaudecker *

December 7, 2013

ABSTRACT

Household investment mistakes are an important concern for researchers and policymakers alike; portfolio underdiversification ranks among those mistakes that are potentially most costly. However, its roots and empirical importance are poorly understood. I estimate quantitatively meaningful diversification statistics and investigate their relationship with key variables. Nearly all households that either score high on financial literacy; or rely on professionals or private contacts for advice; achieve reasonable investment outcomes. Compared to these groups, households with below-median financial literacy that trust their own decision-making capabilities, lose an expected 50bp on average. All group differences stem from the top of the loss distribution.

JEL codes: D14, D12, G11

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

*Universität Bonn, Department of Economics, Adenauerallee 24-42, 53012 Bonn; Germany; hmgaudecker[at]uni-bonn[dot]de; tel. +49 228 73 9357. 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. Special thanks to him for an excellent programming job and many fruitful discussions. Funding from Netspar is gratefully acknowledged. I appreciate helpful comments received from seminar and conference participants at the Universities of Frankfurt and of Bonn, at the 2nd SAVE conference in Deidesheim, at the 2011 Netspar Pension Workshop in Amsterdam, and at the Federal Reserve Bank of Chicago; in particular Tabea Bucher-Koenen, Armin Falk, Dimitris Georgarakos, Michael Haliassos, Rainer Haselmann, Olivia Mitchell, Karl Scholz, Stephen Zeldes, and Michael Ziegelmeyer. Furthermore, I would like to thank CentERdata for help on numerous data questions; 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 Rooij, Lusardi, and Alessie (2011).

Economic theory predicts that households will hold their risky assets in the form of a well-diversified portfolio. The extent to which this prediction holds true has important implications for the regulation of consumer financial products, the design of retirement savings plans and the distribution of consumer well-being in general. For these issues, it is especially significant to know the patterns of underdiversification. If only households that are particularly savvy in financial matters invested their retirement wealth in a few highly correlated stocks, most policymakers would conclude that this was a rational choice. For example, it could reflect superior information (Nieuwerburgh and Veldkamp, 2010). If it were the least financially literate doing the same thing, however, it would be much more likely that these households were simply making poor decisions. In a similar vein, it makes a big difference for a regulator's strategy whether the underdiversified households rely on professional financial advisers; or whether they make their investment choices autonomously.

Only little guidance has come from empirical studies, which can be roughly divided into two strands. One strand has studied detailed diversification properties using administrative data from Scandinavian countries (Calvet, Campbell, and Sodini, 2007, 2009; Grinblatt, Keloharju, and Linnainmaa, 2011). These studies show that—once one accounts for mutual funds on top of stock portfolios—most households reach reasonable investment outcomes, but some may expect a very low rate of return given the amount of risk they take. However, the lack of data on financial literacy and financial advice prevents one from going beyond unconditional distributions of diversification losses with respect to the questions raised above. A second strand has employed simple measures of portfolio diversification drawn from household surveys (Bilias, Georgarakos, and Haliassos, 2009; Kimball and Shumway, 2010; Guiso and Jappelli, 2009; Graham, Harvey, and Huang, 2009). In these studies, the data issues are reversed: While covariates are abundant, the diversification properties do not permit a quantitative analysis.

In this paper, I combine the strengths of both approaches and calculate quantitatively meaningful diversification measures for the respondents to a Dutch household survey. Next to an abundance of background variables, the survey also includes detailed information on financial literacy and on the most important source of financial advice. In my analysis, I employ the return loss (Calvet, Campbell, and Sodini, 2007) as the diversification measure. The return loss is the difference between the maximum expected return that is attainable at a given standard deviation and the actual expected return for a particular portfolio.

The unconditional distribution of the annual return loss reveals that it is limited to half a percentage point or less up to the fourth quintile. In its top quintile, it averages 1.8 percentage points, which is large and comparable to previous results.

My main analyses uncover an important interaction of financial literacy and financial advice in the decision-making process. The return loss of households that seek advice either in their private network or from professionals does not vary with financial literacy. Among households making autonomous decisions, the distribution of the return loss looks very similar only if they are endowed with the maximum level of financial literacy. In this group, a decrease in financial literacy of one standard deviation is associated with an increase in the return loss by 0.7 percentage points on average. I show by means of quantile regressions that these effects stem entirely from very large effects at the upper end of the return loss distribution.

Intrigued by the fact that the outcomes of households relying on professional financial advisers and on recommendations from within their private network look so similar, I further investigate the differences between these groups. The use of professional financial advice is associated with an increase in fees paid across the entire portfolio of 0.3 percentage points. However, the higher fees do not translate into a significantly higher return loss.

This paper is structured as follows. In the first section, I describe the dataset and the mean-standard deviation characteristics of households' portfolios, including the return loss. In Section II, I explore the sources of underdiversification and scrutinize the effects of professional advisers, before discussing these results. Section III concludes.

I. The properties of households' portfolios

Our understanding of the empirical characteristics of households portfolios is based on three distinct sources of data: (1) tax registers from Scandinavian countries; (2) online brokers' databases; and (3) household surveys. Each has its own advantages and drawbacks: The first two sources usually yield extraordinarily detailed portfolio data. However, they do so for highly selective sets of countries and individuals, respectively. The availability of certain covariates is another issue. In surveys the situation is the opposite: While representative samples with a plethora of background information are available all over the world, informa-

tion on household portfolios is rather coarse.[1]

In this study, I combine several strengths of the various approaches by constructing detailed portfolios for the respondents to the Dutch Central Bank Household Survey (DHS). Linking individual portfolio components to historical return series allows me to calculate diversification statistics that are measured in meaningful economic quantities, mirroring studies based on administrative registers or discount broker data. This section establishes the usefulness of this approach; Section II goes beyond previous studies by exploiting the background characteristics present in the DHS data.

A. The DHS data on household portfolios

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 the cost of using the Internet. The CentERpanel is representative of the Dutch population in terms of observable characteristics. It hosts the Dutch Central Bank Household Survey (DHS), which contains particularly detailed information on financial matters. For this reason, the data have been used extensively to describe households' portfolio choice behavior. Alessie, Hochgürtel, and Soest (2002, 2004, 2006) and Dimmock and Kouwenberg (2010); and Korniotis and Kumar (2011) are excellent examples. 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 (Lusardi and Mitchell, 2007; Hung, Parker, and Yoong, 2009; Hung and Yoong, 2010).

My analysis is cross-sectional, but in order to increase the sample size I make use of two waves (2005 and 2006) and adjust all standard errors for clustering within households using standard methods (StataCorp, 2013). Throughout the analysis, I exclude households with less than 1,000 euros in financial assets (8.6% of the sample), which yields 1,604 households to begin with. About a third of these hold risky financial assets, defined as shares, various types of funds, bonds, and options. See Section A.2 (in particular Table A.1) in the Online Appendix for more details.

[1] A detailed review of the literature and of data sources can be found in the Online Appendix, Section A.1.

A unique feature of the dataset is that individuals are not only asked for the number of stocks and mutual funds they possess, but also to report each item's name and the quantity held. I follow the same strategy as Calvet, Campbell, and Sodini (2007, 2009) and connect each item to its time series of returns obtained from Datastream or Morningstar. This allows me to estimate the risk-return characteristics of households' portfolios directly from the data. Importantly, I do not need to invoke any assumptions about the properties of asset classes. Furthermore, I can take the covariances between different items' returns into account in an appropriate way.

While most households' reports of their portfolio constituents appear to be very thorough, I lose about a fourth of the risky financial asset holders because of incomplete reports. Fortunately, I cannot detect any bias in terms of observable characteristics when dropping these households. See Tables A.2 and A.3 in the Online Appendix for all details. My final sample consists of 381 observations.

The households in my sample possess 269 different assets comprised of 99 shares and 170 funds. I observe these assets for more than 11 years on average. The lengths of these return series vary widely, ranging from below six years at the fifth percentile to almost 20 years at the ninety-fifth percentile. I conduct all analyses at a monthly frequency, but report annualized results for easier comparability with the literature.

Of the 170 funds, 106 are equity funds; there are about a dozen funds each investing in sovereign bonds, corporate bonds, and real estate; 28 funds follow a mixed strategy. I also obtained data on the fees of these funds. These are about 130 bp on average, with the fifth and ninety-fifth percentiles at 33 bp and 187 bp. Section A.3 in the Online Appendix describes the characteristics of the portfolios' components in more detail.

B. The mean-variance characteristics of households' portfolios

I follow Calvet, Campbell, and Sodini (2007) and assume that assets are priced according to an international CAPM. Directly estimating the expected return of each household's portfolio would be problematic because of the short return histories for some assets; and because the time series cover different time spans.[2] I take the MSCI Europe Index to proxy the efficient

[2]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 in the 2002-2007 period. The first asset would likely have a much lower estimated α because the market conditions were worse during the earlier period. Pricing assets

market portfolio, which is a natural choice for a member of the Eurozone; results are robust to using the AEX or the MSCI World Index instead. All returns are framed as excess returns over the risk-free rate, which is approximated by the one-month EURIBOR. Before applying the CAPM, I subtract mutual funds' fees from their gross returns and 30 basis points from the benchmark index's returns. This is approximately the cost charged by index funds replicating common benchmarks.

Net of this "fee", the MSCI Europe has an annual excess return of $\mu_b = 5.77\%$ in the period of January 1983 to July 2009. Along with its standard deviation of $\sigma_b = 16.7\%$, this implies a Sharpe ratio $S_b = {}^{\mu_b}/_{\sigma_b}$ of 35%. Imposing the CAPM leads to the following regression for all assets $a \in \{1, 2, \ldots, 269\}$:

$$r^e_{a,t} = \beta_a \cdot r^e_{b,t} + \varepsilon_{a,t}. \tag{1}$$

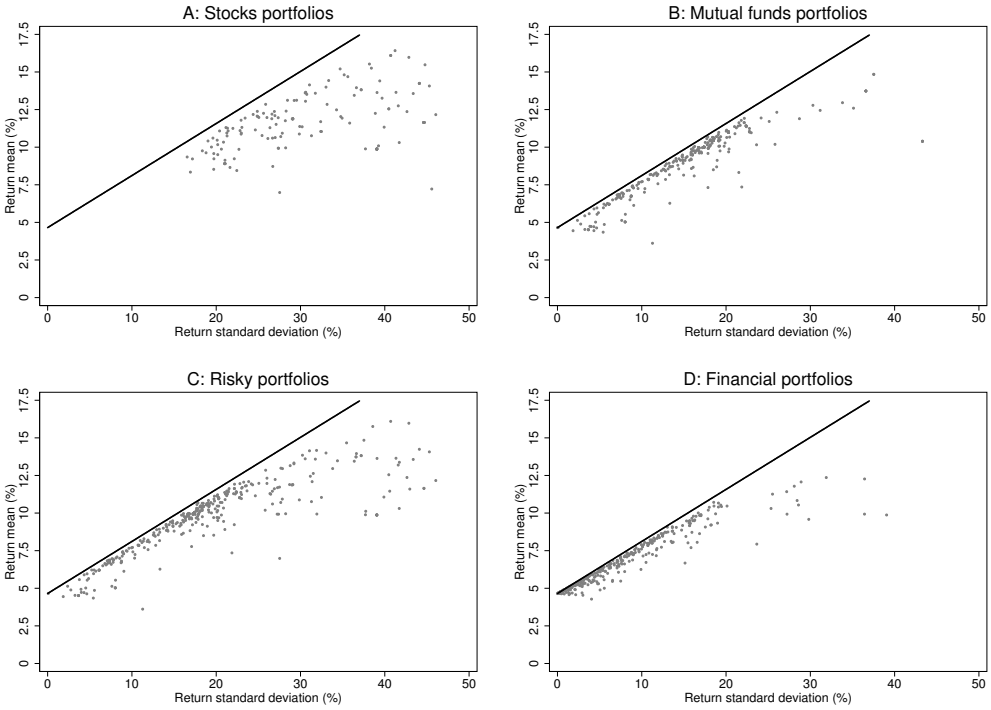
Given the βs of all assets and the portfolio weights for each household, it is straightforward to calculate the expected return μ_h of each household's portfolio.

In order to gain an understanding of the basic characteristics of household portfolios, it is useful to plot them in the mean-standard deviation plane. Panel A of Figure 1 does this for the pure stock portfolios and reveals a picture of strong underdiversification, which is similar to the findings of Calvet, Campbell, and Sodini (2007) for Sweden and Goetzmann and Kumar (2008) for the United States. The mutual fund component of households' portfolios appears to be much better diversified. The major part of the distribution lines up right below the efficient frontier (Panel B of Figure 1). Nevertheless, a substantial fraction of mutual fund portfolios perform significantly worse than the market portfolio at any risk level. 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). Panel C implies that studies forced to focus on stock portfolios (e.g., Grinblatt, Keloharju, and Linnainmaa, 2012; Goetzmann and Kumar, 2008) miss an important part of the picture.

The distance to the efficient frontier shrinks a lot more when holdings of safe assets are

via the CAPM avoids this problem as long as the assets' correlation structure does not change with market conditions.

Figure 1. The mean-standard deviation characteristics of household portfolios



Source: CentERpanel, Datastream, Euroinvestor, own calculations.

taken into account in Panel D. Nevertheless, a few outliers with severe losses remain at high levels of risk; a number of households hold portfolios one to two 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, Campbell, and Sodini, 2007; Tang, Mitchell, Mottola, and Utkus, 2010). Detailed descriptions of mutual fund fees and of other characteristics of the funds and shares in my sample are relegated to Section A.3 in the Online Appendix.

C. The expected costs of underdiversification

A variety of measures are available to reduce the two-dimensional graphs in Figure 1 to a single statistic of diversification losses.[3] The most comprehensive criterion is the return loss, defined by Calvet, Campbell, and Sodini (2007) as the expected return a household loses by not choosing a position on the efficient frontier with the same level of risk as its portfolio. In graphical terms, the return loss is the vertical distance between the efficient frontier and the location of a household's portfolio in Panel D of Figure 1.

The first column in Table I shows the average return loss within each quintile of its distribution. Return losses are limited to half a percentage point per year up to the fourth quintile; they reach 1.8 percentage points annually in the highest quintile. As mentioned before, such a number implies a large reduction in wealth when accumulated over the life cycle.

In order to get an idea of what contributes to the return loss, it is useful to write it as the product of the expected excess return on the market portfolio (which does not vary across households), the risky asset share ω_h, β_h, and the relative difference between the Sharpe ratios of the efficient market portfolio and the household's portfolio:

$$RL_h = \mu_b \cdot \omega_h \cdot \beta_h \cdot \left(\frac{S_b - S_h}{S_h} \right). \tag{2}$$

High values of ω_h and β_h may be signs of efficient risk-taking; the quantity $(S_b - S_h)/S_h$ unambiguously measures a loss from underdiversification.

The remaining columns of Table I contain the three terms that vary across households. Of course, the different components do not add up due to Jensen's inequality; hence this illustration does not qualify as a decomposition. Nevertheless, it remains useful to get a rough idea of the underlying magnitudes. The risky asset share increases quickly until the middle of the return loss distribution and increases modestly afterwards. The β coefficient rises almost linearly over the quintiles, so inefficient and efficient risk-taking at least go hand in hand on average. The diversification loss is fairly constant in the lower quintiles and increases strongly in the top quintile. It is the only one of the three components of RL_h that is significantly different between the fourth and fifth quintiles of RL_h at the 10%-level

[3]Section B in the Online Appendix describes a number of these measures and provides a detailed comparison to the analysis of Calvet, Campbell, and Sodini (2007).

Table I
Return loss and its components by quintile

Return loss quintile	**Return loss** RL_h	**Fraction in risky assets** ω_h	**Beta coefficient** β_h	**Diversification loss*** $\frac{S_b - S_h}{S_h}$
Lowest	0.04	0.08	0.77	0.30
	(0.00)	(0.01)	(0.04)	(0.08)
Second	0.14	0.23	0.83	0.46
	(0.00)	(0.02)	(0.06)	(0.10)
Third	0.28	0.43	0.89	0.27
	(0.01)	(0.03)	(0.04)	(0.04)
Fourth	0.52	0.54	0.91	0.50
	(0.01)	(0.03)	(0.04)	(0.10)
Highest	1.80	0.60	1.01	0.88
	(0.16)	(0.03)	(0.07)	(0.12)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The table entries are average values of the statistics denoted in the table header by return loss quintile. Standard errors of the means are in parentheses.
* The value $(S_b - S_h)/S_h$ becomes extremely high if the expected return on a household's portfolio μ_h is close to zero. I therefore winsorize $(S_b - S_h)/S_h$ above at its ninety-fifth percentile.

(p-value 0.01). Compared to the lower quintiles, the prime driving force behind the highest return losses thus seems to be uncompensated risk-taking. It is important to note that the average size of the financial portfolio is highest in the top two quintiles of the return loss distribution; i.e., losses are not concentrated among small-scale investors.

I have thus shown that the descriptive results of Calvet, Campbell, and Sodini (2007) can be replicated to a large extent for another country and, more importantly, with a type of dataset that is fairly easy to collect anywhere. The greatest advantage of the DHS data, however, lies elsewhere: For the first time, it allows the detailed diversification measures to be related to the covariates the literature deems most important. For example, the Swedish administrative data of Calvet, Campbell, and Sodini (2007, 2009) contain measures of wealth, income, employment, age, household size, education, and immigration status. One of their main findings is that more affluent households invest both more aggressively

and more efficiently. The data do not allow one to discern whether this is because wealthy households purchase better advice or whether they are better equipped to make their own financial decisions. The policy conclusions would be very different. In the former case, one would target the supply of investment advice. If investor sophistication is the key, financial education programs could be of help (Tang, Mitchell, Mottola, and Utkus, 2010). The results in the next section help to better understand such channels.

II. The sources of underdiversification

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., the return loss considered in the previous section. A certainly non-exhaustive list of important inputs identified in the literature includes financial literacy, cognitive abilities and education, sources of financial advice, wealth, age, risk aversion, and several others described below. I approximate the production function by a linear equation:

$$Y = X'b + u. \tag{3}$$

The investment outcome (Y) is observed and relevant for the household as a whole, but many of the inputs concern individuals. For the latter, I use the answers of the households' financial deciders. First of all, these are the individuals who have answered the questions on financial literacy. Second, the results of Smith, McArdle, and Willis (2010) provide evidence that this approach is sensible: They analyze the correlation between cognitive skills and various economic outcomes for older households separately for each partner and show that numeracy of the financial respondent in the HRS data is by far the most important correlate. My analysis identifies subgroups of the population who are at an increased risk of obtaining inferior investment outcomes. It does not follow without further assumption that changing a covariate would lead to a change in Y corresponding to my estimates of b.

As seen in the previous section, the diversification loss is close to negligible for a large part of its distribution; losses only become high in the upper tail. For this reason, I estimate Equation (3) not only 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,

since the return loss (2) is well defined for them: $\omega_h = 0$ or $S_b - S_h = 0$ imply $RL_h = 0$. [4] I estimate (3) using various sets of covariates; my preferred specification includes financial numeracy, the most important source of financial advice, education, age, and the amount of financial assets. A detailed motivation for this choice of covariates may be found in Section A.4 in the Online Appendix, in which I also precisely describe all variables. Here, I briefly discuss financial numeracy and the most important source of financial advice.

One reason for using the portfolio data from the 2005 and 2006 DHS waves is that detailed financial literacy measures are available for that time. Those measures form the basis of Rooij, Lusardi, and Alessie (2011), and the authors kindly provided me with data and code. My financial numeracy score is based on four simple quiz-like math problems framed in financial terms and a question on the time value of money. Using factor-analytic methods, I extract a continuous index and standardize it to have a mean of zero and a variance of one in the overall sample.[5] The financial numeracy index measures whether individuals possess the necessary skills to perform simple numerical computations in financial matters, which is important for informed financial decision-making. Abstracting from agency problems and potential costs, rational households that realize their lack of investment skills would seek external help. In the DHS data, about a quarter of respondents sought help from professional advisers, while another quarter relied upon the advice of family and friends. The remaining half relied on their own financial judgment.

A. Financial literacy, financial advice, and their interaction

Table II contains the average return loss by the primary source of financial advice and a median-split of financial numeracy. The first cell contains individuals who decide by themselves and have fairly low financial numeracy. This group loses about one percentage point per year in expected returns compared to investing in the efficient portfolio with the

[4] The typical way to include non-participants in a least-squares framework would be to model (3) as a two-part process of first deciding whether to invest in risky assets and then deciding how to invest in them (see Pohlmeier and Ulrich (1995) for such a model in another context and Calvet, Campbell, and Sodini (2007) for an application to portfolio choice). Given that the participation decision has been studied extensively—including with the very data used here (Rooij, Lusardi, and Alessie, 2011)—such an approach seems to be an unnecessary complication.

[5] Except for a minor change in the computation described in Section A.4 of the Online Appendix, this is equivalent to the "basic financial literacy" variable of Rooij, Lusardi, and Alessie (2011)

same level of risk. This value is twice as much as the average of the remaining categories. It is the only entry that is significantly higher than the average of the others.[6] The importance of this interaction effect holds up throughout the more formal analyses in this section.

Table II
Average return loss of risky asset holders by financial numeracy and the primary source of financial advice

Primary source of financial advice	**Fin. numeracy < median**		**Fin. numeracy ≥ median**	
	Mean	**(S.E.)**	**Mean**	**(S.E.)**
Self-decider	1.00	(0.24)	0.56	(0.08)
Professional adviser	0.51	(0.21)	0.33	(0.04)
Family / friends	0.46	(0.26)	0.53	(0.12)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The number of observations is 274. The return loss measures the expected return per year (in percentage points) that is foregone by not investing in the efficient portfolio with the same standard deviation of a household's portfolio, see Section I.C for its definition.

The first column of Table III shows the results for an OLS regression of the return loss on my preferred set of covariates in the sample of participants. Financial numeracy does not have an effect for those who seek professional advice: The coefficient in the first row is essentially zero (just one basis point per year) and is precisely estimated (the 95% confidence interval ranges from -8 bp to 9 bp). There is no significant difference from those who rely on their network of family and friends: Both the dummy and the interaction term with financial numeracy are small and insignificant, although the standard errors are a bit higher. The dummy for deciding on the basis of self-collected information takes on a large and significantly positive value – those who rely on their own judgment with a financial numeracy score of zero incur a return loss that is 53 basis points higher on average than the return loss of those who rely on professional advice. The negative interaction term shows that this effect is much worse if the financial numeracy index takes on values below zero. It almost exactly cancels out for those who achieve the highest financial numeracy score of

[6]The value for high-numeracy households that rely on professional advice is significantly lower than the others. However, this result does not survive a change in the definition of fees, see Section II.B.

0.64 (note that the distribution is left-skewed). These households are estimated to incur an insignificant extra return loss of 53bp − 68bp × 0.64 = 9bp on average, compared to those who seek advice from professionals or family/friends and have a financial numeracy score of zero. The coefficients on all other covariates are insignificant and much smaller than those for financial numeracy and advice.

Table III
Contributors to return loss

	OLS	$\mathbf{p_{10}}$	$\mathbf{p_{30}}$	$\mathbf{p_{50}}$	$\mathbf{p_{70}}$	$\mathbf{p_{90}}$
Prof. advice × fin. numeracy	0.01	-0.01	0.01	0.05	0.08	0.11
	(0.04)	(0.02)	(0.03)	(0.07)	(0.16)	(0.43)
Financial advice: Family/friends	0.10	0.01	-0.04	-0.01	-0.02	0.30
	(0.15)	(0.03)	(0.05)	(0.11)	(0.24)	(0.66)
Advice fam./friends × fin. numeracy	0.14	0.05	0.06	0.14	0.21	0.55
	(0.13)	(0.04)	(0.07)	(0.13)	(0.29)	(0.80)
Financial advice: Own judgment	0.53***	0.00	0.02	0.02	0.34*	1.45***
	(0.19)	(0.02)	(0.04)	(0.08)	(0.18)	(0.49)
Own fin. judgment × fin. numeracy	-0.68*	-0.00	-0.02	0.02	-0.39***	-1.48***
	(0.38)	(0.02)	(0.03)	(0.05)	(0.12)	(0.34)
Higher vocational education	0.12	0.01	0.02	0.03	0.05	0.02
	(0.15)	(0.02)	(0.04)	(0.07)	(0.16)	(0.44)
Academic education	0.01	-0.01	-0.04	-0.05	0.09	0.00
	(0.15)	(0.02)	(0.04)	(0.08)	(0.18)	(0.50)
Age 41-64	-0.12	0.00	0.01	0.11	0.09	-0.10
	(0.21)	(0.03)	(0.04)	(0.09)	(0.20)	(0.54)
Age 65+	-0.03	0.04	0.15***	0.27***	0.31	0.31
	(0.24)	(0.03)	(0.05)	(0.10)	(0.22)	(0.60)
Log financial assets	-0.01	0.00	-0.01	-0.03	-0.05	-0.03
	(0.07)	(0.01)	(0.01)	(0.03)	(0.06)	(0.18)
Constant	0.49	-0.02	0.21	0.45	0.73	0.97
	(0.80)	(0.09)	(0.16)	(0.30)	(0.69)	(1.91)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 274. 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.11.

The results of the quantile regressions shown in the remaining columns of Table III demonstrate 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 numeracy among those who seek professional advice. Again, the results are similar for those who turn to family and friends, except for the point estimates at the top of the return loss distribution. The interesting results are again concentrated among those who rely on their own financial judgment. All else equal, their ninetieth percentile of the return loss is 145 basis points higher if they have a financial numeracy index of zero. Again, the effect becomes much worse for those with negative financial numeracy index values and reduces to 50 bp for those with the maximum financial numeracy score. The same pattern holds for the seventieth percentile with substantially smaller magnitudes. The variation of the coefficients across the quantiles is important 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 any form of personal advice nor have a high level of numerical skills as applied to financial matters.

The estimates reported in Table III are valid for the sample of participants in risky asset markets, but they may be different for the general population. Quantile regressions on the entire sample show that there does not seem to be a financial numeracy effect for those who seek others' advice at the seventy-fifth, eightieth, eighty-fifth, ninetieth, or ninety-fifth percentiles.[7] Again, there is no difference between the two groups seeking external advice and there is no relation with financial numeracy among these households. Those taking autonomous decisions with a financial literacy score of zero incur a consistently higher return loss at every quantile considered, reaching a magnitude of 72 basis points at the ninety-fifth percentile. The interaction effect is small and mostly insignificant over the first four quantiles considered; for the ninety-fifth percentile, the previous interpretation continues to hold. Larger financial wealth is associated with higher return loss across the distribution; this contrasts with a fairly precise zero in the sample of participants. While these differences are not significant, they nicely illustrate the varying interpretations in the two samples. The point estimates in Table III imply that participating households with high financial wealth incur similar return losses conditional on the other covariates. In the entire sample, another effect dominates: Participation is much more prevalent among the wealthy and they thus incur c.p. higher return losses.

[7]See Table C.1 in the Online Appendix. Note that the estimator requires the quantile to be strictly positive for all population groups; I thus only consider fairly high percentiles.

These results are robust to a number of other specification choices, which I document in the Online Appendix. Most importantly, there does not seem to be much of a relation between return loss and financial knowledge; including the latter as a regressor neither changes the magnitude or the significance of my findings. The same holds for various other covariates. When considering other measures of diversification loss, the same basic patterns emerge, although they are not always significant.

B. A detailed analysis of professional financial advice

The effects of financial advice have received a lot of attention over the last several years. Inderst and Ottaviani (2012) provide an overview of the regulation side of the market and a theoretical framework; recent empirical papers on the behavior of financial advisers (e.g., Mullainathan, Nöth, and Schoar, 2012) and properties of portfolios managed by advisers (e.g., Kramer, 2012; Hackethal, Haliassos, and Jappelli, 2012) show a rich variety of behavioral patterns. The general theme that appears to emerge is that advisers react to the typical incentive structure in the expected way by recommending high-fee products. The evidence on how this translates into investment outcomes is less clear-cut. My data offer a unique opportunity to complement those studies by investigating whether the same patterns hold for the general population and whether they vary with financial literacy.

Before turning to these issues, I consider the decision to seek professional advice. For example, Collins (2012) argues that financial advice and financial literacy are complements because the more literate have a higher propensity to receive professional advice in his data. I thus run a regression of turning to professional financial advisers on financial numeracy and on the same controls as in Table III. The amount of financial wealth is the only variable that is economically and statistically significant; in particular, the effect of financial numeracy is small and insignificant. This is not too surprising in light of the descriptive statistics, which show that financial numeracy is lowest among those who turn to family and friends for advice and highest among those who rely on their own judgment.[8] In the sense of Collins (2012), professional financial advice and financial numeracy are thus neither complements nor substitutes in my data. In addition, the picture is more complicated because the group that

[8]See Table A.7 in the Online Appendix. This pattern is also confirmed in multinomial logit regressions with the same controls. The only case where financial numeracy becomes significant there is a negative relation with turning to family and friends for advice.

does not seek professional financial advice is made up of two very distinct subpopulations.

In the next step, I regress the fees paid as a percentage of the entire portfolio on financial numeracy, financial advice, their interaction, financial knowledge, and a number of control variables. I relegate the full set of results to Table C.16 in the Online Appendix, as they are very easy to summarize: Average fees paid are about 70 bp per year; they rise by another 30 bp for those who rely on professional financial advice. This is the only effect that is statistically significant. Neither financial numeracy nor financial knowledge play any role.

When analyzing the relation between professional financial advice and expected investment outcomes, the treatment of mutual fund fees deserves special attention because of professional advisers' typical incentive structures. There are at least three different ways of dealing with fees in the return loss calculations. My analysis so far assumed that after-fee returns obey the CAPM. One could also, as Calvet, Campbell, and Sodini (2007) do in their main analysis, completely ignore them. The fact that fees are not return-dependent means that both of these procedures will lead to relatively similar results.[9] Both procedures correspond to fairly favourable views on active fund management. The opposite standpoint taken by many economists—that there is no way to beat an efficient market in terms of expected returns—can be achieved by ignoring fees in the CAPM implementation and adding them to the return loss ex post.

When re-doing the analysis of Table II—a cross-tabulation of the average return loss by source of advice and financial numeracy—under the latter approach, the number in the first cell rises by 35 bp to 1.35 percentage points, while all other numbers are at around 0.8 percentage points.[10] In Table IV, I repeat the analysis from Table III with fees added to the return loss. My previous results are left essentially unchanged. If anything, small differences between the two types of external advice emerge around the middle of the return loss distribution at low levels of financial literacy. Hence, even though professional advisers clearly recommend portfolios with higher average fees, the effects on the resulting portfolio characteristics are small compared to the impact of underdiversification.

[9] To see this, note that in a restriction-free decomposition of total portfolio risk into its idiosyncratic and systematic components (i.e., adding a constant to the basic CAPM regression (1), also see Section B of the Online Appendix), both will yield identical results. Forcing the regression line through the origin breaks the identity, but the results will still be similar in many cases.

[10] See Table C.17 in the Online Appendix.

Table IV
Contributors to return loss, with fees added to the return loss ex post

	OLS	p_{10}	p_{30}	p_{50}	p_{70}	p_{90}
Prof. advice × fin. numeracy	0.07	-0.04	0.06	0.12	0.17	0.21
	(0.05)	(0.04)	(0.08)	(0.09)	(0.19)	(0.54)
Financial advice: Family/friends	-0.02	0.05	-0.18	-0.23*	-0.28	-0.13
	(0.18)	(0.06)	(0.12)	(0.14)	(0.30)	(0.84)
Advice fam./friends × fin. numeracy	0.24	0.13	0.17	0.31*	0.35	0.73
	(0.17)	(0.08)	(0.14)	(0.17)	(0.36)	(1.01)
Financial advice: Own judgment	0.43**	-0.00	-0.08	-0.08	0.19	1.10*
	(0.20)	(0.05)	(0.09)	(0.10)	(0.22)	(0.61)
Own fin. judgment × fin. numeracy	-0.72*	-0.01	-0.11*	-0.08	-0.42***	-1.52***
	(0.39)	(0.03)	(0.06)	(0.07)	(0.15)	(0.43)
Higher vocational education	0.13	-0.03	0.09	0.11	0.10	-0.20
	(0.16)	(0.04)	(0.08)	(0.10)	(0.20)	(0.56)
Academic education	0.06	-0.03	-0.01	0.18*	0.28	-0.08
	(0.17)	(0.05)	(0.09)	(0.11)	(0.23)	(0.64)
Age 41-64	-0.05	0.01	0.10	0.23**	0.36	-0.37
	(0.22)	(0.05)	(0.10)	(0.12)	(0.24)	(0.69)
Age 65+	0.13	0.12**	0.31***	0.57***	0.53*	-0.11
	(0.26)	(0.06)	(0.11)	(0.13)	(0.27)	(0.76)
Log financial assets	-0.00	0.00	-0.06*	-0.09**	-0.07	-0.01
	(0.08)	(0.02)	(0.03)	(0.04)	(0.08)	(0.23)
Constant	0.73	0.04	0.98***	1.32***	1.34	1.94
	(0.83)	(0.19)	(0.34)	(0.41)	(0.86)	(2.41)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets (adding mutual fund fees ex post to the return loss); see Sections I.C and II.B for details. The number of observations is 274. 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.10.

C. Discussion

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. Perhaps the most convincing interpretation of this pattern is overconfidence: The individuals with the highest risk to incur large return losses trust their own capabilities more than those of others and they appear to overestimate the

former. Consistent with this interpretation, self-rated financial knowledge among individuals with below-median financial numeracy is highest among those who rely on their own financial judgment.[11] Coming back to the motivating question at the very end of Section I, the lower return losses among the financially wealthy may be explained by a lower prevalence of households with low financial literacy that exclusively rely on their own judgment.[12]

These results help answer the question posed by Korniotis and Kumar (2013) of whether *portfolio distortions reflect superior information or psychological biases*. Nieuwerburgh and Veldkamp (2010) recently 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 in order to exploit their informational advantage. My results are consistent with such an explanation for portfolios in the middle-upper region of the return loss distribution. However, the basic pattern behind the highest return losses makes it unlikely that superior information was the driving force behind the choice of portfolios. Similar arguments hold for other explanations of low diversification as a rational strategy: for example, Roussanov's (2010) "beat the Jones" argument. Rather, as suggested by Kimball and Shumway (2010), they are more likely to reflect investment mistakes.

In this respect, the consequences of low financial literacy have received a lot of attention recently, especially from the policy side (e.g. Lusardi, 2010). My results suggest that the majority of Dutch households 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.[13]

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 Rooij, Lusardi, and Alessie's (2011) basic financial literacy index—mostly simple maths quizzes worded in financial terms, which is why I have used the term "financial numeracy"—suggests an interpretation as a subcomponent of cognitive functioning. This has also been shown to correlate strongly with the decision

[11] See Table C.21 in the Online Appendix.

[12] See Table C.22 in the Online Appendix.

[13] Corroborative evidence comes from Choi, Laibson, and Madrian (2010), who show that the (self-reported) likelihood of changing one's mind after consulting an investment adviser is highest for those who make poor experimental investment decisions.

to participate in the stock market (Christelis, Jappelli, and Padula, 2010). It becomes increasingly difficult to compensate for low levels of cognitive skills after reaching adolescence (e.g. Cunha, Heckman, and Schennach, 2010, and the references therein); influencing current generations via this channel seems difficult. However, my results suggest 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 (2013) and Grinblatt, Keloharju, and Linnainmaa (2011, 2012), 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. However, note that 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 while unsolicited advice was ignored by investors. Nevertheless, expanding the availability of external guidance seems to be the most promising route. Academic economists' typical advice of investing in low-fee index funds competes with many attempts to guide households' behavior, where the form of the message is designed by professional marketing forces and its content is not necessarily adapted to suit the needs of consumers (Inderst and Ottaviani, 2009). Further research in that direction – how regulation can help shape correctly incentivized marketing forces – seems very promising in this light (see also Campbell, Jackson, Madrian, and Tufano (2011) for an elaboration of this point).

An additional point to keep in mind is that my analysis does not include costs associated with portfolio turnover. If professional advisers suggested mutual funds with higher loads and/or induced higher turnover rates, both of which would be consistent with typical incentive structures, the results may be very different. The evidence regarding portfolio turnover is mixed: Analyzing two samples of customers from an online broker and a bank, Hackethal, Haliassos, and Jappelli (2012) find that monthly turnover is roughly the same for all bank customers regardless of who manages the account. In contrast, turnover for professionally managed accounts was twice as high among the online broker's customers.

Tables C.23 and C.24 contain the results of some back-of-the-envelope calculations that may shed some light on the potential magnitudes. The analysis is crude for two reasons: First, I was only able to obtain data on loads for roughly one third of the funds in my sample

and ended up imputing the remaining ones. Second, I do not observe portfolio rebalancing in the data.I thus assume an annual turnover rate of 60%, which is roughly in line with the numbers reported by Hackethal, Haliassos, and Jappelli (2012). Assigning this to all households, no results are changed except for the scaling baseline values of the return loss. So once more, the effect of professional advisers recommending products with high fees does not seem to be a major factor. If I assign a 60% turnover rate to households that rely on professional advice and 30% to everybody else, notable differences emerge. In particular, the average return loss of all households that rely on professional advice would be in the same range as that of self-deciders with low numeracy, about half a percentage point higher than those in other groups. In order to resolve this ambiguity, future research should strive to collect a comprehensive dataset that contains detailed trading behavior in addition to the variables available in this study.

Switching from wishful thinking to the reality of typically available information, most datasets used in the area of household finance contain much less detail than the DNB Household Survey or the Scandinavian tax registers. Consequently, the question of how to make the most of limited data is of interest to a large number of researchers (see also Calvet, Campbell, and Sodini, 2009, in particular the Online Appendix). In Section D of the Online Appendix, I demonstrate that the detailed portfolio diversification measures constructed in this paper are needed to uncover the detailed patterns described herein. Among the commonly constructed proxies (e.g. those computed in Bilias, Georgarakos, and Haliassos, 2009; Graham, Harvey, and Huang, 2009; Guiso and Jappelli, 2009; or Kimball and Shumway, 2010) the fraction of the risky portfolio invested in mutual funds divided by the number of directly held shares (Guiso and Jappelli, 2009) has the highest correlation with the appropriate measure computed here.

III. Conclusions

Detailed portfolio information can be obtained fairly easily from survey respondents. I have shown that analysing the diversification properties yields results that are very similar to previous analyses on administrative records. This is important for a wide range of future research on household portfolios because until now, comprehensive and quantitatively meaningful analyses of portfolio diversification were limited to Scandinavian countries. A big

advantage of survey data is that they add a wide range of covariates waiting to be analyzed, which are not available in administrative datasets.

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 financial-numerical operations and concepts. The 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. A plausible interpretation of this pattern is overconfidence; the pattern also suggests that underdiversification most often reflects investment mistakes as opposed to optimal strategies. My results suggest that both financial numeracy and increasing the uptake of financial advice—be it from private contacts or professionals—are potential starting points for policies seeking to reduce welfare losses from inferior investment strategies.

References

Agarwal, Sumit and Bhashkar Mazumder (Jan. 2013). "Cognitive Abilities and Household Financial Decision Making". In: *American Economic Journal: Applied Economics* 5.1, pp. 193–207.

Alessie, Rob, Stefan Hochgürtel, and Arthur van Soest (2002). "Household Portfolios in the Netherlands". In: *Household Portfolios*. Ed. by Luigi Guiso, Michael Haliassos, and Tullio Jappelli. MIT Press, pp. 340–388.

— (Aug. 2004). "Ownership of Stocks and Mutual Funds: A Panel Data Analysis". In: *Review of Economics and Statistics* 86.3, pp. 783–796.

— (Aug. 2006). "Non-Take-Up of Tax-Favored Savings Plans: Evidence from Dutch Employees". In: *Journal of Economic Psychology* 27.4, pp. 483–501.

Bilias, Yannis, Dimitris Georgarakos, and Michael Haliassos (Oct. 2009). "Equity Culture and the Distribution of Wealth". Available at `http://ssrn.com/abstract=731805`.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini (2007). "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes". In: *Journal of Political Economy* 115.5, pp. 707–747.

— (2009). "Measuring the Financial Sophistication of Households". In: *American Economic Review* 99.2, pp. 393–398.

Campbell, John Y., Howell E. Jackson, Brigitte C. Madrian, and Peter Tufano (2011). "Consumer Financial Protection". In: *Journal of Economic Perspectives* 25.1, pp. 91–114.

Choi, James J., David Laibson, and Brigitte C. Madrian (Apr. 2010). "Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds". In: *Review of Financial Studies* 23.4, pp. 1405–1432.

Christelis, Dimitris, Tullio Jappelli, and Mario Padula (2010). "Cognitive Abilities and Portfolio Choice". In: *European Economic Review* 54.1, pp. 18–38.

Collins, J. Michael (Apr. 2012). "Financial Advice: A Substitute for Financial Literacy?" Mimeo, University of Wisconsin-Madison. Available at `http://ssrn.com/abstract=2046227`.

Cunha, Flavio, James J. Heckman, and Susanne Schennach (2010). "Estimating the Technology of Cognitive and Noncognitive Skill Formation". In: *Econometrica* 78.3, pp. 883–931.

Dimmock, Stephen G. and Roy R. Kouwenberg (June 2010). "Loss-Aversion and Household Portfolio Choice". In: *Journal of Empirical Finance* 17.3, pp. 441–459.

Goetzmann, William N and Alok Kumar (2008). "Equity Portfolio Diversification". In: *Review of Finance* 12.3, pp. 433–463.

Graham, John R., Campbell R. Harvey, and Hai Huang (July 2009). "Investor Competence, Trading Frequency, and Home Bias". In: *Management Science* 55.7, pp. 1094–1106.

Grinblatt, Mark, Matti Keloharju, and Juhani Linnainmaa (Dec. 2011). "IQ and Stock Market Participation". In: *Journal of Finance* 66.6, pp. 2121–2164.

— (May 2012). "IQ, Trading Behavior, and Performance". In: *Journal of Financial Economics* 104.2, pp. 339–362.

Guiso, Luigi and Tullio Jappelli (Jan. 2009). "Financial Literacy and Portfolio Diversification". University of Naples, CSEF Working Paper no. 212.

Hackethal, Andreas, Michael Haliassos, and Tullio Jappelli (2012). "Financial Advisors: A Case of Babysitters?" In: *Journal of Banking & Finance* 36.2, pp. 509–524.

Hung, Angela A., Andrew M. Parker, and Joanne K. Yoong (Sept. 2009). "Defining and Measuring Financial Literacy". RAND Labor and Population Working Paper 708.

Hung, Angela A. and Joanne K. Yoong (Jan. 2010). "Asking for Help: Survey and Experimental Evidence on Financial Advice and Behavior Change". RAND Labor and Population Working Paper 714.

Inderst, Roman and Marco Ottaviani (June 2009). "Misselling through Agents". In: *American Economic Review* 99.3, pp. 883–908.

— (June 2012). "Financial Advice". In: *Journal of Economic Literature* 50.2, pp. 494–512.

Kimball, Miles S. and Tyler Shumway (Jan. 2010). "Investor Sophistication and the Home Bias, Diversification, and Employer Stock Puzzles". Mimeo, University of Michigan, available at `http://ssrn.com/abstract=1572866`.

Korniotis, George M. and Alok Kumar (2011). "Do Older Investors Make Better Investment Decisions?" In: *Review of Economics and Statistics* 93.1, pp. 244–265.

— (Feb. 2013). "Do Portfolio Distortions Reflect Superior Information or Psychological Biases?" In: *Journal of Financial and Quantitative Analysis* 48.1, pp. 1–45.

Kramer, Marc M. (2012). "Financial Advice and Individual Investor Portfolio Performance". In: *Financial Management* 41.2, pp. 395–428.

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

Lusardi, Annamaria and Olivia S. Mitchell (2007). "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.

Mullainathan, Sendhil, Markus Nöth, and Antoinette Schoar (Mar. 2012). "The Market for Financial Advice: An Audit Study". NBER Working Paper 17929.

Nieuwerburgh, Stijn van and Laura Veldkamp (2010). "Information Acquisition and Under-Diversification". In: *Review of Economic Studies* 77.2, pp. 779–805.

Pohlmeier, Winfried and Volker Ulrich (1995). "An Econometric Model of the Two-Part Decisionmaking Process in the Demand for Health Care". In: *The Journal of Human Resources* 30.2, pp. 339–361.

Rooij, Maarten C. J. van, Annamaria Lusardi, and Rob Alessie (Aug. 2011). "Financial Literacy and Stock Market Participation". In: *Journal of Financial Economics* 101.2, pp. 449–472.

Roussanov, Nikolai (2010). "Diversification and Its Discontents: Idiosyncratic and Entrepreneurial Risk in the Quest for Social Status". In: *Journal of Finance* 65.5, pp. 1755–1788.

Smith, James P., John J. McArdle, and Robert Willis (2010). "Financial Decision Making and Cognition in a Family Context". In: *The Economic Journal* 120.548, F363–F380.

StataCorp (2013). "Stata Statistical Software: Release 13". College Station, TX: StataCorp LP.

Tang, Ning, Olivia S. Mitchell, Gary R. Mottola, and Stephen P. Utkus (Dec. 2010). "The Efficiency of Sponsor and Participant Portfolio Choices in 401(K) Plans". In: *Journal of Public Economics* 94.11-12, pp. 1073–1085.

Online Appendix accompanying:

How does household portfolio diversification vary with financial literacy and financial advice?*

Hans-Martin von Gaudecker †

*Citation format: Hans-Martin von Gaudecker, [year], Internet Appendix to "How does household portfolio diversification vary with financial literacy and financial advice?," Journal of Finance [DOI STRING]. Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the authors of the article.

†Universität Bonn, Department of Economics, Adenauerallee 24-42, 53012 Bonn; Germany; hmgaudecker[at]uni-bonn[dot]de; tel. +49 228 73 9357.

Contents

A. Detailed data description

A.1. Data sources and previous studies on diversification

Recent research has made clear that a significant fraction of households hold widely underdiversified portfolios (Calvet, Campbell, and Sodini, 2007). That study and several related ones (Massa and Simonov, 2006; Calvet, Campbell, and Sodini, 2009a,b) are based on administrative records from Sweden, 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, Keloharju, and Linnainmaa, 2011, 2012). 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, Georgarakos, and Haliassos, 2010) – 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 sheer number of variables is substantial in Sweden, their content often does not exactly cover what a researcher would like to know.

A very popular alternative for investigating individual investment behavior is to obtain data from discount brokers (Odean, 1998; Barber and Odean, 2000, 2001; Goetzmann and Kumar, 2008; Korniotis and Kumar, 2011; Hackethal, Haliassos, and Jappelli, 2012; Ivković, Poterba, and Weisbenner, 2005; Ivković, Sialm, and Weisbenner, 2008). 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 as long as they occurred with this broker. Consequently, many of the just-cited studies have focused on the implications of suboptimal trading behavior for portfolios' performance. These datasets are arguably less than optimal to study diversification issues because often only directly held stocks are observed in detail.[1] Furthermore, it is unknown (a) how much of households' portfolios the individuals' observed accounts cover and (b) to what ex-

[1]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.

tent holders of discount brokerage accounts are representative of the population of interest. Tang, Mitchell, Mottola, and Utkus (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 under-diversification, 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 recent contributions focussing on diversification issues include Christelis, Georgarakos, and Haliassos (2011) or Polkovnichenko (2005). Arguably the main strength of the SCF and general-purpose datasets with a strong module on financial matters[2] 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[3] – some investors achieve good diversification results with a low number of stocks while some mutual funds purposefully concentrate their investments in very specific assets. Calvet, Campbell, and Sodini (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 the relative Sharpe ratio loss (for details on this measure, see Section B in this Online Appendix) 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; one would hope for

[2]Some important examples are the HRS or PSID in the U.S. or the SHARE, ELSA, or ECHP datasets in Europe. Juster, Smith, and Stafford (1999), Christelis, Jappelli, and Padula (2010), Bilias, Georgarakos, and Haliassos (2010), and Christelis, Georgarakos, and Haliassos (2010); or Lusardi and Mitchell (2007a) are some exemplary studies using and describing these datasets for related questions.

[3]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, Harvey, and Huang, 2009), or hybrids thereof (Guiso and Jappelli, 2006, 2009).

better measures when diversification issues play the central role in an analysis.

A.2. Household portfolio data

Table A.1 contains an overview of the relevant asset and debt categories. The 35 entries in the first column are clearly too many to analyze 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. Even though the properties of bonds and options are clearly different, there is no useful information beyond the quantity held in the data for my purposes. In particular, there is no way to link them financial assets with observed properties, so I need to impute the quantities in my main analysis. Fortunately, the magnitude of holdings is fairly small.

Figure A.1 shows risky asset ownership along with unconditional and conditional shares by total net worth.[4] The left panel shows the standard pattern of rising ownership rates in wealth (Guiso, Haliassos, and Jappelli, 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 hump-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, Campbell, and Sodini, 2007, 2009b) and highlights the importance of establishing results along those studies' lines for other countries.

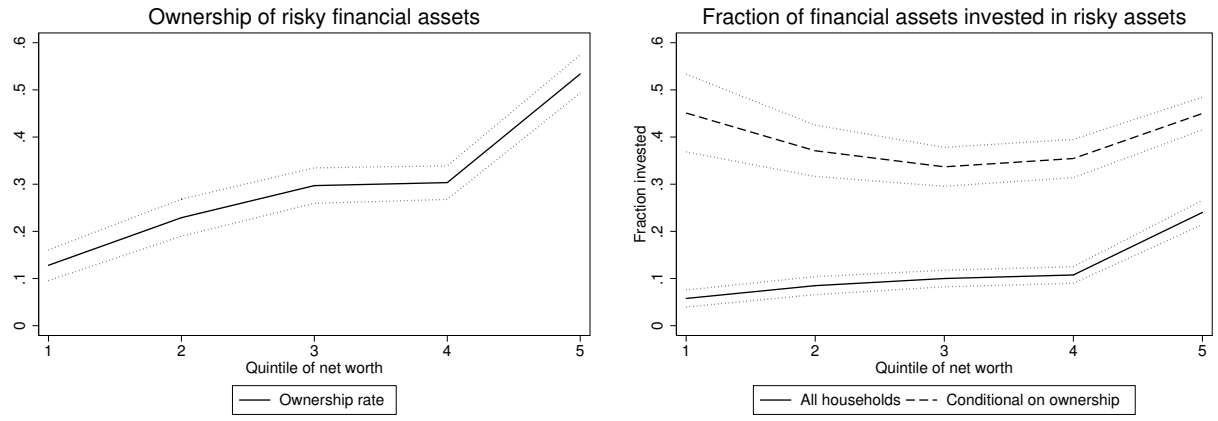
Table A.2 describes in detail the reductions in sample size that occur when considering only risky asset holders and matching the names of individual stocks and mutual funds to return series. Note that these names are not part of the data that are available by default. They may be obtained from CentERdata for a small administrative charge. Subjects are

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

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
Bonds (sovereign, corporate, mortgage)	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

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



Source: CentERpanel, own calculations. The dotted lines represent 95% confidence bands.

asked for the details of their 10 (5, 5) largest positions of stocks (mutual funds, growth funds[5]). The assets that belong to these categories make up by far the largest share of all risky financial assets.

The second and third row of Table A.2 reveal that of the 527 households that 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 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. Polkovnichenko, 2005; Bilias, Georgarakos, and Haliassos, 2009) or Sweden (Calvet, Campbell, and Sodini, 2007).

The next row shows that I can match close to 3 items on average to historical returns on Datastream and Euroinvestor. 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.[6] These are consolidated in row 6 and in all of my analyses,

[5]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.

[6]The questionnaire was fielded to all household members so each should name their personal assets; the

Table A.2
Sample selection and matching of portfolio components to return series

	Variable	# Obs.	Mean	p_5	p_{50}	p_{95}
(1)	Total number of households	1604				
(2)	Owners of risky financial assets	527				
(3)	Owners of shares/funds	499				
(4)	Raw report of individual items	497	3.34	1	2	10
(5)	Raw report (hh. in final sample)	408	3.69	1	2	11
(6)	Matched report of individual items	408	2.96	1	2	9
(7)	Fraction of shares/funds covered	408	0.90	0.28	0.99	1.00
(8)	Fraction of risky fin. assets covered	408	0.84	0.18	0.99	1.00
(9)	Fraction of quantities imputed	408	0.08	0.00	0.00	0.75

Source: CentERpanel, own calculations. All statistics are adjusted for clustering at the household level and use appropriate sampling weights.

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, 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 A.2 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

financial decider of the household additionally was asked for the common belongings as well. There are very few duplicate items, and there is no way to distinguish whether household members misinterpreted the instructions or gave correct answers.

Table A.3
Comparison of the different samples in terms of observable characteristics

Variable	Entire sample # Obs. 1604		Owners of risky financial assets # Obs. 511		Sample with detailed portfolios # Obs. 374	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
No/elementary/secondary education	0.60	.	0.47	.	0.46	.
Higher vocational education	0.26	.	0.31	.	0.31	.
Academic education	0.15	.	0.22	.	0.23	.
Age 26-40	0.31	.	0.18	.	0.19	.
Age 41-64	0.48	.	0.55	.	0.55	.
Age 65+	0.21	.	0.26	.	0.27	.
Value of financial assets	41,005	75,808	88,211	117,859	85,385	115,100
Value of non-financial assets	172,173	198,780	238,710	226,409	235,051	228,165
Value of debt	55,229	83,125	65,790	91,824	65,300	91,369
Net annual household income	31,656	37,466	39,978	61,988	40,024	63,519
Female	0.37	.	0.27	.	0.26	.
Household size	2.36	1.30	2.38	1.33	2.37	1.33
Degree of urbanisation	0.00	1.00	0.01	0.98	0.01	0.98

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 are for the financial decider. The last column contains the statistics for all households whose reports of individual portfolio items cover more than 30% of their entire risky portfolio, hence the number of observations is lower than in Table A.2.

results—which are independent of portfolio size—would arise from these households.

The statistics in Table A.3 show that there does not seem to be selection on observables into reporting detailed portfolio items or not. In particular, comparing the first set of columns on all household with the last set reveals the usual and striking differences between the entire sample and holders of risky assets. But the samples with all owners of risky assets (the second row of Table A.2) and those who end up in my final sample of households with detailed information on the time series of asset returns are virtually identical.

Table A.4
Descriptive statistics on the time series of asset returns and mutual fund fees

Variable	# Obs.	Mean	p_5	p_{50}	p_{95}
Length of time series of returns	269	137.72	57	128	235
Total expense ratio, mutual funds	170	1.30	0.33	1.27	1.87

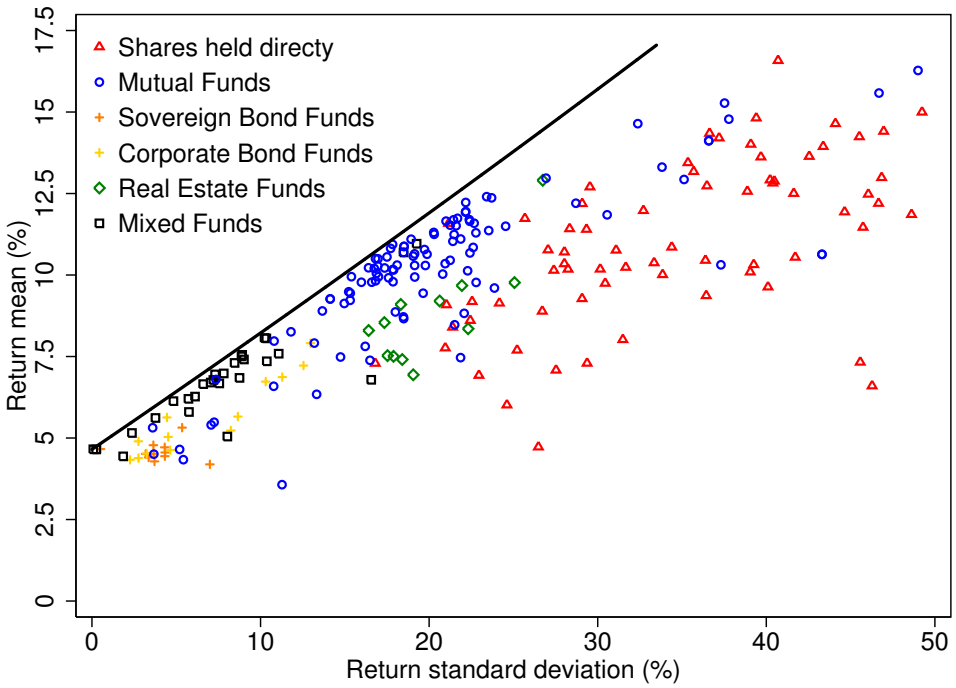
Source: Datastream, Euroinvestor, own calculations. In the case of the time series of asset returns, the number of observations refers to the number of different assets. Returns are observed at a monthly frequency. The total expense ratio is expressed as an annual percentage of the asset value.

A.3. Descriptive statistics on funds and shares

Table A.4 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, Campbell, and Sodini (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 charges of 140 mutual funds via Morningstar or a fund's prospectus. For another 20, I assigned the fee of similar funds managed by the same company. For the remaining 10 funds, I imputed the costs from the distribution of available fees. The second line of Table A.4 reveals charges in the usual range with an annual average of 130 basis points. In the estimations containing mutual fund fees I subtract 30 basis points from the benchmark index. This is approximately the cost charged by index funds replicating common benchmarks.

Figure A.2 and Table A.5 show the diversification properties of these 269 individual items that make up households' portfolios, differentiating between types of funds. Expected returns and standard deviations are calculated as outlined in Section B in the main text. Not surprisingly, single shares bear by far the largest idiosyncratic risk. However, not all mutual funds follow a diversification strategy. The ones that are found at the right of the graph largely consist of IT-funds, for example, investing in a highly correlated set of shares.

Figure A.2. The mean-standard deviation characteristics of shares and funds in the sample



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

The remaining types of funds follow much less risky strategies on average, also yielding lower expected returns. Strikingly, funds investing in a mix of the other categories line up much closer to the efficient frontier than bond funds or real estate funds. This is reflected in the fairly high return losses of bond funds and real estate funds in the last column of Table A.5. These numbers are likely driven more by the choice of the MSCI Europe as the benchmark, which is treating these asset classes unfairly, than anything else. Unfortunately, it proved difficult to find Dutch or European bond and real estate indices that have a sufficiently long time series. In order to still be able to assess robustness to the type of benchmark, I assumed all risk inherent to bond and real estate funds to be systematic. Put differently, I moved each relevant point in Figure A.2 vertically to the efficient frontier. All conclusions remain unchanged, the results are available from the author upon request. .

Table A.5
Descriptive statistics on the shares and funds held by the CentERpanel respondents

	# Obs.	Expected Return	Standard Deviation	Return Loss
Shares held directy	99	12.23	47.06	9.53
Equity Funds	106	10.20	20.98	2.08
Sovereign Bond Funds	11	4.58	3.94	1.51
Corporate Bond Funds	13	5.62	6.82	1.51
Real Estate Funds	12	8.77	20.14	3.21
Mixed Funds	28	6.81	7.95	0.73

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The last three columns show averages for each fund category. All quantities are expressed as annual percentage points.

A.4. Inputs to investment production

In this section, I motivate the sets of covariates I use in my estimations of (3) and provide detailed descriptions of the variables. The financial literacy module, obtained from Rooij, Lusardi, and Alessie (2011), is similar to the one used in Lusardi and Mitchell (2007b); the questions are discussed in detail and compared to other measures in Hung, Parker, and Yoong (2009).

A.4.1. Financial numeracy

A good example of the financial numeracy component (also called "basic financial literacy" by previous authors) is the question on interest compounding:

[Q] Suppose you had €100 in a savings account and the interest rate is 20% per year and you never withdraw money or interest payments. After 5 years, how much would you have on this account in total?

[A]
1. More than €200
2. Exactly €200
3. Less than €200
4. Do not know
5. Refusal

Table A.6 shows that 81% of respondents correctly answer this question. Except for interest compounding, the other questions have a similar structure of straightforward maths quizzes applied to financial matters, see Rooij, Lusardi, and Alessie (2011) for the full list. Similar to that paper, I assume that a single factor is underlying the answers to the five questions and normalize 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 code those answers with the probability that a randomly chosen answer would be correct. For the question above, this would thus be $1/3$. This can be rationalized 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. I use the term "financial numeracy index" for

Table A.6
Detailed documentation of the financial literacy modules

Variable	Entire sample # Obs. 925		Sample with detailed portfolios # Obs. 274	
	Mean	Std. dev.	Mean	Std. dev.
Simplest Numeracy	0.94	0.22	0.97	0.16
Interest Compounding	0.81	0.38	0.86	0.34
Inflation	0.89	0.29	0.95	0.21
Time Value Of Money	0.78	0.40	0.88	0.31
Money Illusion	0.72	0.44	0.74	0.43
Function Of The Stock Market?	0.75	0.40	0.87	0.32
What Does Stock Ownerwhip Mean?	0.70	0.44	0.80	0.40
What Do Mutual Funds Do?	0.78	0.38	0.90	0.28
What Does A Company Bond Do?	0.68	0.42	0.82	0.37
Bond Prices After Int. Rate Change?	0.38	0.42	0.53	0.47
Diversification Stock Vs. Mut. Fund	0.66	0.41	0.77	0.40
Bonds Or Stocks More Risky?	0.77	0.35	0.89	0.29
Equity Premium	0.58	0.45	0.76	0.42
Volatility Different Assets	0.79	0.36	0.92	0.27
Diversification, Direct Question	0.75	0.40	0.87	0.32
Bond Liquidity	0.55	0.40	0.68	0.42

Source: CentERpanel, own calculations. All statistics are adjusted for sampling weights, standard deviations of dummy variables are not shown. Variables refer to the financial deciders of households. The last two columns contain the statistics for all households whose reports of individual portfolio items cover more than 30% of their entire risky portfolio. The exact wording of the questions can be found in Rooij, Lusardi, and Alessie (2011).

the resulting measure, all results are robust to using the exact procedure of Rooij, Lusardi, and Alessie (2011).

Note that the survey instrument resembles to some extent the numeracy component of the cognitive ability score used in Christelis, Jappelli, and Padula (2010) to explain stockholding. The first row of Table A.7 shows that participants in risky asset markets have a significantly higher financial numeracy score, confirming one of the main results of Rooij, Lusardi, and Alessie (2011) and Christelis, Jappelli, and Padula (2010). 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 0.64.

A.4.2. Financial knowledge

The second part of the financial literacy module 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:

[Q] Buying a company stock usually provides a safer return than a stock mutual fund. True or false?

[A]
1. True
2. False
3. Do not know
4. Refusal.

Only two thirds of all respondents give the correct answer, but three quarters of those who participate in risky asset markets do. Based on a total of 11 such questions, I construct an index of financial knowledge in the same way as the financial numeracy index. Row 2 in Table A.7 confirms a main result from Rooij, Lusardi, and Alessie (2011), namely that financial knowledge is positively correlated with holding risky assets. Conceptually, this index might be more problematic as an input in (3) than basic maths skills. The reason is that they may be largely shaped by investor experience – one would expect the probability of a correct answer to the example question to be higher 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

Table A.7
Descriptive statistics on the covariates

Variable	Entire sample # Obs. 925		Sample with detailed portfolios # Obs. 274	
	Mean	Std. dev.	Mean	Std. dev.
Financial numeracy index	-0.01	1.01	0.25	0.77
Financial knowledge index	-0.01	1.00	0.57	0.74
High self-rated fin. knowledge	0.19	.	0.28	.
Financial advice: Professionals	0.23	.	0.23	.
Financial advice: Family/friends	0.24	.	0.14	.
Financial advice: Own judgement	0.52	.	0.62	.
Prof. advice × fin. numeracy	0.00	1.05	0.24	0.96
Advice fam./friends × fin. numeracy	-0.15	1.03	0.17	0.64
Own fin. judgement × fin. numeracy	0.06	0.98	0.28	0.72
No/elementary/secondary education	0.60	.	0.46	.
Higher vocational education	0.25	.	0.31	.
Academic education	0.15	.	0.23	.
Age 26-40	0.31	.	0.19	.
Age 41-64	0.48	.	0.55	.
Age 65+	0.21	.	0.27	.
Value of financial assets	39,200	72,557	85,385	115,100
Log financial assets	9.66	1.38	10.72	1.19
Value of non-financial assets	169,724	198,164	235,051	228,165
Log non-financial assets	10.19	2.93	11.00	2.62
Value of debt	54,867	82,755	65,300	91,369
Log debt	7.45	3.97	7.95	3.99
Net annual household income	31,521	37,659	40,024	63,519
Log net household income	10.21	0.49	10.41	0.51
High tolerance for risky investm.	-0.00	1.00	0.47	0.95
Female	0.37	.	0.26	.
Household size	2.35	1.30	2.37	1.33
Degree of urbanisation	-0.00	1.00	0.01	0.98

Source: CentERpanel, own calculations. All statistics are adjusted for sampling weights, standard deviations of dummy variables are not shown. The interaction terms in the third part of the table give averages of financial literacy within each category of financial advice. Variables relating to individuals rather than the household are for the financial decider. The numbers of observations shown in the table header refer to the preferred specification. When including all covariates, the numbers reduce to 743 and 238, respectively. The last two columns contain the statistics for all households whose reports of individual portfolio items cover more than 30% of their entire risky portfolio.

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 dichotomize the four-point rating into a dummy variable by merging the top two categories (very knowledgeable, knowledgeable) and the bottom two categories (more or less knowledgeable, not knowledgeable). Self-rated financial knowledge is high for 19% (28%) of all households (participants in risky asset markets).

A.4.3. Financial advice

The most important source of financial advice is directly asked for in the DHS questionnaire. Individuals were asked about their most important source of financial information and could pick one of the categories shown in the second part of Table A.7. The numbers there reveal that about a quarter of respondents seek help from professional advisers 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 adviser; advertisements; financial computer programs; the Internet; other. I label the aggregate category "reliance on own financial judgment".[7] 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. Comparing the level of financial literacy among the different groups of advice-seeking is especially interesting. It is significantly lower among those who ask their friends and family compared to any of the two other groups. The same pattern emerges for holders of risky assets, although the sample size compromises statistical significance.

A.4.4. Other variables

The remaining inputs to the production function are additional controls that serve to sharpen the interpretation of the financial literacy and advice variables. First, the financial literacy

[7]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 advisers. Presumably, a financial institution's advisers and brochures would recommend similar investment strategies. I prefer the classification I chose because brochures seem to focus on advertising specific investments, while advisers would (hopefully) make recommendations based on the entire portfolio. See also Inderst and Ottaviani (2012). In any case, all results survive a reclassification.

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 financial numeracy index) or general 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, 2011; Agarwal, Driscoll, Gabaix, and Laibson, 2009) and the need for advice may be lower. Third, one may be worried that high diversification losses predominantly occur for households with very low financial wealth; for example because other assets are much more important for them so they allocate little thought to their investments. I thus include the logarithm of financial wealth in the main analysis.

The covariates mentioned so far comprise my preferred specification; I include more measures in the robustness checks. 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.:

[Q] It is more important to have a safe investment with guaranteed returns than taking risk.

[A] 1. Totally disagree 2. Disagree 3. Partly disagree 4. Neither agree or disagree 5. Partly agree 6. Agree 7. Totally agree.

I add up the answers and standardize 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 A.7 once more. Including these variables in the regression has the drawback of reducing the sample size by about 15%; doing so adds additional credibility to the results because preference-based risk-taking implies greater scope for incurring diversification losses, particularly in case of the return loss measure. I include a gender dummy since women are more risk averse than men on average (Croson and Gneezy, 2009) and for comparisons with the literature.

In order to compare my results to those of Calvet, Campbell, and Sodini (2007), I furthermore include measures of household size, the degree of urbanisation, household income, additional wealth components, and liabilities in various other specifications. The reason for not incorporating these variables in my standard specification is that their interpretation

in the production function framework (3) 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. households in urban areas might have easier access to financial advice). I discuss the results in the text and all corresponding tables can be found in the Appendix.

B. Detailed comparison with Calvet, Campbell, and Sodini (2007, 2009b)

I follow the exposition in Calvet, Campbell, and Sodini (2007) rather closely and sketch their measures of diversification loss briefly in order to keep the text self-contained. The interested reader is referred to Calvet, Campbell, and Sodini (2007) and its Online Appendix for further details.

In a first step, I decompose total portfolio risk into a systematic and an idiosyncratic component. As in my main analysis, 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 $r^e_{b,t}$:

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

The MSCI Europe index continues to be the benchmark; results are robust to using the excess returns of the AEX or the MSCI World Index instead. The decomposition of the household's total portfolio risk σ^2_h into a systematic σ^2_b and an idiosyncratic $\sigma^2_{h,\text{idios.}}$ component is then given by:

$$\sigma^2_h = \beta^2_h \cdot \sigma^2_b + \sigma^2_{h,\text{idios.}} \tag{1}$$

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.

Once we estimate a CAPM, an alternative to the return loss is the relative Sharpe ratio loss, defined as

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

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. The relative Sharpe ratio loss has a number of attractive features for the analysis of risky financial asset portfolios (see Calvet, Campbell, and Sodini, 2007). However, the return loss is more comprehensive, since a poorly diversified risky portfolio will not lead an investor far astray from the efficient frontier if the share invested in risky assets is very

low. Put differently, RSL_h is independent of the risky asset share; this is not necessarily desirable.

Another measure that is closely related to the return loss on the entire financial portfolio is the return loss on the risky financial asset portfolio only. It amounts to setting ω_h to 1 in (2); in graphical terms, it is the vertical distance between a household's portfolio and the efficient frontier in Panel C of Figure 1 in the main text.

B.1. Results for the risk decompositions

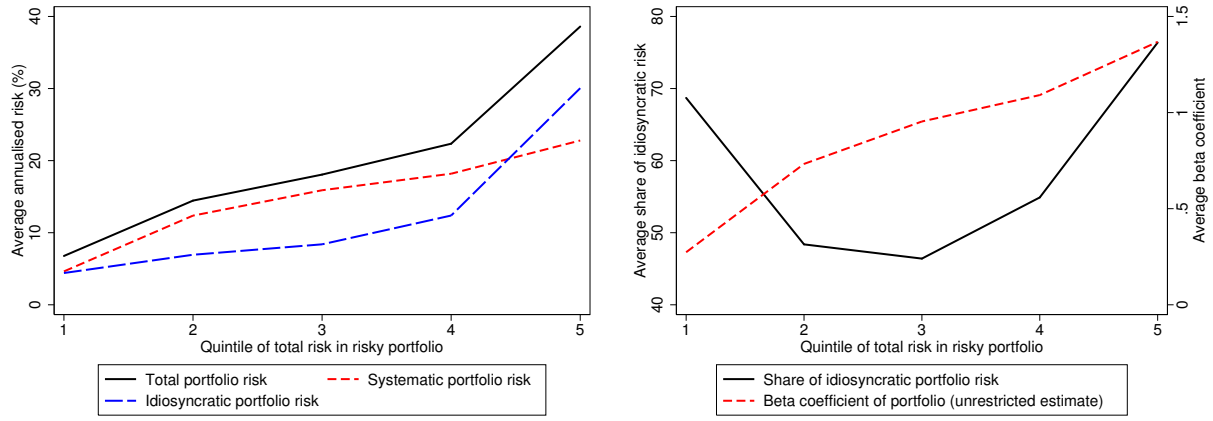
For all participants in risky asset markets, Figure B.1 contains plots of several measures for each quintile of the distribution of total portfolio risk, as inferred from equation (1). Shown in the left panel, total portfolio risk rises from less than 10% annually in the lowest quintile to almost 40% 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, Campbell, and Sodini (2007) for Swedish households – they report 11% (19.5%, 36.4%) at the 10^{th} (50^{th}, 90^{th}) percentiles for total portfolio risk.[8] 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 funds, which display a low correlation with the benchmark index. This is also reflected in the average β coefficient inferred from (1), which rises strongly in total portfolio risk. Again, the magnitudes are very close to those reported in Calvet, Campbell, and Sodini (2007).

B.2. Results for CAPM-based measures

The quintile-specific values of the relative Sharpe ratio loss (2) and of the return loss (2) are plotted in Figure B.2. 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

[8]Given the huge sample size, Calvet, Campbell, and Sodini (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 B.1. Portfolio risk components by quintile of total portfolio risk



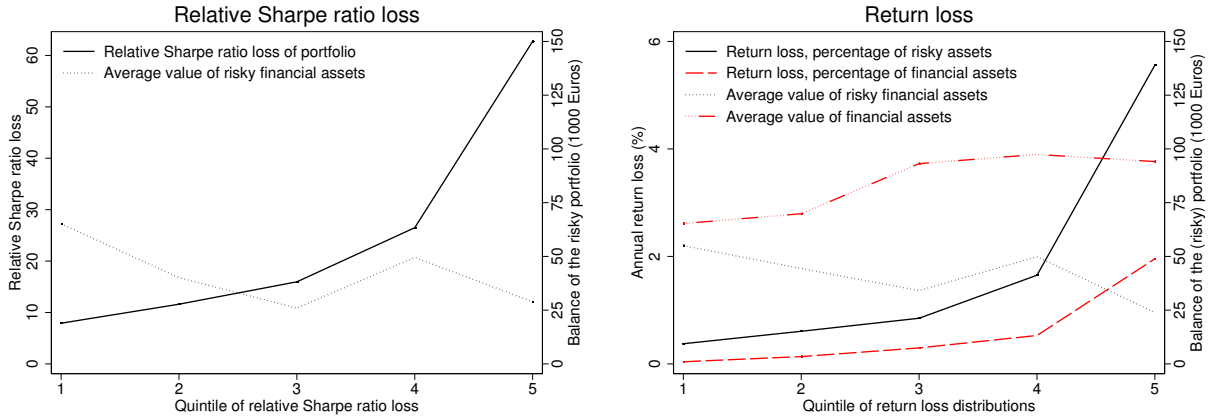
Source: CentERpanel, Datastream, Euroinvestor, own calculations.

upper quintiles. Again, this pattern mirrors the findings of Calvet, Campbell, and Sodini (2007) very closely: on the one hand, most households minimize inefficient risk-taking; on the other hand, almost two thirds of all the risk the average household in the top quintile takes remains uncompensated. I also plot average financial wealth in each quintile in the same figure. It shows that the least efficient portfolios are fairly small on average, but by no means negligible.

The right panel of Figure B.2 contains two measures of return loss. The solid black line just considers the risky portfolio. The average return households lose on their risky portfolio compared to an efficient investment equals 178 basis points per year, which is just above the number reported by Calvet, Campbell, and Sodini (2007) for the unhedged world index as the benchmark. It is below 100 basis points for the bottom three quintiles and 585 basis points in the highest.

Mirroring the results from Table I in the main text, these return losses are far lower when the entire portfolio is taken as the basis. This is depicted by the dashed red line in the right panel of Figure 1. Again, return losses are very limited for the first four quintiles (52 basis points in the fourth), but reach the substantial amount of more than 180 bp in the highest quintile. The average is about 55 basis points, substantially less than the 178 basis points for the risky portfolio multiplied with the risky asset share of .38 (see Figure A.1). This implies a negative covariance between the risky asset share on the one hand; and the

Figure B.2. Mean-variance measures of diversification losses



Source: CentERpanel, Datastream, Euroinvestor, own calculations.

product of the losses from underdiversification and the household portfolio's β on the other hand.[9] It 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 two lines on the size of the risky and financial portfolios show that return losses are not concentrated among small-scale investors.

B.3. Results of return loss regressions

Calvet, Campbell, and Sodini (2007) take the logarithm of (2) 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 β, 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, the scaling of the logarithmic regressions seems less useful to me than a direct estimate: Return losses on the

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

financial portfolio are 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 4 bp at the 10^{th} percentile and of 10 bp at the 25^{th} percentile are almost as far apart from each other as the 61 bp at the 75^{th} percentile and the 149 bp at the 90^{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.

Table B.1 contains the results from a regression model that mirrors Table 5 in Calvet, Campbell, and Sodini (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, Campbell, and Sodini (2007) is 29.3; for a sample size of 373, this would reduce to 1.79 on average. 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, Campbell, and Sodini (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. Having seen the plots in Figure A.1, this does not come as a surprise. Table B.2 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.

Table B.1
Contributors to return loss and its components in logarithms, covariates similar to Calvet, Campbell, and Sodini (2007)

	RL_h	ω_h	β_h	$\frac{S_b - S_h}{S_h}$
Higher vocational education	0.13	0.16	-0.04	-0.02
	(0.17)	(0.14)	(0.12)	(0.13)
Academic education	-0.00	0.15	-0.14	-0.01
	(0.23)	(0.19)	(0.19)	(0.17)
Age 41-64	0.18	0.19	-0.14	0.09
	(0.23)	(0.18)	(0.12)	(0.15)
Age 65+	0.51*	0.50**	-0.20	0.15
	(0.26)	(0.23)	(0.16)	(0.20)
Female	-0.31*	-0.22	-0.01	-0.04
	(0.19)	(0.17)	(0.10)	(0.13)
Household size	-0.09	-0.03	-0.00	-0.04
	(0.06)	(0.06)	(0.05)	(0.05)
Degree of urbanisation	0.04	0.00	-0.08	0.08
	(0.08)	(0.07)	(0.07)	(0.06)
Log net household income	-0.25	-0.15	-0.05	-0.06
	(0.16)	(0.15)	(0.11)	(0.11)
Log financial assets	-0.08	0.06	0.02	-0.14**
	(0.08)	(0.06)	(0.05)	(0.06)
Log non-financial assets	0.01	-0.01	-0.00	0.01
	(0.05)	(0.04)	(0.03)	(0.04)
Log debt	-0.00	-0.00	0.02	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)
Constant	-2.61*	-0.70	-0.02	0.79
	(1.53)	(1.51)	(0.84)	(1.07)
Adjusted R^2	0.035	0.028	0.002	0.024

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variables are the natural logarithms of the left-hand side of Equation (2) in the main text and of its constituents. See Section I.C for details. All regressions are estimated by OLS on both waves of data, adjusting for clustering at the household level and using sampling weights. The number of observations is 373; standard errors are in parentheses.

Table B.2
Contributors to return loss and its components in logarithms

	RL_h	ω_h	β_h	$\frac{S_b - S_h}{S_h}$
External advice × fin. numeracy	0.22**	0.30	-0.07	-0.02
	(0.10)	(0.22)	(0.09)	(0.16)
Financial advice: Own judgment	0.27	0.06	0.11	0.12
	(0.17)	(0.17)	(0.16)	(0.15)
Own fin. judgment × fin. numeracy	-0.24	-0.24	0.17	-0.14
	(0.22)	(0.15)	(0.19)	(0.14)
Higher vocational education	0.05	0.05	0.02	-0.04
	(0.20)	(0.16)	(0.13)	(0.15)
Academic education	-0.17	0.10	-0.24	-0.03
	(0.24)	(0.21)	(0.24)	(0.20)
Age 41-64	0.10	0.12	-0.13	0.08
	(0.26)	(0.21)	(0.16)	(0.18)
Age 65+	0.59**	0.56**	-0.20	0.16
	(0.28)	(0.23)	(0.19)	(0.21)
Log financial assets	-0.03	0.08	-0.02	-0.09
	(0.09)	(0.07)	(0.05)	(0.07)
Constant	-6.04***	-2.66***	-0.09	-0.53
	(0.95)	(0.77)	(0.53)	(0.68)
Adjusted R^2	0.033	0.051	0.013	0.004

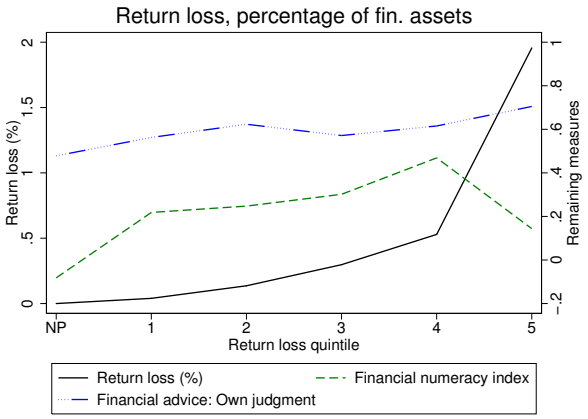
Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variables are the natural logarithms of the left-hand side of Equation (2) in the main text and of its constituents. See Section I.C for details. All regressions are estimated by OLS on both waves of data, adjusting for clustering at the household level and using sampling weights. The number of observations is 274; standard errors are in parentheses.

C. Additional material relating to Section II

C.1. Bivariate relations between financial literacy, financial advice, and return loss

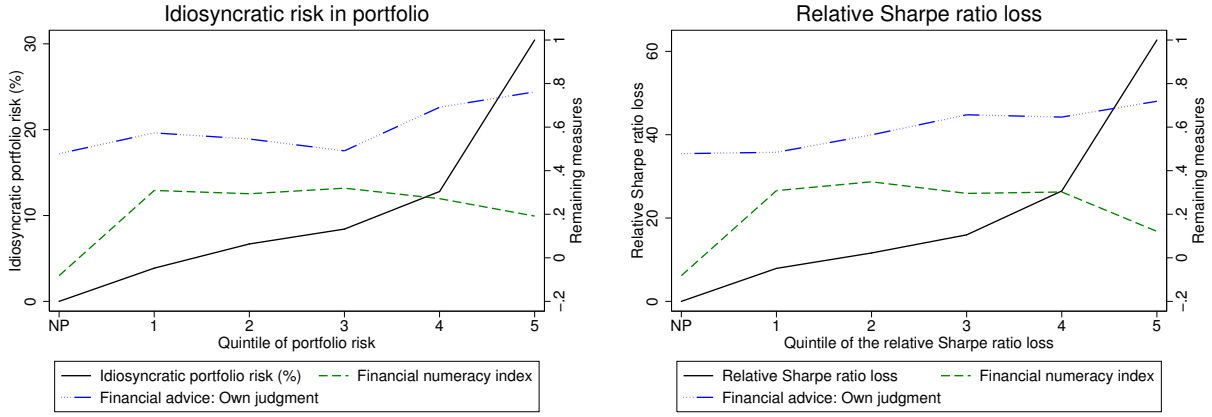
Figure C.1 plots the quintile-specific averages of the return loss, adding non-participants in risky asset markets (for whom $RL_h = 0$ since $w_h = 0$). For each return loss quintile, the figure furthermore adds the average values of the financial numeracy index and the share of households relying on their own financial judgment. Financial numeracy shows a hump-shaped pattern – it is lowest for the non-participants (previously shown by Rooij, Lusardi, and Alessie (2011) and Christelis, Jappelli, and Padula (2010)) and 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 (Table I in the main text), this suggests that low investment skills may play an important role in determining the worst outcomes. The same hump-shaped pattern is found for the amount of idiosyncratic risk or the relative Sharpe ratio loss, see Figure C.2. The fraction of individuals relying on their own financial judgments is generally rising in the return loss, although the high value in the second quintile is an exception to the rule.

Figure C.1. Financial literacy, financial advice, and diversification losses



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

Figure C.2. Financial literacy, financial advice, and diversification losses – alternative measures of diversification loss



Source: CentERpanel, Datastream, Euroinvestor, own calculations. The quintile values "NP" stand for non-participants.

C.2. Background information and robustness checks for Section II.A

C.2.1. Quantile regressions of the preferred specification for the entire sample

To see why results may differ between quantile regressions in the sample of participants and the entire population, assume there are two groups. 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 first group compared to the second group. Not doing so will reverse the order except for the highest quantiles. Table C.1 presents the results of including the non-participants in the estimation sample. The quantiles under consideration need to be strictly positive for the estimator to be well-defined, so results of this exercise are presented for every fifth percentile starting with the 75^{th}. As discussed in the main text, a comparison of Table III there and Table C.1 shows similar results for the main variables of interest; there appears to be an interesting selection effect for financial wealth.

Table C.1
Contributors to return loss. Non-participants are included in the estimation sample

	p_{75}	p_{80}	p_{85}	p_{90}	p_{95}
Prof. advice × fin. numeracy	0.01	0.01	0.01	0.02	-0.01
	(0.03)	(0.04)	(0.05)	(0.13)	(0.32)
Financial advice: Family/friends	0.01	0.03	0.04	0.03	0.15
	(0.05)	(0.06)	(0.08)	(0.19)	(0.47)
Advice fam./friends × fin. numeracy	0.00	0.01	0.01	0.00	0.02
	(0.03)	(0.04)	(0.06)	(0.13)	(0.32)
Financial advice: Own judgment	0.05	0.09*	0.12*	0.22	0.72*
	(0.04)	(0.05)	(0.07)	(0.15)	(0.38)
Own fin. judgment × fin. numeracy	0.02	0.03	0.04	0.07	-0.59***
	(0.02)	(0.03)	(0.04)	(0.09)	(0.22)
Higher vocational education	0.07*	0.08*	0.09	0.06	0.44
	(0.04)	(0.05)	(0.06)	(0.15)	(0.37)
Academic education	-0.01	-0.00	0.03	-0.04	-0.01
	(0.05)	(0.06)	(0.08)	(0.18)	(0.46)
Age 41-64	0.00	-0.01	-0.01	-0.02	-0.06
	(0.04)	(0.05)	(0.07)	(0.16)	(0.40)
Age 65+	0.08*	0.10*	0.11	0.13	0.38
	(0.05)	(0.06)	(0.08)	(0.18)	(0.46)
Log financial assets	0.06***	0.08***	0.09***	0.13***	0.19
	(0.01)	(0.01)	(0.02)	(0.05)	(0.12)
Constant	-0.50***	-0.62***	-0.71***	-0.93**	-1.40
	(0.12)	(0.15)	(0.20)	(0.47)	(1.18)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 925; standard errors are in parentheses. 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 C.2
Contributors to return loss, two categories of financial advice

	OLS	$\mathbf{p_{10}}$	$\mathbf{p_{30}}$	$\mathbf{p_{50}}$	$\mathbf{p_{70}}$	$\mathbf{p_{90}}$
External advice × fin. numeracy	0.04	-0.01	0.04	0.05	0.08	0.17
	(0.05)	(0.02)	(0.03)	(0.06)	(0.13)	(0.48)
Financial advice: Own judgment	0.49**	-0.00	0.06	0.02	0.33**	1.34**
	(0.19)	(0.02)	(0.04)	(0.07)	(0.14)	(0.53)
Own fin. judgment × fin. numeracy	-0.68*	-0.01	-0.02	0.01	-0.40***	-1.46***
	(0.38)	(0.02)	(0.03)	(0.06)	(0.12)	(0.44)
Higher vocational education	0.12	0.01	0.02	0.05	0.06	0.01
	(0.15)	(0.02)	(0.04)	(0.07)	(0.15)	(0.56)
Academic education	0.01	-0.01	-0.04	-0.05	0.07	0.01
	(0.15)	(0.02)	(0.04)	(0.08)	(0.18)	(0.64)
Age 41-64	-0.13	0.01	0.02	0.12	0.11	-0.39
	(0.20)	(0.03)	(0.05)	(0.09)	(0.19)	(0.69)
Age 65+	-0.04	0.04	0.13**	0.23**	0.30	-0.24
	(0.24)	(0.03)	(0.05)	(0.10)	(0.21)	(0.77)
Log financial assets	-0.01	-0.00	-0.01	-0.03	-0.05	-0.11
	(0.07)	(0.01)	(0.02)	(0.03)	(0.06)	(0.23)
Constant	0.59	0.03	0.20	0.49	0.76	2.26
	(0.75)	(0.09)	(0.16)	(0.30)	(0.64)	(2.35)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 274. 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.11.

C.2.2. Robustness checks

Tables C.2 and C.3—corresponding to Tables III and C.1, respectively—show that merging the two groups relying on advice from other persons does not change any of the results. This is not surprising since they were practically indistinguishable in all previous analyses. I thus include only one group that I call "external financial advice" in the robustness checks of this section in order to keep the tables more legible.

I first consider robustness to including a number of additional covariates in regressions similar to those underlying Tables C.2 and Tables C.3. In Tables C.4 and C.5, I add the

Table C.3
Contributors to return loss, two categories of financial advice. Non-participants are included in the estimation sample

	p_{75}	p_{80}	p_{85}	p_{90}	p_{95}
External advice × fin. numeracy	0.00	0.00	0.01	0.00	-0.00
	(0.02)	(0.03)	(0.04)	(0.09)	(0.24)
Financial advice: Own judgment	0.04	0.06	0.10*	0.20	0.63*
	(0.03)	(0.04)	(0.05)	(0.12)	(0.33)
Own fin. judgment × fin. numeracy	0.02	0.03	0.05	0.07	-0.60***
	(0.02)	(0.03)	(0.04)	(0.09)	(0.23)
Higher vocational education	0.06	0.07	0.08	0.07	0.40
	(0.04)	(0.05)	(0.06)	(0.15)	(0.38)
Academic education	-0.01	-0.00	0.00	-0.04	-0.01
	(0.05)	(0.06)	(0.08)	(0.18)	(0.48)
Age 41-64	0.01	-0.01	-0.02	-0.02	-0.01
	(0.04)	(0.05)	(0.07)	(0.16)	(0.41)
Age 65+	0.07	0.09	0.10	0.13	0.46
	(0.05)	(0.06)	(0.08)	(0.18)	(0.47)
Log financial assets	0.06***	0.08***	0.09***	0.12***	0.20
	(0.01)	(0.01)	(0.02)	(0.05)	(0.12)
Constant	-0.49***	-0.60***	-0.65***	-0.90**	-1.43
	(0.12)	(0.14)	(0.19)	(0.44)	(1.16)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 925; standard errors are in parentheses. 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.

financial knowledge index, self-rated financial knowledge, the willingness to take financial risk, and gender. None of the extra financial literacy variables has an effect for the selected sample. This is also true for their interactions with financial advice in an unreported regression. In the entire sample (Table C.5), the higher financial knowledge among participants is reflected in small and mostly significant effects across the distribution. However, 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 near doubling of the adjusted R^2 in OLS the specification of Table C.4 compared to Table C.2 (leaving out the other extra variables does not seriously change the adjusted R^2). None of the main results regarding financial literacy and advice is altered by adding the covariates.

The latter is also true for the specification in Tables C.6 and C.7, which adds only the question on diversification from the financial knowledge battery to the variables used in Tables C.2 and C.3. I expect this to have the strongest relation with my outcome; the notes on endogeneity from the previous paragraph apply all the more strongly. As expected, all coefficients are negative, they are significant for the OLS regression and for the upper quantiles. The picture is different for the general population, where all coefficients are small and insignificant. This is consistent with non-participation as a rational decision for many households: Little knowledge is associated with worse outcomes conditional on participation; uninformed households staying out leads to similar return losses across the distribution.

My main results also continue to hold when adding household size, urbanisation, income, and additional wealth measures to the preferred specification in Tables C.8 and C.9 (although the coefficients in the latter table become insignificant). Following the suggestion of referees, I also added specifications that include house ownership and income (Table C.10, no changes whatsoever) and experience with risky assets. The latter is a little bit tricky because the attrition rate in the CentERpanel is relatively high. I thus include two specifications: The number of years I observe a household holding risky assets in the CentERpanel (Table C.11) and that number as a fraction of all years I observe a household (Table C.12). None of my results change.

Finally, I consider a number of different measures of diversification loss as dependent variables. Most results are broadly similar to those from Table C.2. In particular, reliance on one's own financial judgment clearly leads to more uncompensated risk-taking, regardless of the diversification measure. When using the share of idiosyncratic risk as the dependent variable (Table C.13), the effect of financial literacy appears to be present for all households at the 90^{th} percentile. At the 70^{th} percentile, the coefficients are the same as before, although only borderline significant. Tables C.14 and C.15 contain the results for the relative Sharpe ratio loss and the return loss measured as the fraction of risky financial assets (i.e. setting $w_h = 1$). 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 with annual portfolio standard deviations below 2% from the analysis. The point estimates display the same patterns as for the return loss; standard errors are much larger and statistical significance is often compromised.

Table C.4
Contributors to return loss, adding financial knowledge and risk tolerance

	OLS	$\mathbf{p_{10}}$	$\mathbf{p_{30}}$	$\mathbf{p_{50}}$	$\mathbf{p_{70}}$	$\mathbf{p_{90}}$
External advice × fin. numeracy	0.02	-0.01	0.04	0.06	0.11	0.13
	(0.07)	(0.01)	(0.03)	(0.07)	(0.11)	(0.48)
Financial knowledge index	0.06	0.01	-0.01	-0.00	-0.06	-0.03
	(0.16)	(0.01)	(0.03)	(0.05)	(0.09)	(0.40)
High self-rated fin. knowledge	-0.05	-0.03*	-0.04	-0.03	0.03	-0.14
	(0.14)	(0.02)	(0.04)	(0.08)	(0.13)	(0.58)
Financial advice: Own judgment	0.54**	-0.01	-0.01	0.10	0.69***	1.32**
	(0.22)	(0.02)	(0.04)	(0.08)	(0.13)	(0.57)
Own fin. judgment × fin. numeracy	-0.95**	-0.01	-0.03	-0.11	-1.08***	-1.67***
	(0.46)	(0.02)	(0.03)	(0.07)	(0.12)	(0.53)
Higher vocational education	0.02	-0.00	-0.00	0.01	-0.08	-0.54
	(0.14)	(0.02)	(0.04)	(0.08)	(0.14)	(0.58)
Academic education	-0.03	-0.01	0.01	-0.02	0.01	-0.37
	(0.14)	(0.02)	(0.04)	(0.09)	(0.15)	(0.66)
Age 41-64	-0.13	-0.01	0.06	0.12	0.00	-0.44
	(0.21)	(0.02)	(0.04)	(0.10)	(0.16)	(0.70)
Age 65+	0.05	-0.00	0.17***	0.27**	0.29	-0.13
	(0.23)	(0.02)	(0.05)	(0.11)	(0.18)	(0.80)
Log financial assets	-0.03	-0.00	-0.02	-0.03	-0.04	-0.03
	(0.07)	(0.01)	(0.01)	(0.03)	(0.05)	(0.23)
High tolerance for risky investm.	0.12*	0.02*	0.05***	0.08**	0.13**	0.17
	(0.06)	(0.01)	(0.02)	(0.04)	(0.06)	(0.27)
Female	0.11	-0.01	-0.02	0.06	0.16	0.24
	(0.15)	(0.02)	(0.04)	(0.09)	(0.15)	(0.66)
Constant	0.74	0.07	0.29*	0.45	0.77	1.60
	(0.69)	(0.07)	(0.16)	(0.33)	(0.57)	(2.45)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 239. 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.18.

Table C.5
Contributors to return loss, adding financial knowledge and risk tolerance. Non-participants are included in the estimation sample

	p_{75}	p_{80}	p_{85}	p_{90}	p_{95}
External advice × fin. numeracy	-0.00	0.00	0.02	0.01	0.04
	(0.03)	(0.03)	(0.05)	(0.15)	(0.22)
Financial knowledge index	0.04	0.05*	0.06	0.06	0.17
	(0.02)	(0.03)	(0.05)	(0.12)	(0.18)
High self-rated fin. knowledge	0.01	-0.01	-0.02	-0.01	-0.06
	(0.04)	(0.05)	(0.08)	(0.23)	(0.34)
Financial advice: Own judgment	0.01	0.04	0.05	0.35*	1.04***
	(0.04)	(0.05)	(0.08)	(0.21)	(0.30)
Own fin. judgment × fin. numeracy	0.02	0.04	-0.02	-0.45***	-1.75***
	(0.03)	(0.04)	(0.06)	(0.16)	(0.24)
Higher vocational education	0.03	0.01	-0.03	-0.04	-0.19
	(0.05)	(0.05)	(0.09)	(0.24)	(0.35)
Academic education	-0.03	-0.05	-0.10	-0.11	-0.26
	(0.06)	(0.06)	(0.11)	(0.29)	(0.42)
Age 41-64	-0.00	0.00	-0.02	-0.00	-0.14
	(0.05)	(0.06)	(0.10)	(0.26)	(0.38)
Age 65+	0.08	0.10	0.11	0.18	0.11
	(0.06)	(0.07)	(0.11)	(0.30)	(0.43)
Log financial assets	0.05***	0.06***	0.08***	0.08	0.07
	(0.02)	(0.02)	(0.03)	(0.08)	(0.11)
High tolerance for risky investm.	0.06***	0.10***	0.13***	0.17	0.24
	(0.02)	(0.02)	(0.04)	(0.11)	(0.15)
Female	0.01	0.04	0.07	0.11	0.03
	(0.04)	(0.05)	(0.09)	(0.23)	(0.34)
Constant	-0.34**	-0.38**	-0.45	-0.44	0.14
	(0.15)	(0.17)	(0.28)	(0.76)	(1.10)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 754; standard errors are in parentheses. 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 C.6
Contributors to return loss, adding diversification questions

	OLS	**p10**	**p30**	**p50**	**p70**	**p90**
External advice × fin. numeracy	0.08	-0.01	0.04	0.08	0.11	0.21
	(0.06)	(0.02)	(0.03)	(0.06)	(0.14)	(0.44)
Fin. knowl. - diversification, direct question	-0.55**	-0.01	-0.09*	-0.16*	-0.17	-1.43**
	(0.24)	(0.03)	(0.05)	(0.09)	(0.23)	(0.71)
Financial advice: Own judgment	0.52***	0.00	0.05	0.05	0.40**	1.13**
	(0.19)	(0.02)	(0.03)	(0.06)	(0.15)	(0.47)
Own fin. judgment × fin. numeracy	-0.64*	-0.00	-0.02	0.02	-0.43***	-1.13***
	(0.38)	(0.02)	(0.03)	(0.05)	(0.13)	(0.39)
Higher vocational education	0.13	0.00	0.02	0.03	0.06	0.23
	(0.15)	(0.02)	(0.04)	(0.07)	(0.16)	(0.50)
Academic education	0.06	-0.01	-0.03	-0.05	0.08	0.14
	(0.15)	(0.03)	(0.04)	(0.08)	(0.19)	(0.57)
Age 41-64	-0.14	0.01	0.02	0.11	0.09	-0.02
	(0.20)	(0.03)	(0.05)	(0.08)	(0.20)	(0.62)
Age 65+	-0.06	0.04	0.13**	0.25***	0.26	0.26
	(0.23)	(0.03)	(0.05)	(0.09)	(0.22)	(0.68)
Log financial assets	0.00	0.00	-0.01	-0.03	-0.04	-0.10
	(0.07)	(0.01)	(0.01)	(0.03)	(0.07)	(0.20)
Constant	0.86	0.00	0.24	0.58**	0.81	3.05
	(0.75)	(0.09)	(0.16)	(0.28)	(0.69)	(2.12)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 274. 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.13.

Table C.7
Contributors to return loss, adding diversification questions. Non-participants are included in the estimation sample

	p_{75}	p_{80}	p_{85}	p_{90}	p_{95}
External advice × fin. numeracy	-0.01	-0.00	-0.00	-0.01	-0.00
	(0.02)	(0.03)	(0.04)	(0.09)	(0.25)
Fin. knowl. - diversification, direct question	0.05	0.05	0.07	0.04	-0.01
	(0.04)	(0.05)	(0.07)	(0.17)	(0.46)
Financial advice: Own judgment	0.03	0.06	0.09*	0.19	0.63*
	(0.03)	(0.04)	(0.05)	(0.13)	(0.34)
Own fin. judgment × fin. numeracy	0.02	0.03	0.04	0.06	-0.60**
	(0.02)	(0.03)	(0.04)	(0.09)	(0.24)
Higher vocational education	0.06	0.07	0.07	0.07	0.40
	(0.04)	(0.05)	(0.06)	(0.15)	(0.40)
Academic education	-0.01	-0.01	0.00	-0.04	-0.01
	(0.05)	(0.06)	(0.08)	(0.19)	(0.51)
Age 41-64	0.00	0.00	0.00	-0.02	-0.01
	(0.04)	(0.05)	(0.06)	(0.16)	(0.43)
Age 65+	0.08*	0.13**	0.16**	0.11	0.46
	(0.05)	(0.06)	(0.07)	(0.18)	(0.49)
Log financial assets	0.06***	0.07***	0.08***	0.12***	0.20
	(0.01)	(0.01)	(0.02)	(0.05)	(0.13)
Constant	-0.49***	-0.59***	-0.69***	-0.93**	-1.42
	(0.12)	(0.14)	(0.18)	(0.45)	(1.22)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 925; standard errors are in parentheses. 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 C.8
Contributors to return loss, adding household size, urbanisation, and financial variables to the preferred specification

	OLS	p_{10}	p_{30}	p_{50}	p_{70}	p_{90}
External advice × fin. numeracy	0.06	-0.01	0.02	0.04	0.10	0.21
	(0.06)	(0.02)	(0.03)	(0.06)	(0.15)	(0.49)
Financial advice: Own judgment	0.49**	0.00	0.02	-0.01	0.31**	1.41***
	(0.20)	(0.02)	(0.03)	(0.06)	(0.16)	(0.53)
Own fin. judgment × fin. numeracy	-0.67*	-0.00	-0.01	0.01	-0.39***	-1.59***
	(0.39)	(0.01)	(0.03)	(0.05)	(0.13)	(0.44)
Higher vocational education	0.14	0.02	0.01	0.03	0.07	-0.18
	(0.14)	(0.02)	(0.04)	(0.07)	(0.17)	(0.56)
Academic education	0.03	-0.01	-0.06	-0.06	0.06	-0.13
	(0.15)	(0.02)	(0.04)	(0.08)	(0.19)	(0.66)
Age 41-64	-0.09	0.00	0.07	0.05	0.15	-0.18
	(0.22)	(0.02)	(0.05)	(0.08)	(0.21)	(0.70)
Age 65+	-0.10	0.01	0.16***	0.15	0.26	-0.02
	(0.24)	(0.03)	(0.05)	(0.10)	(0.24)	(0.83)
Female	-0.05	-0.01	-0.06	-0.03	-0.08	-0.01
	(0.17)	(0.02)	(0.04)	(0.07)	(0.18)	(0.61)
Household size	-0.03	-0.01	-0.01	-0.03	-0.00	0.04
	(0.05)	(0.01)	(0.01)	(0.02)	(0.06)	(0.21)
Degree of urbanisation	0.02	0.01	0.02	0.01	0.02	0.02
	(0.08)	(0.01)	(0.02)	(0.03)	(0.08)	(0.26)
Log net household income	-0.27	-0.04**	-0.01	0.04	-0.05	-0.23
	(0.21)	(0.02)	(0.04)	(0.07)	(0.18)	(0.63)
Log financial assets	0.02	0.01	-0.02	-0.05*	-0.04	0.01
	(0.08)	(0.01)	(0.02)	(0.03)	(0.07)	(0.24)
Log non-financial assets	-0.00	0.01**	-0.00	0.02	-0.00	-0.09
	(0.04)	(0.00)	(0.01)	(0.02)	(0.04)	(0.14)
Log debt	0.00	-0.00	0.00	-0.01	-0.00	-0.01
	(0.02)	(0.00)	(0.01)	(0.01)	(0.02)	(0.08)
Constant	3.21*	0.32*	0.42	0.31	1.30	4.32
	(1.77)	(0.19)	(0.39)	(0.70)	(1.76)	(5.99)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 273. 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.11.

Table C.9
Contributors to return loss, adding household size, urbanisation, and financial variables to the preferred specification. Non-participants are included in the estimation sample

	p_{75}	p_{80}	p_{85}	p_{90}	p_{95}
External advice × fin. numeracy	0.00	0.00	-0.00	0.02	0.03
	(0.02)	(0.03)	(0.04)	(0.10)	(0.28)
Financial advice: Own judgment	0.04	0.08*	0.09*	0.13	0.57
	(0.03)	(0.04)	(0.05)	(0.13)	(0.38)
Own fin. judgment × fin. numeracy	0.02	0.03	0.04	0.03	-0.46*
	(0.02)	(0.03)	(0.04)	(0.09)	(0.27)
Higher vocational education	0.06	0.06	0.08	0.09	0.23
	(0.04)	(0.05)	(0.06)	(0.15)	(0.44)
Academic education	-0.03	-0.02	0.01	-0.05	-0.03
	(0.05)	(0.06)	(0.08)	(0.19)	(0.56)
Age 41-64	-0.01	-0.01	-0.01	0.02	-0.09
	(0.04)	(0.05)	(0.07)	(0.17)	(0.48)
Age 65+	0.05	0.10	0.10	0.17	0.39
	(0.05)	(0.06)	(0.08)	(0.20)	(0.59)
Female	-0.03	-0.04	-0.05	-0.09	-0.12
	(0.04)	(0.04)	(0.06)	(0.14)	(0.40)
Household size	-0.00	-0.01	0.00	0.01	-0.01
	(0.01)	(0.02)	(0.02)	(0.06)	(0.16)
Degree of urbanisation	0.01	0.02	0.01	0.02	-0.04
	(0.02)	(0.02)	(0.03)	(0.07)	(0.20)
Log net household income	0.02	0.01	0.02	-0.09	-0.14
	(0.04)	(0.05)	(0.07)	(0.16)	(0.47)
Log financial assets	0.06***	0.07***	0.09***	0.14***	0.21
	(0.01)	(0.02)	(0.02)	(0.05)	(0.15)
Log non-financial assets	-0.00	-0.00	-0.01	-0.01	-0.03
	(0.01)	(0.01)	(0.01)	(0.03)	(0.09)
Log debt	0.00	0.00	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.02)	(0.06)
Constant	-0.57	-0.60	-0.78	-0.01	0.37
	(0.38)	(0.45)	(0.61)	(1.49)	(4.33)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 908; standard errors are in parentheses. 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 C.10
Contributors to return loss, including house ownership and household income

	OLS	$\mathbf{p_{10}}$	$\mathbf{p_{30}}$	$\mathbf{p_{50}}$	$\mathbf{p_{70}}$	$\mathbf{p_{90}}$
External advice × fin. numeracy	0.08	-0.01	0.04	0.06	0.09	0.15
	(0.05)	(0.02)	(0.03)	(0.06)	(0.13)	(0.41)
Financial advice: Own judgment	0.49***	-0.01	0.04	0.02	0.33**	1.14**
	(0.19)	(0.02)	(0.03)	(0.07)	(0.15)	(0.45)
Own fin. judgment × fin. numeracy	-0.66*	-0.01	-0.03	0.02	-0.39***	-1.59***
	(0.38)	(0.01)	(0.03)	(0.06)	(0.12)	(0.37)
Higher vocational education	0.08	0.02	0.04	0.01	0.05	0.08
	(0.13)	(0.02)	(0.04)	(0.07)	(0.16)	(0.48)
Academic education	0.09	-0.02	-0.02	-0.06	0.13	0.22
	(0.14)	(0.02)	(0.04)	(0.08)	(0.18)	(0.54)
Age 41-64	-0.13	0.02	0.03	0.14	0.15	-0.20
	(0.23)	(0.02)	(0.04)	(0.09)	(0.20)	(0.59)
Age 65+	-0.02	0.04	0.13***	0.26***	0.32	0.01
	(0.23)	(0.03)	(0.05)	(0.10)	(0.21)	(0.65)
Log financial assets	-0.05	-0.01	-0.03*	-0.06*	-0.06	-0.11
	(0.07)	(0.01)	(0.02)	(0.03)	(0.07)	(0.20)
Log net household income	-0.17	-0.00	0.00	0.00	-0.03	-0.10
	(0.18)	(0.02)	(0.04)	(0.08)	(0.17)	(0.51)
Has primary housing (household)	0.13	-0.00	-0.01	0.00	-0.00	-0.13
	(0.14)	(0.02)	(0.04)	(0.07)	(0.16)	(0.48)
Constant	2.71*	0.20	0.36	0.76	1.18	3.19
	(1.55)	(0.20)	(0.38)	(0.77)	(1.67)	(5.03)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 266. 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.14.

Table C.11
Contributors to return loss, including years of experience with risky assets

	OLS	$\mathbf{p_{10}}$	$\mathbf{p_{30}}$	$\mathbf{p_{50}}$	$\mathbf{p_{70}}$	$\mathbf{p_{90}}$
External advice × fin. numeracy	0.04	-0.01	0.03	0.05	0.09	0.12
	(0.05)	(0.02)	(0.03)	(0.06)	(0.13)	(0.50)
Financial advice: Own judgment	0.49**	-0.01	0.02	0.02	0.32**	1.55***
	(0.19)	(0.02)	(0.04)	(0.07)	(0.14)	(0.55)
Own fin. judgment × fin. numeracy	-0.68*	-0.00	0.00	0.01	-0.40***	-1.44***
	(0.38)	(0.01)	(0.03)	(0.06)	(0.12)	(0.45)
Higher vocational education	0.12	0.02	0.02	0.05	0.07	0.01
	(0.15)	(0.02)	(0.04)	(0.07)	(0.15)	(0.58)
Academic education	0.02	-0.01	-0.03	-0.05	0.08	-0.07
	(0.15)	(0.02)	(0.05)	(0.08)	(0.17)	(0.67)
Age 41-64	-0.14	0.01	0.02	0.11	0.11	-0.58
	(0.21)	(0.02)	(0.05)	(0.09)	(0.19)	(0.73)
Age 65+	-0.06	0.03	0.13**	0.22**	0.30	-0.31
	(0.25)	(0.03)	(0.06)	(0.10)	(0.21)	(0.83)
Log financial assets	-0.02	-0.00	-0.02	-0.03	-0.05	-0.09
	(0.07)	(0.01)	(0.02)	(0.03)	(0.06)	(0.24)
Years of experience with risky assets	0.01	0.00	0.00	0.00	0.00	0.02
	(0.02)	(0.00)	(0.01)	(0.01)	(0.02)	(0.08)
Constant	0.56	0.05	0.26	0.49	0.78	2.00
	(0.74)	(0.08)	(0.17)	(0.30)	(0.63)	(2.44)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 274. 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.11.

Table C.12
Contributors to return loss, including fraction of years in panel with experience with risky assets

	OLS	$\mathbf{p_{10}}$	$\mathbf{p_{30}}$	$\mathbf{p_{50}}$	$\mathbf{p_{70}}$	$\mathbf{p_{90}}$
External advice × fin. numeracy	0.04	-0.01	0.02	0.05	0.08	0.17
	(0.05)	(0.01)	(0.03)	(0.06)	(0.13)	(0.49)
Financial advice: Own judgment	0.49**	0.00	0.02	0.03	0.32**	1.52***
	(0.19)	(0.02)	(0.04)	(0.07)	(0.15)	(0.54)
Own fin. judgment × fin. numeracy	-0.68*	-0.00	-0.00	0.01	-0.40***	-1.40***
	(0.38)	(0.01)	(0.03)	(0.06)	(0.12)	(0.45)
Higher vocational education	0.11	-0.00	0.02	0.03	0.06	0.17
	(0.15)	(0.02)	(0.04)	(0.07)	(0.16)	(0.58)
Academic education	0.01	-0.02	-0.02	-0.05	0.09	-0.01
	(0.15)	(0.02)	(0.04)	(0.08)	(0.18)	(0.66)
Age 41-64	-0.13	0.01	0.04	0.12	0.11	-0.23
	(0.21)	(0.02)	(0.05)	(0.09)	(0.19)	(0.71)
Age 65+	-0.05	0.04	0.16***	0.25**	0.30	0.05
	(0.24)	(0.02)	(0.05)	(0.10)	(0.21)	(0.79)
Log financial assets	-0.02	-0.00	-0.01	-0.03	-0.05	-0.15
	(0.08)	(0.01)	(0.02)	(0.03)	(0.06)	(0.24)
Fraction of years of experience with risky assets	0.16	0.05	0.08	0.06	-0.02	0.68
	(0.32)	(0.03)	(0.08)	(0.15)	(0.33)	(1.23)
Constant	0.52	0.02	0.13	0.42	0.83	1.88
	(0.69)	(0.07)	(0.17)	(0.31)	(0.67)	(2.47)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 274. 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.11.

Table C.13
Contributors to idiosyncratic portfolio risk

	OLS	**p10**	**p30**	**p50**	**p70**	**p90**
External advice × fin. numeracy	-1.14	-0.15	-0.79	-0.43	0.79	-7.06
	(1.93)	(0.40)	(0.62)	(0.82)	(2.06)	(4.41)
Financial advice: Own judgment	3.59**	0.71	0.68	2.66***	6.62***	7.43
	(1.60)	(0.44)	(0.68)	(0.89)	(2.26)	(4.84)
Own fin. judgment × fin. numeracy	-1.20	-0.70*	-0.37	-1.75**	-3.38*	-7.38*
	(1.90)	(0.37)	(0.56)	(0.74)	(1.87)	(4.01)
Higher vocational education	-0.18	0.71	1.15	0.46	-0.92	0.40
	(1.56)	(0.47)	(0.72)	(0.95)	(2.41)	(5.16)
Academic education	-2.51	-0.16	-0.52	-0.54	-2.46	-7.49
	(1.53)	(0.54)	(0.82)	(1.09)	(2.75)	(5.89)
Age 41-64	-1.14	-0.29	1.27	1.49	-3.52	-0.07
	(1.99)	(0.58)	(0.89)	(1.17)	(2.96)	(6.34)
Age 65+	-1.99	-0.22	0.99	0.50	-4.42	-7.81
	(2.11)	(0.64)	(0.98)	(1.29)	(3.27)	(7.01)
Log financial assets	-1.04	0.00	0.03	-0.21	-1.90*	-1.76
	(0.68)	(0.19)	(0.29)	(0.38)	(0.97)	(2.08)
Constant	23.72***	3.96**	4.81	9.03**	35.82***	47.69**
	(7.55)	(1.96)	(3.01)	(3.97)	(10.03)	(21.50)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the idiosyncratic portfolio risk; see Section B in this Online Appendix for details. The number of observations is 274. 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.05.

Table C.14
Contributors to the relative Sharpe ratio loss

	OLS	**p10**	**p30**	**p50**	**p70**	**p90**
External advice × fin. numeracy	2.94	-1.03	0.60	1.58	2.92	9.40
	(2.12)	(1.34)	(1.35)	(2.71)	(4.20)	(10.66)
Financial advice: Own judgment	4.15	0.37	0.50	2.48	8.82*	-5.32
	(3.27)	(1.43)	(1.44)	(2.89)	(4.48)	(11.38)
Own fin. judgment × fin. numeracy	-3.90	-0.14	0.69	1.03	-4.27	-7.57
	(3.03)	(1.13)	(1.14)	(2.28)	(3.54)	(8.98)
Higher vocational education	-1.19	-0.70	-0.87	0.48	-2.19	2.25
	(3.08)	(1.48)	(1.49)	(2.98)	(4.62)	(11.72)
Academic education	-1.24	0.41	0.30	3.43	0.49	-5.09
	(4.47)	(1.77)	(1.78)	(3.56)	(5.52)	(14.02)
Age 41-64	1.14	0.90	-1.05	-4.80	-5.69	2.46
	(3.96)	(1.99)	(2.00)	(4.01)	(6.21)	(15.77)
Age 65+	1.12	1.33	-0.41	-2.06	-3.70	-6.77
	(4.55)	(2.17)	(2.18)	(4.38)	(6.79)	(17.24)
Log financial assets	-1.57	-0.28	-1.66***	-3.37***	-3.39*	1.52
	(1.40)	(0.57)	(0.58)	(1.16)	(1.79)	(4.55)
Constant	37.71***	10.26*	30.20***	53.86***	63.66***	37.68
	(14.07)	(5.91)	(5.95)	(11.92)	(18.49)	(46.94)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the relative Sharpe ratio loss; see Section B in this Online Appendix for details. The number of observations is 197 (households with an annual standard deviation of the financial portfolio below 2% have been excluded). 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.01.

Table C.15
Contributors to return loss as a fraction of risky financial assets

	OLS	**p10**	**p30**	**p50**	**p70**	**p90**
External advice × fin. numeracy	0.13	-0.10	-0.05	-0.00	0.27	0.97
	(0.13)	(0.06)	(0.07)	(0.26)	(0.52)	(1.29)
Financial advice: Own judgment	1.22***	-0.02	0.13	0.27	1.18**	1.89
	(0.34)	(0.06)	(0.08)	(0.28)	(0.56)	(1.37)
Own fin. judgment × fin. numeracy	-0.65	-0.00	0.06	0.04	-1.08**	-0.73
	(0.51)	(0.05)	(0.06)	(0.22)	(0.44)	(1.08)
Higher vocational education	0.18	-0.03	0.04	0.08	0.08	1.99
	(0.39)	(0.07)	(0.08)	(0.29)	(0.57)	(1.41)
Academic education	-0.29	-0.06	-0.05	-0.05	0.17	0.41
	(0.40)	(0.08)	(0.10)	(0.34)	(0.69)	(1.69)
Age 41-64	-0.50	0.10	0.14	0.01	-0.82	-2.41
	(0.53)	(0.09)	(0.11)	(0.39)	(0.77)	(1.90)
Age 65+	-0.90*	0.08	0.20	0.08	-0.81	-3.48*
	(0.54)	(0.10)	(0.12)	(0.42)	(0.84)	(2.08)
Log financial assets	-0.13	-0.01	-0.08**	-0.09	-0.40*	-0.33
	(0.17)	(0.03)	(0.03)	(0.11)	(0.22)	(0.55)
Constant	3.25*	0.50*	1.27***	1.64	6.48***	9.07
	(1.76)	(0.26)	(0.33)	(1.15)	(2.30)	(5.66)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of risky financial assets; see Section B in this Online Appendix for details. The number of observations is 197 (households with an annual standard deviation of the financial portfolio below 2% have been excluded). 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.06.

C.3. Additional tables for Section II.B

Table C.16
Contributors to fees paid

	Plain	Interacted	Fin. Knowledge
Financial numeracy index	0.03	0.03	0.01
	(0.06)	(0.07)	(0.06)
Financial advice: Professionals	0.29**	0.29***	0.29**
	(0.12)	(0.11)	(0.12)
Financial advice: Family/friends	0.08	0.06	0.09
	(0.10)	(0.11)	(0.10)
Higher vocational education	-0.03	-0.03	-0.04
	(0.10)	(0.10)	(0.10)
Academic education	0.06	0.06	0.05
	(0.12)	(0.12)	(0.12)
Age 41-64	0.13	0.13	0.14
	(0.11)	(0.12)	(0.11)
Age 65+	0.01	0.01	0.01
	(0.12)	(0.12)	(0.12)
Log financial assets	-0.00	-0.00	-0.01
	(0.04)	(0.04)	(0.04)
Prof. advice × fin. numeracy		-0.02	
		(0.12)	
Advice fam./friends × fin. numeracy		0.07	
		(0.18)	
Financial knowledge index			0.05
			(0.07)
Constant	0.70	0.69	0.73*
	(0.42)	(0.43)	(0.42)
Adjusted R^2	0.018	0.014	0.018

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the annual fee (in percentage points) paid on the risky assets portfolio. The regressions have been estimated on both waves of data and standard errors are clustered at the household level and use sampling weights. The number of observations is 274; standard errors are in parentheses.

Table C.17
Average return loss of risky asset holders by financial numeracy and the primary source of financial advice, mutual fund fees are added to the return loss ex post

Primary source of financial advice	Fin. numeracy $<$ median		Fin. numeracy $\geq$ median	
	Mean	(S.E.)	Mean	(S.E.)
Self-decider	1.34	(0.25)	0.83	(0.08)
Professional adviser	0.79	(0.24)	0.77	(0.06)
Family / friends	0.72	(0.29)	0.80	(0.17)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The number of observations is 274. The return loss measures the expected return per year (in percentage points) that is foregone by not investing in the efficient portfolio with the same standard deviation of a household's portfolio, see Section I.C for its definition.

Table C.18
Contributors to return loss, with fees added to the return loss ex post. Non-participants are included in the estimation sample

	p_{75}	p_{80}	p_{85}	p_{90}	p_{95}
Prof. advice × fin. numeracy	0.20	0.21	0.23	0.21	0.33
	(0.19)	(0.20)	(0.26)	(0.54)	(1.10)
Financial advice: Family/friends	-0.04	-0.08	-0.18	-0.13	0.08
	(0.30)	(0.31)	(0.41)	(0.84)	(1.71)
Advice fam./friends × fin. numeracy	0.54	0.59	0.54	0.73	1.05
	(0.36)	(0.37)	(0.49)	(1.01)	(2.06)
Financial advice: Own judgment	0.46**	0.74***	0.92***	1.10*	2.41*
	(0.22)	(0.22)	(0.30)	(0.61)	(1.25)
Own fin. judgment × fin. numeracy	-0.99***	-1.28***	-1.48***	-1.52***	-2.63***
	(0.16)	(0.16)	(0.21)	(0.43)	(0.88)
Higher vocational education	0.14	0.12	0.09	-0.20	0.27
	(0.20)	(0.20)	(0.27)	(0.56)	(1.14)
Academic education	0.32	0.39*	0.39	-0.08	-0.53
	(0.23)	(0.23)	(0.31)	(0.64)	(1.30)
Age 41-64	0.32	0.17	0.15	-0.37	-0.35
	(0.25)	(0.25)	(0.34)	(0.69)	(1.40)
Age 65+	0.58**	0.56**	0.36	-0.11	0.04
	(0.27)	(0.28)	(0.37)	(0.76)	(1.55)
Log financial assets	-0.08	-0.09	-0.07	-0.01	0.13
	(0.08)	(0.08)	(0.11)	(0.23)	(0.46)
Constant	1.55*	1.83**	1.83	1.94	0.79
	(0.87)	(0.88)	(1.18)	(2.41)	(4.91)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets (adding mutual fund fees ex post to the return loss); see Sections I.C and II.B for details. The number of observations is 274; standard errors are in parentheses. 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 C.19
Contributors to return loss, ignoring mutual fund fees

	OLS	$\mathbf{p_{10}}$	$\mathbf{p_{30}}$	$\mathbf{p_{50}}$	$\mathbf{p_{70}}$	$\mathbf{p_{90}}$
Prof. advice × fin. numeracy	0.01	-0.01	0.01	0.05	0.08	0.11
	(0.05)	(0.02)	(0.04)	(0.07)	(0.16)	(0.45)
Financial advice: Family/friends	0.11	0.01	-0.04	-0.01	-0.02	0.32
	(0.16)	(0.03)	(0.06)	(0.11)	(0.25)	(0.70)
Advice fam./friends × fin. numeracy	0.15	0.05	0.06	0.15	0.22	0.58
	(0.14)	(0.04)	(0.07)	(0.13)	(0.31)	(0.84)
Financial advice: Own judgment	0.56***	0.00	0.02	0.02	0.35*	1.53***
	(0.20)	(0.03)	(0.04)	(0.08)	(0.19)	(0.51)
Own fin. judgment × fin. numeracy	-0.71*	-0.00	-0.02	0.03	-0.41***	-1.56***
	(0.40)	(0.02)	(0.03)	(0.06)	(0.13)	(0.36)
Higher vocational education	0.12	0.01	0.02	0.03	0.05	0.02
	(0.16)	(0.02)	(0.04)	(0.07)	(0.17)	(0.47)
Academic education	0.01	-0.01	-0.05	-0.05	0.10	0.00
	(0.16)	(0.03)	(0.04)	(0.08)	(0.19)	(0.53)
Age 41-64	-0.13	0.00	0.01	0.12	0.10	-0.12
	(0.22)	(0.03)	(0.05)	(0.09)	(0.21)	(0.57)
Age 65+	-0.03	0.04	0.16***	0.28***	0.33	0.32
	(0.25)	(0.03)	(0.05)	(0.10)	(0.23)	(0.63)
Log financial assets	-0.01	0.00	-0.01	-0.03	-0.05	-0.03
	(0.08)	(0.01)	(0.02)	(0.03)	(0.07)	(0.19)
Constant	0.52	-0.02	0.22	0.47	0.77	1.01
	(0.84)	(0.10)	(0.16)	(0.32)	(0.73)	(2.01)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 274. 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.11.

Table C.20
Contributors to return loss, ignoring mutual fund fees. Non-participants are included in the estimation sample

	p_{75}	p_{80}	p_{85}	p_{90}	p_{95}
Prof. advice × fin. numeracy	0.01	0.01	0.01	0.02	-0.01
	(0.04)	(0.04)	(0.06)	(0.13)	(0.34)
Financial advice: Family/friends	0.01	0.04	0.04	0.03	0.16
	(0.05)	(0.06)	(0.08)	(0.20)	(0.49)
Advice fam./friends × fin. numeracy	0.00	0.01	0.01	0.00	0.02
	(0.04)	(0.04)	(0.06)	(0.13)	(0.34)
Financial advice: Own judgment	0.06	0.09*	0.13*	0.23	0.76*
	(0.04)	(0.05)	(0.07)	(0.16)	(0.41)
Own fin. judgment × fin. numeracy	0.02	0.03	0.05	0.08	-0.63***
	(0.02)	(0.03)	(0.04)	(0.09)	(0.23)
Higher vocational education	0.07*	0.08*	0.09	0.07	0.46
	(0.04)	(0.05)	(0.07)	(0.16)	(0.39)
Academic education	-0.02	-0.00	0.02	-0.05	-0.01
	(0.05)	(0.06)	(0.08)	(0.19)	(0.49)
Age 41-64	0.00	-0.01	-0.02	-0.02	-0.07
	(0.04)	(0.05)	(0.07)	(0.17)	(0.42)
Age 65+	0.09*	0.11*	0.11	0.14	0.41
	(0.05)	(0.06)	(0.08)	(0.19)	(0.49)
Log financial assets	0.07***	0.08***	0.10***	0.13***	0.20
	(0.01)	(0.02)	(0.02)	(0.05)	(0.13)
Constant	-0.53***	-0.66***	-0.75***	-0.98**	-1.47
	(0.13)	(0.15)	(0.21)	(0.49)	(1.24)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the return loss as a percentage of total financial assets; see Section I.C for details. The number of observations is 925; standard errors are in parentheses. 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.

C.4. Detailed tables for Section II.C

Table C.21
Fraction rating their own financial knowledge as high; by financial literacy and the primary source of financial advice

Primary source of financial advice	Fin. numeracy < median		Fin. numeracy ≥ median	
	Mean	**(S.E.)**	**Mean**	**(S.E.)**
Self-decider	0.37	(0.07)	0.47	(0.04)
Professional adviser	0.00	(–)	0.31	(0.06)
Family / friends	0.33	(0.13)	0.19	(0.07)

Source: CentERpanel, own calculations. The number of observations is 274.

Table C.22
Average financial wealth by financial literacy and the primary source of financial advice

Primary source of financial advice	Fin. numeracy < median		Fin. numeracy ≥ median	
	Mean	**(S.E.)**	**Mean**	**(S.E.)**
Self-decider	50802	(9368)	109619	(12313)
Professional adviser	80211	(16199)	125449	(19871)
Family / friends	63513	(16772)	64023	(13121)

Source: CentERpanel, own calculations. The number of observations is 274.

Table C.23

Average return loss of risky asset holders by financial numeracy and the primary source of financial advice, mutual fund fees and loads (assuming 60% annual turnover) are added to the return loss ex post

Primary source of financial advice	Fin. numeracy < median		Fin. numeracy ≥ median	
	Mean	(S.E.)	Mean	(S.E.)
Self-decider	1.90	(0.24)	1.55	(0.10)
Professional adviser	1.66	(0.31)	1.48	(0.10)
Family / friends	1.60	(0.40)	1.40	(0.18)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The number of observations is 274. The return loss measures the expected return per year (in percentage points) that is foregone by not investing in the efficient portfolio with the same standard deviation of a household's portfolio, see Section I.C for its definition.

Table C.24

Average return loss of risky asset holders by financial numeracy and the primary source of financial advice, mutual fund fees and loads (assuming 60% annual turnover if advice is from professionals and 30% otherwise) are added to the return loss ex post

Primary source of financial advice	Fin. numeracy < median		Fin. numeracy ≥ median	
	Mean	(S.E.)	Mean	(S.E.)
Self-decider	1.62	(0.24)	1.19	(0.09)
Professional adviser	1.66	(0.31)	1.48	(0.10)
Family / friends	1.16	(0.32)	1.10	(0.17)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The number of observations is 274. The return loss measures the expected return per year (in percentage points) that is foregone by not investing in the efficient portfolio with the same standard deviation of a household's portfolio, see Section I.C for its definition.

D. The performance of proxies for portfolio diversification

Finding out whether one can uncover similar patterns to those of the main paper on the basis of simpler diversification measures is important for at least two reasons. First, few if any other surveys currently contain the same amount of detail on portfolios as the DHS; gauging the limits of their usefulness is of great interest. Second, the process of constructing the detailed portfolio characteristics is quite cumbersome; finding a measure that allows to avoid this procedure would be very welcome. I thus repeat the regressions for my preferred specification using various other diversification measures that have been considered by previous authors. All these measures only concern the risky portfolio, hence the return loss is no appropriate statistic for comparison. Instead, I use the Relative Sharpe Ratio Loss (2), see Section B.

Calvet, Campbell, and Sodini (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 D.1 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", financial numeracy is insignificant regardless of the main source of financial advice. The quantile regressions are meaningless except for the 7^{th} decile because the measure takes on the values zero or one for many portfolios. At the 7^{th} decile, 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, Campbell, and Sodini (2009b).[10]

A 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 $1 - \nu_h / \min\{1, N_{h,\text{shares}}\}$, where ν_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 D.2. Again, the OLS coefficient only picks up the higher diversification losses of those relying on their own judgment. However, the results at the 70^{th} percentile now show the same effects as those for the more general measures: Higher diversification losses for

[10] Calvet, Campbell, and Sodini (2009a) consider the share of funds and the Sharpe ratio, which leads to exactly the same figures as using the inverse share and the relative Sharpe ratio loss.

Table D.1
Contributors to the fraction in shares

	OLS	$\mathbf{p_{10}}$	$\mathbf{p_{30}}$	$\mathbf{p_{50}}$	$\mathbf{p_{70}}$	$\mathbf{p_{90}}$
External advice × fin. numeracy	-0.006	0.000	0.000	-0.000	0.067	0.000
	(0.063)	(.)	(.)	(0.086)	(0.126)	(0.017)
Financial advice: Own judgment	0.156***	0.000	0.000	0.147	0.576***	0.000
	(0.057)	(.)	(.)	(0.094)	(0.139)	(0.019)
Own fin. judgment × fin. numeracy	-0.000	0.000	0.000	0.049	-0.014	0.000
	(0.057)	(.)	(.)	(0.078)	(0.115)	(0.015)
Higher vocational education	0.020	0.000	0.000	0.010	0.008	0.000
	(0.062)	(.)	(.)	(0.100)	(0.148)	(0.020)
Academic education	-0.075	0.000	0.000	-0.004	-0.204	-0.046**
	(0.069)	(.)	(.)	(0.114)	(0.168)	(0.023)
Age 41-64	-0.102	0.000	0.000	-0.001	-0.081	0.000
	(0.084)	(.)	(.)	(0.123)	(0.182)	(0.024)
Age 65+	-0.062	0.000	0.000	-0.001	0.050	0.000
	(0.092)	(.)	(.)	(0.136)	(0.201)	(0.027)
Log financial assets	0.001	0.000	0.000	0.007	-0.017	0.000
	(0.027)	(.)	(.)	(0.040)	(0.059)	(0.008)
Constant	0.303	0.000	0.000	-0.067	0.524	1.000***
	(0.285)	(.)	(.)	(0.418)	(0.615)	(0.082)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the fraction of risky assets invested directly in shares; see Section D for details. The number of observations is 274. 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.03.

the "self-deciders", which get much worse with below-average financial literacy and almost disappear for those with the highest level of financial literacy.

Table D.2
Contributors to the inverse of the diversification index D_1 from Guiso and Jappelli (2009)

	OLS	$\mathbf{p_{10}}$	$\mathbf{p_{30}}$	$\mathbf{p_{50}}$	$\mathbf{p_{70}}$	$\mathbf{p_{90}}$
External advice × fin. numeracy	-0.032	0.000	0.000	-0.000	0.003	-0.199
	(0.071)	(.)	(.)	(0.027)	(0.073)	(0.161)
Financial advice: Own judgment	0.062	0.000	0.000	0.073***	0.171**	0.145
	(0.050)	(.)	(.)	(0.028)	(0.076)	(0.166)
Own fin. judgment × fin. numeracy	-0.016	0.000	0.000	-0.012	-0.126**	0.049
	(0.046)	(.)	(.)	(0.023)	(0.062)	(0.136)
Higher vocational education	0.018	0.000	0.000	-0.000	0.076	0.324*
	(0.049)	(.)	(.)	(0.030)	(0.080)	(0.177)
Academic education	-0.016	0.000	0.000	-0.000	0.041	0.073
	(0.054)	(.)	(.)	(0.034)	(0.091)	(0.200)
Age 41-64	-0.072	0.000	0.000	-0.000	-0.099	-0.183
	(0.066)	(.)	(.)	(0.037)	(0.098)	(0.215)
Age 65+	-0.036	0.000	0.000	-0.000	-0.124	-0.056
	(0.071)	(.)	(.)	(0.041)	(0.109)	(0.240)
Log financial assets	-0.040*	0.000	0.000	0.000	-0.036	-0.122*
	(0.023)	(.)	(.)	(0.012)	(0.032)	(0.071)
Constant	0.638**	0.000	0.000	-0.000	0.553*	1.920***
	(0.258)	(.)	(.)	(0.125)	(0.334)	(0.735)

Source: CentERpanel, Datastream, Euroinvestor, own calculations. The dependent variable is the inverse of the diversification index D_1 from Guiso and Jappelli (2009); see Section D for details. The number of observations is 270. 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 cross-section. All regressions use sampling weights; standard errors are in parentheses. The adjusted R^2 of the OLS regression is 0.03.

References

Agarwal, Sumit, John C. Driscoll, Xavier Gabaix, and David Laibson (Fall 2009). "The Age of Reason: Financial Decisions over the Life Cycle and Implications for Regulation". In: *Brookings Papers on Economic Activity*, pp. 51–117.

Barber, Brad M. and Terrance Odean (2000). "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors". In: *The Journal of Finance* 55.2, pp. 773–806.

— (Feb. 2001). "Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment". In: *Quarterly Journal of Economics* 116.1, pp. 261–292.

Bergstresser, Daniel and James Poterba (Aug. 2004). "Asset Allocation and Asset Location: Household Evidence from the Survey of Consumer Finances". In: *Journal of Public Economics* 88.9-10, pp. 1893–1915.

Bilias, Yannis, Dimitris Georgarakos, and Michael Haliassos (Oct. 2009). "Equity Culture and the Distribution of Wealth". Available at `http://ssrn.com/abstract=731805`.

— (2010). "Portfolio Inertia and Stock Market Fluctuations". In: *Journal of Money, Credit and Banking* 42.4, pp. 715–742.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini (2007). "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes". In: *Journal of Political Economy* 115.5, pp. 707–747.

— (Feb. 2009a). "Fight or Flight? Portfolio Rebalancing by Individual Investors". In: *Quarterly Journal of Economics* 124.1, pp. 301–348.

— (2009b). "Measuring the Financial Sophistication of Households". In: *American Economic Review* 99.2, pp. 393–398.

Campbell, John Y. (2006). "Household Finance". In: *Journal of Finance* 61, pp. 1553–1604.

Christelis, Dimitris, Dimitris Georgarakos, and Michael Haliassos (Sept. 2010). "Differences in Portfolios across Countries: Economic Environment versus Household Characteristics". Available at `http://ssrn.com/abstract=1089802`.

— (Aug. 2011). "Stockholding: Participation, Location, and Spillovers". In: *Journal of Banking & Finance* 35.8, pp. 1918–1930.

Christelis, Dimitris, Tullio Jappelli, and Mario Padula (2010). "Cognitive Abilities and Portfolio Choice". In: *European Economic Review* 54.1, pp. 18–38.

Croson, Rachel and Uri Gneezy (June 2009). "Gender Differences in Preferences". In: *Journal of Economic Literature* 47.2, pp. 448–474.

Goetzmann, William N and Alok Kumar (2008). "Equity Portfolio Diversification". In: *Review of Finance* 12.3, pp. 433–463.

Graham, John R., Campbell R. Harvey, and Hai Huang (July 2009). "Investor Competence, Trading Frequency, and Home Bias". In: *Management Science* 55.7, pp. 1094–1106.

Grinblatt, Mark, Matti Keloharju, and Juhani Linnainmaa (Dec. 2011). "IQ and Stock Market Participation". In: *Journal of Finance* 66.6, pp. 2121–2164.

Grinblatt, Mark, Matti Keloharju, and Juhani Linnainmaa (May 2012). "IQ, Trading Behavior, and Performance". In: *Journal of Financial Economics* 104.2, pp. 339–362.

Guiso, Luigi, Michael Haliassos, and Tullio Jappelli, eds. (2002). *Household Portfolios*. Cambridge, MA: MIT Press.

Guiso, Luigi and Tullio Jappelli (Oct. 2006). "Information Acquisition and Portfolio Performance". University of Naples, CSEF Working Paper No. 167.

— (Jan. 2009). "Financial Literacy and Portfolio Diversification". University of Naples, CSEF Working Paper no. 212.

Hackethal, Andreas, Michael Haliassos, and Tullio Jappelli (2012). "Financial Advisors: A Case of Babysitters?" In: *Journal of Banking & Finance* 36.2, pp. 509–524.

Hung, Angela A., Andrew M. Parker, and Joanne K. Yoong (Sept. 2009). "Defining and Measuring Financial Literacy". RAND Labor and Population Working Paper 708.

Inderst, Roman and Marco Ottaviani (June 2012). "Financial Advice". In: *Journal of Economic Literature* 50.2, pp. 494–512.

Ivković, Zoran, James Poterba, and Scott Weisbenner (Dec. 2005). "Tax-Motivated Trading by Individual Investors". In: *American Economic Review* 95.5, pp. 1605–1630.

Ivković, Zoran, Clemens Sialm, and Scott Weisbenner (2008). "Portfolio Concentration and the Performance of Individual Investors". In: *Journal of Financial and Quantitative Analysis* 43.03, pp. 613–655.

Juster, F. Thomas, James P. Smith, and Frank Stafford (1999). "The Measurement and Structure of Household Wealth". In: *Labour Economics* 6.2, pp. 253–275.

Kimball, Miles S. and Tyler Shumway (Jan. 2010). "Investor Sophistication and the Home Bias, Diversification, and Employer Stock Puzzles". Mimeo, University of Michigan, available at `http://ssrn.com/abstract=1572866`.

Korniotis, George M. and Alok Kumar (2011). "Do Older Investors Make Better Investment Decisions?" In: *Review of Economics and Statistics* 93.1, pp. 244–265.

Lusardi, Annamaria and Olivia S. Mitchell (Jan. 2007a). "Baby Boomer retirement security: The roles of planning, financial literacy, and housing wealth". In: *Journal of Monetary Economics* 54.1, pp. 205–224.

— (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.

Massa, Massimo and Andrei Simonov (2006). "Hedging, Familiarity and Portfolio Choice". In: *Review of Financial Studies* 19.2, pp. 633–685.

Odean, Terrance (1998). "Are Investors Reluctant to Realize Their Losses?" In: *The Journal of Finance* 53.5, pp. 1775–1798.

Polkovnichenko, Valery (2005). "Household Portfolio Diversification: A Case for Rank-Dependent Preferences". In: *Review of Financial Studies* 18.4, pp. 1467–1502.

Rooij, Maarten C. J. van, Annamaria Lusardi, and Rob Alessie (Aug. 2011). "Financial Literacy and Stock Market Participation". In: *Journal of Financial Economics* 101.2, pp. 449–472.

Tang, Ning, Olivia S. Mitchell, Gary R. Mottola, and Stephen P. Utkus (Dec. 2010). "The Efficiency of Sponsor and Participant Portfolio Choices in 401(K) Plans". In: *Journal of Public Economics* 94.11-12, pp. 1073–1085.