ISER Working Paper Series

# Interviewer Effects and the Measurement of Financial Literacy

Thomas F. Crossley
University of Essex and Institute for Fiscal Studies

Tobias Schmidt
Deutsche Bundesbank

Panagiota Tzamourani
Deutsche Bundesbank

Joachim K. Winter
University of Munich

No. 2017-06
May 2017

INSTITUTE FOR SOCIAL
& ECONOMIC RESEARCH

University of Essex

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

www.iser.essex.ac.uk

## Non-Technical Summary


- *Differences across survey respondents in a measure of financial literacy are explained in part by which interviewer conducted the interview.*

- *These 'interviewer effects' are larger for financial literacy than for other variables collected in the survey.*

- *It is possible to use information on which interviewer conducted the interview to improve estimates of the effect of financial literacy on financial choices and outcomes.*

- *Corrected estimates of the effect of financial literacy are larger than uncorrected estimates.*

Survey measures of financial literacy are based on responses to a series of questions that test respondents' financial literacy. Previous research suggests the resulting scores suffer from important measurement error. This study examined the role of the interview in both generating and moderating measurement error in a financial literacy score.

We used data from the first wave of the German Panel on Household Finances (PHF), conducted in 2010/2011. This survey contains a standard set of questions used to measure financial literacy. Unusually, we were also able to group respondents by interviewer. We studied 1,705 interviews conducted by 160 interviewers. We then used econometric techniques to gauge the role of the interviewer in the elicited responses. In particular we tested whether some interviewers elicited better responses.

Insufficient saving and poor financial decision-making are major policy concerns, particularly in the face of increasingly complex financial markets and increasing reliance on individual financial provision for old age. One explanation for inadequate financial decisions that has attracted considerable interest is a lack of financial literacy. Financial literacy may be amenable to being altered by public policy. Current research on financial literacy is based on financial literacy scores derived from survey responses. It is important to understand the quality of those measures.

# Interviewer Effects
# and the Measurement of Financial Literacy

**Thomas F. Crossley***
University of Essex and Institute for Fiscal Studies

**Tobias Schmidt**
Deutsche Bundesbank

**Panagiota Tzamourani**
Deutsche Bundesbank

**Joachim K. Winter**
University of Munich

This version: 28 April 2017

**Abstract:** In this paper we ask whether interviewers influence the answers to a standard set of survey questions on financial literacy. We study data from Germany's wealth survey, Panel on Household Finances (PHF). We have access to extensive paradata, including interviewer identifiers, background characteristics of interviewers, and measures of interviewer activity through the survey. We find that interviewer effects explain a significant fraction of the variance of the financial literacy score, and inter-interviewer correlations are notably larger for the financial literacy score than for other survey variables. We explore how accounting for interviewer effects can improve estimates of the effects of financial literacy on financial behaviours and outcomes.

**Keywords:** financial literacy, interviewer effects, measurement error

**Acknowledgements:** We wish to thank participants of the 2015 European Survey Research Conference (ESRA), the 2015 Household Finance and Consumption Conference, and seminars at the University of Essex and the London School of Economics for useful comments and suggestions. Crossley acknowledges support from the ESRC through the ESRC-funded Centre for Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies (reference RES-544-28-5001), through the Research Centre on Micro-Social Change (MiSoC) at the University of Essex, (reference ES/L009153/1), and through a grant to Essex University for "Understanding Household Finance through Better Measurement" (reference ES/N006534/1). Winter acknowledges support from the Deutsche Forschungsgemeinschaft via SFB/TR 190.

**Disclaimer:** The views expressed in this paper represent the authors' personal opinions and do not necessarily reflect the views of the Deutsche Bundesbank or its staff.

* Corresponding author: Thomas Crossley, Department of Economics, University of Essex, Wivenhoe Park, Colchester, U.K., CO4 3SQ. tfcrossley@gmail.com

# 1. Introduction

Insufficient saving and poor financial decision-making are major policy concerns, particularly in the face of increasingly complex financial markets and increasing reliance on individual financial provision for old age. While these concerns have been raised for decades (see, Engen, Gale and Scholz, 1996; Skinner, 2007), recent research has highlighted the limitations of households' decision processes. One explanation for inadequate financial decisions that has attracted considerable interest is a lack of financial literacy (van Rooij et al., 2011, 2012; Hastings et al., 2013; Lusardi and Mitchell, 2014, 2015). This emerging literature argues that poor financial literacy is both causally responsible for suboptimal financial choices of households and individuals, and amenable to being altered by public policy.

Much of the current knowledge about the predictors and effects of financial literacy is based on survey data. Lusardi and Mitchell (2008) proposed a short list of questions on interest rate compounding, on the effects of inflation, and on diversification of securities that can be integrated into existing surveys at low cost. The premise is that individuals should know the answers to these questions in order to make sound decisions on issues of household finance. Indeed, a variety of studies have shown that measures of financial literacy based on the responses to such simple survey questions are correlated with the quality of households' financial decisions and also with long-term financial outcomes, even after controlling for socio-economic characteristics and for cognitive ability. This holds for teenagers who are just beginning to make their own financial decisions as well as for young and older adults, and across both developed and developing countries.[1]

Despite the recent advances in the analysis of financial literacy, measurement error arising from the survey response process is an important concern. Lusardi and Mitchell (2014) summarize studies that use instrumental variables (IV) techniques to estimate models where financial literacy is a right-hand side variable. They observe that IV estimates of the effects of financial literacy in these studies are typically larger than OLS estimates, and conclude that "the noninstrumented estimates of financial literacy may underestimate the true effect" (p. 27).[2] While econometric methods such as IV can resolve endogeneity that arises from measurement error, they are not ideal for several reasons, perhaps the most important of which is the fact that credible instruments are often hard to come by. In this paper, we explore how the survey response process induces measurement error in measures of financial literacy, with the ultimate goals of improving the econometric analysis of the effects of financial literacy and of constructing better survey measures. Specifically, we focus on the role of the survey interviewer.

[1] The existing evidence is reviewed by Hastings et al. (2013), Lusardi and Mitchell (2014), Mitchell and Lusardi (2015); Lührmann et al. (2015) show evidence for teenagers.

[2] In principle, financial literacy measures might suffer from several types of endogeneity. Measurement error in financial literacy will tend attenuate estimates of the effects of financial literacy on behavior and outcomes. On the other hand, reverse causation, from financial behavior to financial literacy would lead simple regressions to overstate the causal effect of financial literacy, as would omitted variables that are correlated with both financial literacy and financial choices. The fact that IV estimates are typically much larger (not smaller) than ordinary regression estimates suggests that measurement error is the key empirical problem.

The survey methodology literature argues that interviewers might affect survey responses in three different ways: unit nonresponse, item nonresponse, and the response itself.[3] We highlight the latter channel, i.e., the possibility that interviewers induce differential measurement error. For example, interviewers might help respondents to better comprehend complex survey questions or they might help respondents to find strategies that enhance the reporting of quantities that are not easily recalled. It is likely that measurement error induced in such ways is heterogeneous across interviewers. While this possibility has been recognized for some time, the survey methodology literature has mostly focused on the implications for variance estimation and item nonresponse. The objective of the present paper is to consider the effects of interviewer-induced measurement error on coefficient estimates in regression models.[4] While the application is concerned with survey measures of financial literature, the general approach is applicable to other settings as well.

As measures of financial literacy are often used in regression models, both as dependent variables and as regressors, it seems important to understand the statistical properties of interviewer-induced measurement error. We develop a tractable analytic framework for thinking about (i) interviewers both as a source of error in survey responses but also as a moderator of respondent errors, (ii) the consequences of interviewer effects for the kinds of models estimated in the financial literacy literature, and (iii) how information on interviewers or interviewer effects might be used to improve estimates of the effects of financial literacy on financial choices and outcomes.

We apply this framework to data on financial literacy collected as part of the German Panel on Household Finances (PHF). The PHF is a large survey on household finance that is representative of the German population. We use data from the first wave of the PHF which was conducted in 2010/11. The questionnaire focuses on households' financial and non-financial assets and debts. It includes the standard financial literacy questions on interest rate compounding, the effect of inflation, and diversification of securities developed by Lusardi and Mitchell (2008).

The PHF survey data we analyze in this paper are unusual in that they not only allow us to identify the interviewers – we also obtained data on a number of interviewer characteristics, including gender, age and education level, from the survey firm that conducted the fieldwork. From detailed contact records we are also able to compute measures of interviewers' contact behaviour and workload. Our analysis will make extensive use of such paradata, and the results highlight the usefulness of paradata (see also Couper, 1998, and Kreuter, 2015). We use these data to test for the independent interviewer effects and for a moderating effect of interviewers on respondent error; to explore whether interviewer effects in financial literacy questions are related to interviewer characteristics (including those, such as interviewer

[3] See Biemer (1980), Platek and Gray (1983), West and Olson (2010), and West, Kreuter and Jaenichen (2013), among others.

[4] A large literature in survey methodology has shown that interviewers may lead to complications in variance estimation (O'Muircheartaigh and Campanelli, 1998; Schnell and Kreuter, 2005) and to nonresponse biases (Durrant et al., 2010). See also West, Kreuter, and Jaenichen (2013), and further references therein.

experience, that might be controlled by survey field agencies); and to evaluate the strategies for mitigating the consequences of interviewer effects that our analytic framework suggests.

We find significant interviewer effects in the financial literacy score, with inter-interviewer correlations notably larger than for other survey variables we examine. We find interview effects in both mean (location) and variance (scale); the later suggests a moderating effect of interviewers on respondent errors. Estimated interviewer effects are weakly related to interview characteristics. There is some evidence that older interviewers elicit responses that indicate higher financial literacy on average, and which are less variable. Different approaches to using the paradata to improve estimates of the effect of financial literacy on outcomes and behaviour give different results. This suggests that the measurement error induced by interviewers has a rich structure.

These results have a number of important implications. Most fundamentally, they reinforce the need for providing survey paradata – specifically, interviewer identifiers and perhaps also interviewer characteristics – along with any household survey dataset. They also highlight the need for more research on the relative importance of interviewer effects across different surveys and survey questions. Finally, as the bias introduced in regression coefficients is difficult to correct after the fact, mitigating interviewer effects appears to be crucial.

The remainder of the paper proceeds as follows. Before turning to the data, we first present a tractable framework for thinking about interview effects. This is presented in Section 2. In Section 3, we describe our data. Section 4 contains our empirical results. We discuss the implications our results have extensively in the concluding section of this paper.

## 2. Statistical Framework

To help organize our interpretation of the data, consider the following statistical framework for thinking about interview effects. We will first develop a measurement model for financial literacy, which we then combine with a simple model of financial decision making with financial literacy as the explanatory variable of interest. Both models can include additional regressors, but we will initially suppress them to simplify the exposition.

Beginning with the measurement model, $FL_{ij}$ is a measure of the variable we are interested in (financial literacy), with true value $FL^*_{ij} = \theta + v_i$ . The subscript $i$ indexes respondents, and $j$ indexes interviewers. The overall mean of true financial literacy is given by $\theta$, and heterogeneity in the true value given by $v_i$, so that $V\left[FL^*_{ij}\right] = \sigma^2_v$. Below we will allow $\theta$ to be a function of observed covariates.

Our model of measured responses is:

$$FL_{ij} = \theta + v_i + \pi_j \omega_i + u_j \tag{1}$$

Response error is $\pi_j \omega_i + u_j$, where $u_j$ is interviewer-level error and $\omega_i$ is individual reporting error, which is moderated by interviewers. So interviewers affect both the mean and variance of measurement response error.[5] An interviewer who (for example, through clarity in posing the questions) reduces respondent error has $\pi_j < 1$ (and an interviewer that exacerbates respondent error has $\pi_j > 1$.) We assume that $\pi_j$, $\omega_i$, $v_i$ and $u_j$ are independent.

A testable restriction on this model is $\pi_j = 1$. This restriction implies common within-group variances, and a more familiar error component structure,

$$FL_{ij} = \theta + v_i + \omega_i + u_j \,, \tag{2}$$

with an interviewer error component ($u_j$) and an individual error component ($\epsilon_j = v_i + \omega_i$) that contains both error and genuine heterogeneity. Given our assumption of independent errors, $\sigma_\epsilon^2 = \sigma_{v+\omega}^2 = \sigma_v^2 + \sigma_\omega^2$. However, as we will show below, the assumption of common within-group variances is rejected in our data.

Carrying on with our more general frame work,

$$FL_{ij} = \theta + v_i + \pi_j\omega_i + u_j$$

with $\omega_i$, $v_i$, $u_j$ and $\pi_j$ independent, it is straightforward to show that:

$$V\left[FL_{ij}\right] = \sigma_u^2 + \sigma_v^2 + (\sigma_\pi^2 + \mu_\pi^2)\sigma_\omega^2 \tag{3}$$

where $\mu_\pi^2 = \left(E\left[\pi_j\right]\right)^2$. The reliability, *R*, of the financial literacy measure is therefore:

$$R = \frac{V[FL_{ij}^*]}{V[FL_{ij}]} = \frac{\sigma_v^2}{\sigma_u^2 + \sigma_v^2 + (\sigma_\pi^2 + \mu_\pi^2)\sigma_\omega^2} \tag{4}$$

Thus, the reliability of the financial literacy measure depends in part on interviewer effects: *R* is decreasing in $\sigma_u^2$, $\sigma_\pi^2$ and $\mu_\pi^2$.

Now note that:

$$E\left[FL_{ij}\middle|j\right] = \theta + u_j \tag{5}$$

$$V\left[E\left[FL_{ij}\middle|j\right]\right] = \sigma_u^2 \tag{6}$$

And:

$$FL_{ij} - E\left[FL_{ij}\middle|j\right] = v_i + \pi_j\omega_i \tag{7}$$

Taking these results together gives:

$$V\left[FL_{ij} - E\left[FL_{ij}\middle|j\right]\right] = V\left[v_i + \pi_j\omega_i\right] = \sigma_v^2 + (\sigma_\pi^2 + \mu_\pi^2)\sigma_\omega^2 \tag{8}$$

so that the intra-class correlation (ICC) is:

$$\frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2 + (\sigma_\pi^2 + \mu_\pi^2)\sigma_\omega^2} \tag{9}$$

And:

[5] Brunton-Smith et al. (2017) also present a framework in which interviewers have both scale and location effects.

$$(1-ICC)=\frac{\sigma_v^2+(\sigma_\pi^2+\mu_\pi^2)\sigma_\omega^2}{\sigma_u^2+\sigma_v^2+(\sigma_\pi^2+\mu_\pi^2)\sigma_\omega^2}=R+\frac{(\sigma_\pi^2+\mu_\pi^2)\sigma_\omega^2}{\sigma_u^2+\sigma_v^2+(\sigma_\pi^2+\mu_\pi^2)\sigma_\omega^2} \tag{10}$$

The key point here is that $(1-ICC)$ provides an upper bound on the reliability $R$ of the financial literacy measure, and that this quantity can be obtained from analysis of variance.

Next, we explore the implications of these results for substantive regressions that contain financial literacy as the independent variable of interest. Suppose the equation of interest is given by:

$$y_{ij}=\alpha FL_{ij}^{*}+e_{ij} \tag{11}$$

where $y_{ij}$ is typically a measure of financial behaviour, such as stock market participation. We assume that $y_{ij}$ is well measured, though we will return to this assumption below. Substituting measured financial literacy for true financial literacy gives:

$$y_{ij}=\alpha(FL_{ij}-\pi_j\omega_i-u_j)+e_{ij} \tag{12}$$

The substantive regression is thus subject to the usual measurement error problem that the independent variable is correlated with components of the error term. Let $\hat{\alpha}$ be the OLS estimate of $\alpha$ in this equation. It is straightforward to show that the coefficient estimate is attenuated:

$$plim\ \hat{\alpha}=\alpha\frac{\sigma_v^2}{\sigma_u^2+\sigma_v^2+(\sigma_\pi^2+\mu_\pi^2)\sigma_\omega^2}=\alpha R \tag{13}$$

Interviewer effects lead to asymptotic bias in the point estimate of the effect of financial literacy on the outcome of interest.

Interviewer effects are often thought of as being similar to design effects or clustering of respondents at sampling points. As such, they are seen as a challenge to inference: variance estimation must account for the correlation structure. However, with this framework we emphasize that interviewer effects can also compromise point estimates because they affect the reliability of the measure of a quantity whose effects are being studied (in our case, financial literacy). Note that correlated *true* heterogeneity in financial literacy (in $v_i$) – as might arise from complex sample designs – does not directly affect the reliability and hence does not lead to inconsistent estimates.[6]

It is well known that in the presence of measurement error in an independent variable, rescaling the least squares estimate by the reliability of the measure gives a better estimate (see Goldstein and French (2015) for a recent example). Here rescaling by *(1-ICC)* improves the estimate because *(1-ICC)* gives a lower bound to the reliability.

$$plim\ \frac{\hat{\alpha}}{1-ICC}=\alpha\frac{\sigma_v^2}{\sigma_v^2+(\sigma_\pi^2+\mu_\pi^2)\sigma_\omega^2} \tag{14}$$

and:

$$\frac{\sigma_v^2}{\sigma_v^2+(\sigma_\pi^2+\mu_\pi^2)\sigma_\omega^2}>\frac{\sigma_v^2}{\sigma_u^2+\sigma_v^2+(\sigma_\pi^2+\mu_\pi^2)\sigma_\omega^2} \tag{15}$$

[6] Note however that, for a given level of response error, if a complex sampling design led to lower total variance of true financial literacy (smaller $\sigma_v^2$) , this would lead to lower reliability.

Intuition may be helped by special cases. If $\pi_j = 1$, then $\frac{1}{(1-ICC)} = \frac{\sigma_v^2+\sigma_\omega^2+\sigma_u^2}{\sigma_v^2+\sigma_\omega^2}$ , $plim\ \hat{\alpha} = \alpha \frac{\sigma_v^2}{\sigma_v^2+\sigma_\omega^2+\sigma_u^2}$ and $plim\ \frac{\hat{\alpha}}{1-ICC} = \alpha \frac{\sigma_v^2}{\sigma_v^2+\sigma_\omega^2}$ . Again $(1-ICC)$ gives an upper bound on the attenuation and $\frac{1}{(1-ICC)}$ gives a lower bound to the required correction factor. If, further, there is no individual component to reporting error $\sigma_\omega^2 = 0$ and $\frac{1}{(1-ICC)} = \frac{\sigma_v^2+\sigma_u^2}{\sigma_v^2} = 1/R$ . In this case, $plim\ \frac{\hat{\alpha}}{1-ICC} = \alpha$. This is the textbook, classical measurement error case, and in this case rescaling by $\frac{1}{(1-ICC)}$ eliminates the asymptotic bias.

This analysis tells us that rescaling the least-squares estimate by one minus the ICC will improve the estimates; it will do so more effectively if (a) interviewers have little moderating effect on individual reporting errors ($\pi_j$ close to 1 for all *j*); and it will completely offset the attenuation if there is no individual component to reporting error.

Estimation is also improved by any "within" transformation that eliminates the interviewer effect. For example, one can take deviations from interviewer means (or equivalently condition on interview dummies). Denoting deviations from interviewer means by $\Delta^j$:

$$y_{ij} = \alpha(FL_{ij} - \pi_j\omega_i - u_j) + e_{ij} \tag{16}$$

$$\Delta^j y_{ij} = \alpha(\Delta^j FL_{ij} - \pi_j \Delta^j \omega_i - 0) + \Delta^j e_{ij} \tag{17}$$

Let $\hat{\alpha}^\Delta$ be the estimate of $\alpha$ obtained from least-squares estimation of the transformed equation. Our independence assumptions imply that $E[v_i|j] = E[v_j] = 0$ and similarly for $\omega_i$, so that $\sigma^2{}_{\Delta v} = \sigma_v^2$ and similarly for $\omega_i$. It is then straightforward to show that:

$$plim\ \hat{\alpha}^\Delta = \alpha \frac{\sigma_v^2}{\sigma_v^2 + (\sigma_\pi^2 + \mu_\pi^2)\sigma_\omega^2} \tag{18}$$

Thus $\hat{\alpha}^\Delta$ suffers from less attenuation than the untransformed estimator and in fact is identical to the untransformed estimator scaled by $\frac{1}{(1-ICC)}$. This means that researchers can improve estimates by sweeping out interviewer effects if interviewer identifiers are available, or by scaling estimates by a factor that depends on the intra-class correlation (ICC), and these should have the same effect. A comparison of these two procedures therefore provides a possible test of the measurement error assumptions laid out above.

*Adding covariates*

The model of interest would typically also involve other covariates, $X$, and as usual in models with measurement error, the bias is more complicated in the presence of additional covariates. In this case we have

$$y = \alpha FL^* + \gamma X + e \tag{19}$$

where $X$ can be a set of covariates (a matrix) and we assume these are well measured. As before assume $FL = FL^* + v$ with *v* independent of $FL^*, X$ and $e$. Then:

$$\alpha^{OLS} = [\, FL' M_X FL]^{-1} [FL' M_X y]$$
$$= [(FL^* + v) M_X (FL^* + v)]^{-1} [(FL^* + v) M_X (\alpha FL^* + \gamma X + e)] \tag{20}$$

where $M_X = I - X(X'X)^{-1}X'$, and so

$$plim(\alpha^{ols}) = [V(M_x FL^*) + V(v)]^{-1} [\alpha V(M_x FL^*)] \tag{21}$$

Note that if $V(v) = 0$ (there is no measurement error) then $plim(\alpha^{ols}) = \alpha$. Also if $M_X FL^* = FL^*$ (if $FL^*$ is orthogonal to *X*) then

$$plim(\alpha^{ols}) = \alpha[V(FL^*) + V(v)]^{-1}[V(FL^*)] = \alpha R \tag{22}$$

where *R* is the reliability of *FL*. However if $FL^*$ is not orthogonal to *X* then the appropriate rescaling involves an adjustment for *X*, and will vary with the choice of covariates.

In our empirical analysis we will allow that response errors associated with respondents might be predicted by individual or household characteristics ($X_i$); similarly, response errors associated with interviewers might be associated with interviewer characteristics ($Z_j$). In addition, conditional on financial literacy, individual or household characteristics may predict financial choices (capturing heterogeneity in choice sets or preferences). Thus our full framework has both a measurement model and a choice model, and a multilevel structure:

$$\begin{gathered} y_i = \alpha FL_i^* + X_i \beta + e_i \\ FL_{ij} = FL_i^* + \pi_j \omega_i + u_j \\ FL_i^* = X_i \theta + v_i \\ \omega_i = X_i \gamma^\omega + \widetilde{\omega}_i \\ u_j = Z_j \gamma^u + \tilde{u}_j \\ \left(\pi_j\right)^2 \sigma_\omega^2 = Z_j \gamma^\pi + \widetilde{\pi}_j \end{gathered} \tag{23}$$

We begin by estimating the ICC for financial literacy questions and a financial literacy score. We also investigate how estimates of $\alpha$ are affected by adjusting for the reliability of the financial literacy score or by conditioning on interviewer identifiers. Finally, we report estimates of $\gamma^u$ and $\gamma^\pi$, which capture the relationship between interviewer characteristics and response errors. Further estimation details are given below.

## 3. Data

The German "Panel on Household Finances (PHF)" is a face-to-face CAPI survey focused on measuring household wealth. It was carried out on behalf of the Deutsche Bundesbank by infas GmbH Bonn in 2010/2011.[7] The PHF field phase consisted of two major parts. In our study we include only the 1,705 interviews from the first part, since in the second part the allocation of selected households/addresses to interviewers was not completely random. The first part of the PHF field phase started in September 2010 and lasted until February 2011**.** In part one 178 interviewers conducted at least one interview. For the analysis we only consider interviewers with at least three interviews, which reduces our sample to 160 interviewers.[8]

The PHF survey provides a representative picture of the population of non-institutionalized households in Germany and focuses mainly on their financial and non-financial assets and liabilities (secured and unsecured debt). It also collects information on income, employment and pensions. The core questionnaire program is supplemented with, among others, questions about financial literacy. It includes the standard questions on interest rate, the effect of inflation and diversification of securities developed by Lusardi and Mitchell (2008):

(1) ***Let us assume that you have a balance of €100 on your savings account. This balance bears interest at a rate of 2% per year and you leave it for 5 years on this account. How high do you think your balance will be after 5 years?***

| | |
|---|---|
| **1 - More than €102** | **-1 - Don't know** |
| **2 - Exactly €102** | **-2 - No answer** |
| **3 - Less than €102** | **-3 - Question filtered** |

*(2)* ***Let us assume that your savings account bears interest at a rate of 1% per year and the rate of inflation is 2% per year. Do you think that in one year's time the balance on your savings account will buy the same as, more than or less than today***

| | |
|---|---|
| **1 - More** | **-1 - Don't know** |
| **2 - The same** | **-2 - No answer** |
| **3 - Less than today** | **-3 - Question filtered** |

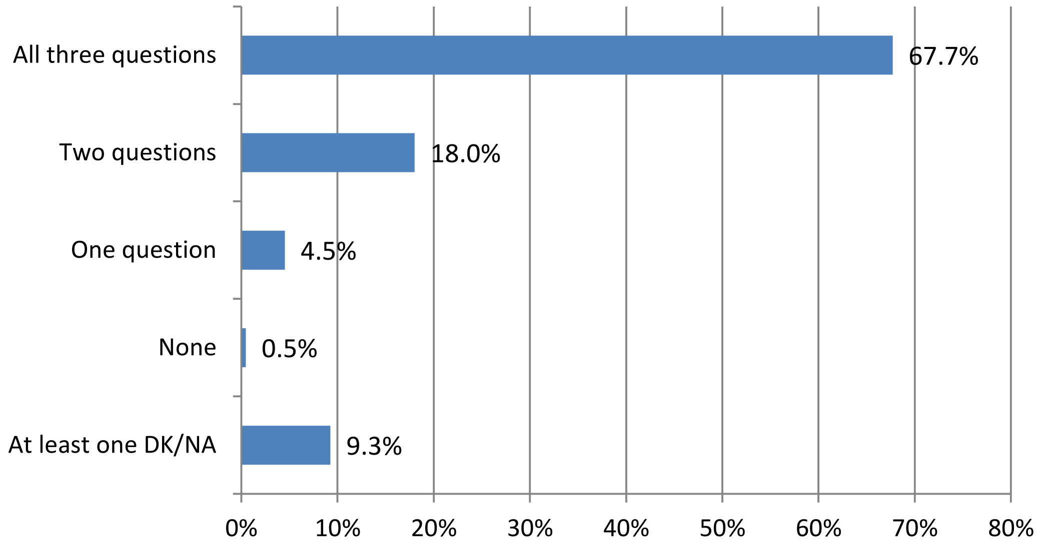
*(3)* ***Do you agree with the following statement: "Investing in shares of one company is less risky than investing in a fund containing shares of similar companies"?***

| | |
|---|---|
| **1 - Agree** | **-1 - Don't know** |
| **2 - Disagree** | **-2 - No answer** |
| | **-3 - Question filtered** |

[7] See von Kalckreuth et al. (2010) for a detailed description of the survey.
8 We dropped 15 interviewers with only one interview and another 15 interviewers with exactly two interviews.

The questions are administered to one person in each household, the reference person, who is selected as being the most knowledgeable on the household's finances. We follow Bucher-Koenen and Lusardi (2011) and Bucher-Koenen and Ziegelmeyer (2014) and aggregate the answers of this reference person to these three questions into a "financial literacy score", which is the number of correct answers. Missing values (DK/NA) are treated as incorrect responses in the baseline specification (as either an incorrect response or a "Don't know" indicate a lack of knowledge). Figure 1 below shows that almost 68% of respondents in our sample provide correct answers to all three literacy questions, 18% get two right, 4.5% one. About 10% have missing values for at least one question.[9] As a robustness check we also conduct our analysis excluding missing answers in the calculation of the literacy score.

**Figure 1: Number of correctly answered financial literacy questions**




*Source: PHF 2010/2011.*

The survey is accompanied by a large collection of paradata, among them the interviewers' id-numbers, as well as interviewers' gender, age and level of education.[10] From contact protocols of the field work and the survey data we can construct indicators of interviewers' contact behaviour (average number of contact attempts per case), their performance (number of successful interviews as a share of addresses issued), and two indicators of quality (percentage of DK/NA responses to all survey questions; interview-time in seconds per item/question). All these indicators are meant to characterise the interviewer in general and are therefore calculated based on information from both part one and two of the PHF survey.

To mitigate the possibility that within-interview correlations reflect clustering of similar households at sample points we also control for various observable factors that have been found to be

[9] See the Appendix for detailed information on individual questions.
[10] See the Appendix for summary statistics on interviewers' characteristics.

related to financial literacy, as reported, for example, in Lusardi and Mitchell (2014). We include respondent's personal socio-demographic characteristics, i.e. age, gender, education, employment status, nationality, as well as household characteristics, i.e. household size and household income.

To put the estimated size of the interviewer effects on financial literacy into perspective we also estimate interviewer effects for an number of additional survey measures, e.g. an 11-point Likert-scale question on life-satisfaction, a question on total household net income (in Euro) and a question about qualitative inflation expectations ("change in the general price level") in the next 12 months.

Finally, we estimate substantive regressions with financial literacy as an explanatory variable. Our choice of models and dependent variables for this exercise follows some of the most prominent examples in the literature: Christelis, Jappelli, and Padula (2010) showed that more financially literate individuals are more likely to own stocks. Bucher-Koenen and Lusardi (2011) showed a positive effect of financial literacy on retirement planning. Van Rooij et al. (2012) showed that financial literacy is positively related to retirement planning and the development of a savings plan and to stock-holding, and through these channels it has a positive effect on wealth accumulation.

# 4. Results

## 4.1 The magnitude and correlates of interviewer effects

We now implement the statistical model developed in Section 2 above. The starting point is the equation for financial literacy:

$$FL_{ij} = X_i\theta + v_i + \pi_j\omega_i + u_j \tag{24}$$

We begin by documenting the size of interview effects in financial literacy questions, relative to other typical questions in a household finance survey. Our measure of the size of interviewer effects is the intraclass correlation (ICC), which we compute from a random effects regressions estimated by GLS. As described in Section 2, the ICC is the ratio of the variance of interviewer effects to the total variance of the financial literacy score. Our results are presented in Table 1.

The first column of Table 1 shows the ICC computed from a random effects regression with no covariates (just a constant). Starting from the top, the ICC for our financial literacy score indicates that 18% of the variance in this score in the PHF can be attributed to interviewers. The second through fourth row report the ICC for the individual financial literacy questions which contribute to the overall score. The ICC is large for all three questions, so the large ICC in the overall score is not driven by a single question. The ICC is largest for the inflation question and smallest for the portfolio diversification question. The remaining rows report the ICC for typical variables from a household finance survey. ICCs for other questions are significantly smaller than for the financial literacy questions.

One possible concern with these results is that ICC is actually capturing within-interviewer correlation in true financial literacy (the $v$ in Section 2) as respondents are not randomly assigned to interviewers. In particular, interviewers are typically associated with particular sampling points. In the absence of an interpenetrating design, it is possible that the interviewer effects we estimate are actually design effects.[11] To address this concern we re-estimate the ICC adding a rich set of individual and household characteristics $X_i$, to the random effects model (as in equation 24). The results are in the second column of Table 1. Note that this has almost no effect on the ICCs for the financial literacy variables. Design effects would arise primarily from the clustering of similar households at sampling points, and should therefore be diminished by controlling for a rich set of respondent observables. Consistent with this view, the ICCs for many other variables *do* fall significantly when we control for observables (see for example, car ownership and life satisfaction.) Thus these findings support our interpretation of the ICC for financial literacy variables as measuring genuine interviewer effects.

Table 2 repeats this analysis for the nonresponse in the financial literacy questions. In Table 1 the Don't Knows/No Answers (DK/NA) are treated as incorrect answers. To check the effect of non-response, we constructed an indicator taking the value of 1 if the question had a DK/NA and 0 otherwise and computed the ICC's for these indicators. The ICC's are generally much smaller than those reported for the financial literacy questions in Table 1, except for the portfolio diversification question. This suggests that for the first two questions very little of the interviewer variation in Table 1 is driven by the DK/NA responses.

We next add interviewer fixed effects to the mean and estimate

$$\widehat{FL_{ij}} = X_i\hat{\theta} + \hat{u}_j \tag{26}$$

We then take the estimated interviewer fixed effects, $\hat{u}_j$, and fit them to interviewer characteristics (following equation 23 above):

$$\hat{u}_j \quad = Z_j\gamma^u \quad (+error\,) \tag{27}$$

The results are presented in Table 3. We find few significant predictors of the interviewer effects on the mean, with only interviewer age being statistical significant, and less than 10% of the cross-interviewer variation explained by interviewer characteristics.

The statistical model in Section 2 allows for the possibility that interviewers moderate response errors, and do so with differing ability. This generates an interaction between individual mean responses and an interviewer parameter ($\pi_j\omega_i$); there are interviewer effects in the variance as well as the mean. Referring back to equation (24), note that the residuals from this model with interviewer fixed effects in estimate:

$$residuals = v_i \widehat{+\pi_j}\omega_i \tag{28}$$

Given our assumption that

$$\sigma_\omega^2 \; and \;\; \sigma_v^2 \; are \; constant \; across \; interviewers \; j, \tag{30}$$

[11] Studies that have had interpenetrating designs have found that interviewer effects are as large (O'Muircheartaigh and Campanelli, 1998) or larger (Schnell and Kreuter, 2005) than cluster/design/sampling point effects. See also Vassallo et al. (2017) for discussion of separating interviewer and design effects.

we can test for interview effects on the scale (or variance), $Ho\colon \pi_j = \pi$ , by testing the equality of the variance of the residuals across individuals:

$$Ho\colon (\pi_j)^2 \sigma_\omega^2 + \sigma_v^2 = constant\ across\ interviewers\ j, \tag{29}$$

In Table 4 each row presents alternative tests of variance equality. The first column gives results for the financial literacy score (the number of correct answers) and the following columns contain those for each individual financial literacy test. In every case, the null is rejected at conventional levels of statistical significance. We interpret these results as evidence of interview effects in variance (not just the mean), which in turn lends support to the idea that interviewers moderate individual response errors, perhaps with heterogeneous skill.

To explore these interviewer effects in scale further, we fit the squared residuals to interviewer characteristics

$$(residuals)^2 = Z_j \gamma^\pi \quad (+error) \tag{32}$$

to estimate

$$(\pi_j)^2 \sigma_\omega^2 = Z_j \gamma^\pi \quad (+error) \tag{33}$$

Under the maintained assumption that $\sigma_\omega^2$ is constant across interviewers, the coefficients are $\gamma^\pi$ are proportional to the interviewer effects on variance. The results are reported in Table 5. We find more significant predictors of the interviewer effects on variance (scale) than we did for interviewer effects in mean (level). Interviewer age and percentage of DK/NA responses associated with the surveys conducted by an interviewer are statistically significant. Note however that we again fail to explain much of the cross-interviewer variation.

The bottom line from this analysis seems to be that there are significant interviewer effects, in both mean and variance, but find little relationship between these effects and observable interviewer characteristics.

### 4.2 Estimating the effects of financial literacy on outcomes

We now consider the impact of these interviewer effects on estimates of the impact of financial literacy on a variety of financial outcomes – that is, we put the financial literacy measure on the right-hand side of "substantive" regressions. We also explore alternative approaches to improving those estimates. Starting with substantive equation

$$y_{ij} = \alpha FL_i^* + X_i \beta + e_{ij} \tag{34}$$

We use the orthogonal projection matrix, $M_X$, to eliminate $X$ giving:

$$M_x y_{ij} = \alpha M_x FL_i^* + e_{ij} \tag{35}$$

(under the maintained assumption that $e_{ij}$ is orthogonal to $X$.) We fit this equation, and then also this equation augmented by interviewer fixed effects (which are also assumed orthogonal to $X$):

$$M_x y_{ij} = \alpha' M_x FL_i^* + u_j + e_{ij} \tag{36}$$

Recall that the statistical model developed in Section 2 implies two things: First, that the interviewer effects lead to attenuation of estimates of the effects of financial literacy on financial behaviours and outcomes, and that, so long as outcomes are well-measured, the degree of attenuation is independent of the outcome under study, $y_{ij}$. Second, the attenuation can be reduced either by sweeping out the interviewer effects with an appropriate fixed-effects estimator (to give $\alpha'$ above), or by rescaling the OLS estimates by the ICC (to give $\frac{\alpha}{(1-ICC)}$ ). Further, either approach should give the same answer (as long as we deal with covariates appropriately). Thus comparing $\alpha'$ and $\frac{\alpha}{(1-ICC)}$ to the uncorrected estimates ($\alpha$) reveals the impact of the interviewer effects on estimates of the effect of financial literacy on behaviour and outcomes, while comparing $\alpha'$ *to* $\frac{\alpha}{(1-ICC)}$ allows us to assess the adequacy of the statistical model developed in Section 2.

Table 6 presents estimated FL effects from four models of the form given by equation (34). The financial behaviours we consider are participation in mutual funds, participation in bonds, participation in equities and participation in a private pension. The first column gives OLS estimates from a linear probability model. The second column then corrects these estimates with (one minus) the estimated ICC from Table 1 (allowing for covariates).  This of course raises the estimated effects, offsetting the presumed attenuation. By construction, the proportional correction is the same for each outcome.

The third column of Table 6 then presents estimates these models that accounting for interview fixed-effects. In three of the four cases the estimated effect increases, consistent with attenuation due to the interviewer effects. This suggests that using paradata to account interviewer fixed effects or rescale estimates by the ICC is a useful strategy for minimizing the effects of measurement error in empirical studies of financial literacy. However, the magnitude of the change is identical to re-scaling by the *ICC* in only one of four cases. This suggests that the structure of measurement errors in the present data is even richer than we assume; we leave further exploration of this possibility to future research.

## 5. Conclusion

We present a tractable model of measurement error arising from interviewer effects. We allow for interview effects in both the mean and variance of responses. The latter captures heterogeneity in the ability of interviewers to moderate respondent errors. The model clarifies how interviewer effects lead to inter-interviewer correlation of survey responses (similar to the well-known algebra of intra-class correlation). We further show that these correlations lead to biased regression coefficients when the dependent variable is subject to interviewer effects. We derive a correction factor that rescales regression coefficients so that they are purged from the effects of intra-interviewer correlations. The model is straightforward to estimate if interviewer identifiers are included in the data, but in a regression context the correction factors depend on the covariates.

An informal test of the validity of this response model can be performed by comparing the rescaled coefficient estimates to those obtained by estimation with interviewer fixed effects. These approaches should give the same result if the multi-level response model we propose is correctly specified. In our data, this test fails. Both approaches suggests that the uncorrected estimates are attenuated, but the empirical difference in corrections implies that, at least in these data, the structure of response error is more complicated than allowed for by our model. It seems unlikely that such errors are amenable to econometric correction using instrumental variable techniques (which in their standard implementation require measurement error to be classical).

Our results might be specific to our data. Even though we believe that the financial literacy application on which we focus is very typical of situations in which interviewer effects might affect applied economic analysis, the importance of interviewer effects should be studied in other fields of applied research that use survey data as well. We thus urge applied researchers to estimate similar models with other data.

Interestingly, we find much larger interviewer effects for financial literacy questions than for a wide range of other questions in the same survey. Financial literacy questions are unusual in that they are *testing* the respondent's knowledge, and in that it is very likely that the interviewer knows the answer, and thus is unusually able to guide the respondent. This is not the case, for example, with questions about a respondent's attitudes, or financial circumstances (the interviewer does not know the respondent's income or wealth). This raises the possibility that interviewer effects may also be important in other survey questions which share these characteristics, such as cognitive tests, or measures of health literacy. Thus these should be priority areas for further investigation.

In any event, our theoretical analysis and our results suggest that every effort should be made to avoid measurement error arising from interviewer effects from the outset. An attempt to correct them after the fact seems to be feasible only in a best-case scenario in which restrictive assumptions on the process that generates interviewer effects hold and certain survey paradata were released.[12] Avoiding or reducing interviewer effects could be achieved either by altering interviewer behaviour or by moving away from personal interviews in favor of self-completion survey modes.

[12] Survey administrators may be unwilling or unable to release interview identifiers to data users. But they may be willing to release intra-class-correlations for key variables, including financial literacy. Some survey organizations are unwilling or unable to release cluster information but do release variance inflation factors to allow users to account for design effects. Releasing an estimate of intra-interviewer correlation would be similar. Note however that the exact correction required depends on covariates in the regression model (see Section 2).

## References


BANKS, J., C. O'DEA, AND Z. OLDFIELD (2010): "Cognitive Function, Numeracy and Retirement Saving Trajectories," *Economic Journal*, 120(548), F381-F410.

BANKS, J., AND Z. OLDFIELD (2007): "Understanding Pensions: Cognitive Function, Numerical Ability and Retirement Saving," *Fiscal Studies*, 28(2), 143-170.

BIEMER, P. (1980): ''A Survey Error Model Which Includes Edit and Imputation Error.'' *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 616–21.

BROWNING, M., T. F.CROSSLEY, AND J. K. WINTER (2014): "The measurement of household consumption expenditures," *Annual Review of Economics*, 6(1), 475-501.

BRUNTON-SMITH, I., P. STURGIS, AND G. LECKIE (2017): "Detecting and understanding interviewer effects on survey data by using a cross-classified mixed effects location-scale model," *Journal of the Royal Statistical Society*, Series A, 180(2), 551-568

BUCHER-KOENEN, T. (2009): "Financial literacy and private old-age provision in Germany – Evidence from SAVE 2008," *Mannheim Research Institute for the Economics of Aging*, Discussion Paper No. 192.

BUCHER-KOENEN, T., AND A. LUSARDI (2011): "Financial literacy and retirement planning in Germany," *Journal of Pension Economics and Finance*, 10(4), 565-584.

BUCHER-KOENEN, T., AND M. ZIEGELMEYER (2014): "Once Burned, Twice Shy? Financial Literacy and Wealth Losses during the Financial Crisis," *Review of Finance*, 18 (6), 2215-2246.

COUPER, M. (1998): "Measuring survey quality in a CASIC environment," *Proceedings of the Section on Survey Research Methods of the American Statistical Association*.

CHRISTELIS, D., T. JAPPELLI, AND M. PADULA (2010): "Cognitive abilities and portfolio choice," *European Economic Review*, vol. 54 (1), 18-39.

ENGEN, E. M., W. G. GALE, AND J. K. SCHOLZ (1996): "The illusory effects of saving incentives on saving," *The Journal of Economic Perspectives*, 10(4), 113-138.

DAVIS, R. E., M. P. COUPER, N. K. JANZ, C. H. CALDWELL, AND K. RESNICOW (2010): "Interviewer effects in public health surveys," *Health Education Research*, 25(1), 14-26.

DURRANT G., R. M. GROVES, L. STAETSKY, AND F. STEELE (2010): "Effects of interviewer attitudes and behaviours on refusal in household surveys," *Public Opinion Quarterly*, 74, 1-36.

GOLDSTEIN H. AND R.FRENCH (2015): "Differential educational progress and measurement error", *Longitudinal and Life Course Studies*, 6, 331---376, http://dx.doi.org/10.14301/llcs.v6i3.361

HASTINGS, J. S., B. C. MADRIAN, AND W. L. SKIMMYHORN (2013): "Financial Literacy, Financial Education and Economic Outcomes," *Annual Review of Economics*, 2013(5), 347-373.

KREUTER, F. (2015): "The use of paradata." In: U. Engel et al. (Ed.), *Improving survey methods. Lessons from recent research*, New York: Routledge, 303-315.

LE BLANC, J., AND T. SCHMIDT (2013): "Knocking on respondents' doors - interviewers and unit non-response in a large wealth survey," *JSM Proceedings, Survey Research Methods Section. Alexandria, VA: American Statistical Association*, 943-957. Retrieved from https://www.amstat.org/sections/SRMS/Proceedings/y2013/Files/308000_80087.pdf.

LÜHRMANN, M., M. SERRA-GARCIA, AND J. K. WINTER (2015): "Teaching teenagers in finance: Does it work?," *Journal of Banking and Finance*, 54, 160-174.

LUSARDI, A., AND O. S. MITCHELL (2008): "Planning and Financial Literacy: How Do Women Fare?," *American Economic Review*, 98(2), 413-417.

LUSARDI A., AND O. MITCHELL (2009): "How ordinary consumers make complex economic decisions: Financial literacy and retirement readiness," National Bureau of Economic Research Working Paper 15350.

LUSARDI, A., AND O. S.MITCHELL (2014): “The Economic Importance of Financial Literacy: Theory and Evidence,” *Journal of Economic Literature*, 52(1), 5-44.

MITCHELL, O.S., AND A. LUSARDI (2015): Financial Literacy and Economic Outcomes: Evidence and Policy Implications,” Working Paper, Pension Research Council PRC WP2015-01

O'MUIRCHEARTAIGH, C., AND P. CAMPANELLI (1998): “The relative impact of interviewer effects and sample design effects on survey precision,” *Journal of the Royal Statistical Society*, Series A, 161(1), 63-77.

PLATEK, R. AND G: B. GRAY (1983): ‘‘Imputation Methodology.’’ In *Incomplete Data in Sample Surveys*, eds. W.G. Madow, I. Olkin, and D.B. Rubin, 255–94. New York: Academic Press.

SCHNELL, R., AND F. KREUTER (2005): “Separating interviewer and sampling-point effects,” *Journal of Official Statistics*, 21(3), 389-410.

SKINNER, J. (2007): “Are You Sure You’re Saving Enough for Retirement?,” *Journal of Economic Perspectives*, 21(3), 59–80.

VAN ROOIJ, M. C. J., A. LUSARDI, AND R. J. M. ALESSIE (2011): “Financial literacy and stock market participation,” *Journal of Finance*, 101:2, 449-472.

VAN ROOIJ, M. C. J., A. LUSARDI, AND R. J. M. ALESSIE (2012): “Financial literacy, retirement planning and household wealth,” *The Economic Journal*, 122, 449-478.

VASSALLO, R., G. DURRANT, AND P. W. F. SMITH (2017): “Separating Interviewer and Area Effects Using a Cross-Classified Multilevel Logistic Model: Simulation Findings and Implications for Survey Designs,” *Journal of the Royal Statistical Society*, Series A, 180(2):531-550.

VON KALCKREUTH, U., M. EISELE, J. LE BLANC, T. SCHMIDT, AND J. ZHU (2012): “The PHF: a comprehensive panel survey on household finances and wealth in Germany,” Deutsche Bundesbank Discussion Paper, No 13/2012.

WEST, B. T., AND K. OLSON (2010): “How much of interviewer variance is really nonresponse error variance?,” *Public Opinion Quarterly*, 74, 1004-1026.

WEST, B. T., F. KREUTER, AND U. JAENICHEN (2013): ““Interviewer” Effects in Face-to-Face Surveys: A Function of Sampling, Measurement Error, or Nonresponse?,” *Journal of Official Statistics*, Vol. 29 (2), 277–297

## Tables and Figures

**Table 1: Intraclass Correlation Coefficients for Financial Literacy (FL) and Other Variables**

| Variable of interest | (a) Interviewer Random Effects only<br>RHO_RE | (b) Household Characteristics + Interviewer Random Effects<br>RHO_HC_RE |
|---|---|---|
| FL score: sum of correct answers | 17.9 | 14.5 |
| FL score: binary, 1 if all correct | 18.0 | 16.2 |
| Q1 Interest rate | 16.9 | 15.6 |
| Q2 Effects of inflation | 17.4 | 12.9 |
| Q3 Diversification | 9.2 | 7.5 |
| Has saving accounts | 4.0 | 0.3 |
| Has mutual funds | 4.6 | 0.1 |
| Has bonds | 0.0 | 0.0 |
| Has shares | 5.5 | 1.8 |
| Has private pension plan | 0.8 | 3.7 |
| Price expectations | 7.7 | 5.6 |
| Has mortgage | 0.0 | 0.0 |
| Has credit cards | 10.7 | 7.4 |
| Has cars | 7.6 | 2.5 |
| Discretionary saving | 0.0 | 0.0 |
| Satisfaction with life | 4.6 | 1.1 |
| Self-assessment: risk | 2.7 | 1.1 |
| Self-assessment: patience | 3.3 | 3.3 |
| Visit to religious service | 8.7 | 8.3 |

Notes:

(a) ICCs are estimated from a random effects model estimated by GLS. See text for further details.
(b) Individual/HH Characteristics included (column 2 only): Reference Person: born in Germany (dummy), female (dummy), Age (<35, 35-44, 45-54, 54-64, 65+), Employment (1 gainfully employed, 2 self-employed, 3 other), Education (1-low, 2-medium, 3-high), Household Characteristics: gross household income (quintiles), HH-Size (1, 2, 3, 4+), Stratum indicator.

**Table 2: Intra-Class Correlations for Non-Response to Financial Literacy**

| Variable of interest | (a) Interviewer Random Effects only | (b) Household Characteristics + Interviewer Random Effects |
|---|---|---|
| Don't Know/ No Answer in: | | |
| At least one Financial Literacy question | 9.5% | 7.7% |
| Q1 Interest rate | 0.0% | 0.0% |
| Q2 Effects of inflation | 3.5% | 2.4% |
| Q3 Diversification | 10.1% | 8.5% |

Notes:

(a) ICCs are estimated from a random effects model estimated by GLS. See text for further details.

(b) Individual/HH Characteristics included: Reference person characteristics: born in Germany (dummy), female (dummy), Age (<35, 35-44, 45-54, 54-64, 65+), Employment (1 gainfully employed, 2 self-employed, 3 other), Education (1-low, 2-medium, 3-high), Household characteristics: gross household income (quintiles), HH-Size (1, 2, 3, 4+), Stratum indicator.

**Table 3: Fit of Interviewer fixed effects in means to interviewer characteristics**.

| | model 1 | model 2 | model 3 |
|---|---|---|---|
| INT: Female | -0.04 | -0 | -0.03 |
| | (0.058) | (0.059) | (0.061) |
| INT: Age 45-64 | 0.085 | 0.06 | 0.069 |
| | (0.060) | (0.061) | (0.062) |
| INT: Age 65+ | 0.189* | 0.181* | 0.177* |
| | (0.073) | (0.074) | (0.077) |
| INT: Medium Education ("Mittlere Reife") | 0.071 | 0.08 | 0.086 |
| | (0.093) | (0.093) | (0.094) |
| INT: High Education ("Abitur, Hochschule") | 0.089 | 0.09 | 0.095 |
| | (0.091) | (0.091) | (0.092) |
| INT: DK/NA percentage | | -4 | -3.82 |
| | | (2.274) | (2.293) |
| INT: Number of contact attempts | | 0 | 0 |
| | | (0.000) | (0.000) |
| INT: Share of interviews in number of addresses received | | -0.1 | -0.04 |
| | | (0.345) | (0.363) |
| INT: Interview-time in seconds per item | | 0 | 0.002 |
| | | (0.005) | (0.005) |
| stratum: wealthy small municipality | | | 0.023 |
| | | | (0.070) |
| stratum: large municipality, wealthy street | | | 0.087 |
| | | | (0.078) |
| stratum: large municipality, other street | | | 0.025 |
| | | | (0.078) |
| Constant | 0.539*** | 0.618** | 0.563** |
| | (0.102) | (0.201) | (0.210) |
| N | 160 | 160 | 160 |
| R-sqr | 0.056 | 0.09 | 0.097 |
| AIC | 105.3 | 108 | 112.2 |
| BIC | 123.8 | 138 | 152.2 |

Notes:

(a) Interviewer (INT) fixed effects obtained from financial literacy (FL) regression on respondent/household characteristics and interviewer fixed effects. See text for further details.

(b) Standard errors in parenthesis, * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

**Table 4: Tests Interviewer Effects in Scale**

**(p-values)**

| | Number of correct answers | Interest rate | Inflation | Diversification |
|---|---|---|---|---|
| P-value for Brown and Forsythe's F statistics (trimmed mean) | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| P-value for Brown and Forsythe's F statistic (median) | <0.0001 | <0.0001 | <0.0001 | 0.00061 |
| P-value for Leven's F-statistic | <0.0001 | <0.0001 | <0.0001 | <0.0001 |

Notes:

(a) F-Tests for equality of residual variance, across interviewers, from models with interviewer effects in mean. See text for further details.

**Table 5: Fit of Interviewer fixed effects in variances to interviewer characteristics**

| | model 1 | | model 2 | |
|---|---|---|---|---|
| INT: Female | 0.015 | | 0.014 | |
| | (0.044) | | (0.036) | |
| INT: Age 45-64 | -0.046 | | -0.034 | |
| | (0.048) | | (0.046) | |
| INT: Age 65+ | -0.132 | * | -0.101 | * |
| | (0.056) | | (0.050) | |
| INT: Medium Education ("Mittlere Reife") | -0.006 | | 0.024 | |
| | (0.058) | | (0.051) | |
| INT: High Education ("Abitur, Hochschule") | -0.009 | | 0.02 | |
| | (0.049) | | (0.048) | |
| INT: DK/NA percentage | | | 5.306 | * |
| | | | (2.153) | |
| INT: Number of contact attempts | | | 0 | |
| | | | (0.000) | |
| INT: Share of interviews in number of addresses received | | | 0.071 | |
| | | | (0.302) | |
| INT: Interview-time in seconds per item | | | -0.001 | |
| | | | (0.004) | |
| Stratum: wealthy small municipality | | | -0.145 | ** |
| | | | (0.046) | |
| Stratum: large municipality, wealthy street | | | -0.18 | *** |
| | | | (0.047) | |
| Stratum: large municipality, other street | | | -0.109 | * |
| | | | (0.052) | |
| Constant | 0.358 | *** | 0.372 | ** |
| | (0.064) | | (0.139) | |
| | | | | |
| N | 1705 | | 1705 | |
| R-sqr | 0.008 | | 0.03 | |
| AIC | 3203.1 | | 3177.6 | |
| BIC | 3235.7 | | 3248.3 | |

Notes:

a) Standard errors in parenthesis: * p<0.05, ** p<0.01, *** p<0.001

**Table 6: Effects of Financial Literacy on Financial Behaviours**

| | | A<br>Estimated Effect, no interviewer fixed effects | **B**<br>Corrected with ICC | **C**<br>Estimated Effect, Interviewer fixed effects |
|---|---|---|---|---|
| Has mutual funds | Fl coeff | 0.062 | 0.073 | 0.055 |
| | s.e. | (0.014) | | (0.016) |
| Has bonds | Fl coeff | 0.032 | 0.038 | 0.036 |
| | s.e. | (0.008) | | (0.010) |
| Has shares | Fl coeff | 0.026 | 0.030 | 0.026 |
| | s.e. | (0.015) | | (0.015) |
| Has private pensions | Fl coeff | 0.020 | 0.024 | 0.048 |
| | s.e. | (0.017) | | (0.018) |

## Appendix

**Table A1 – Answers to financial literacy questions in PHF 2010**

| Literacy Question | | Percentage of respondents in % |
|---|---|---|
| Q1 Interest rate | Correct | 87.7 |
| | Incorrect | 10.9 |
| | Don't know | 1.2 |
| | No answer | 0.2 |
| Q2 Effects of inflation | Correct | 92.0 |
| | Incorrect | 6.4 |
| | Don't know | 1.3 |
| | No answer | 0.4 |
| Q3 Diversification | Correct | 77.1 |
| | Incorrect | 14.2 |
| | Don't know | 7.7 |
| | No answer | 1.0 |
| Number of correctly answered questions | None | 0.5 |
| | One question | 4.5 |
| | Two questions | 18.0 |
| | All three questions | 67.7 |
| | At least one "Don't know" | 8.0 |
| | At least one "No answer" | 1.3 |

**Table A2 – Interviewer Characteristics**

| | | Interviewers | Percent |
|---|---|---|---|
| Interviewers' education | Low (Volks-/Hauptschulabschluss) | 18 | 11.3 |
| | Medium (Mittlere Reife) | 56 | 35.0 |
| | High (Fachhochschulreife, Abitur, Hochschule) | 86 | 53.8 |
| Interviewers' gender | Male | 99 | 61.9 |
| | Female | 61 | 38.1 |
| Interviewers' age group | Up to 54 | 63 | 39.4 |
| | 55-64 | 63 | 39.4 |
| | 65+ | 34 | 21.3 |
| All | | 160 | 100.0 |

**Table A3 – Descriptive Statistics for all Variables used in the Analysis**

| | Mean | St. deviation |
|---|---|---|
| ***Interviewer- Level (n=160)*** | | |
| INT: female (%) | 38.2 | 48.7 |
| INT: Age | 56.2 | 10.5 |
| INT: Age 18-44 (%) | 39.4 | 49.0 |
| INT: Age 45-64 (%) | 39.4 | 49.0 |
| INT: Age 65+ (%) | 21.3 | 41.0 |
| INT: Low Education (%) | 11.3 | 31.7 |
| INT: Medium Education (%) | 35.0 | 47.8 |
| INT: High Education (%) | 53.7 | 50.0 |
| INT: DK/NA percentage (%) | 1.4 | 1.2 |
| INT: Number of contact attempts | 352.6 | 266.0 |
| INT: Share of interviews in number of addresses received (%) | 18.0 | 8.1 |
| INT: Interview-time in seconds per item | 29.9 | 5.5 |
| ***Household-Level (n=1,705)*** | | |
| Gross household income (annual, in Euro) | 62,017 | 71,902 |
| Household (hh) size | 2.26 | 1.14 |
| HH Size 1 (%) | 25.5 | 43.6 |
| HH Size 2 (%) | 44.2 | 49.7 |
| HH Size 3 (%) | 15.4 | 35.6 |
| HH Size 4+ (%) | 14.9 | 35.6 |
| Stratum 1: “poor” small municipality (<100 000 inhabitants) (%) | 35.5 | 47.9 |
| Stratum 2: “wealthy” small municipality (<100 000 inhabitants) (%) | 30.2 | 45.9 |
| Stratum 3: wealthy street section in large municipality (>100 000 inhabitants) (%) | 17.0 | 37.5 |
| Stratum 4: poor street section in large municipality (>100 000 inhabitants) (%) | 17.3 | 37.8 |
| ***Asset Ownership and Savings*** | | |
| Has mutual fund (%) (n=1,682) | 23.2 | 42.3 |
| Has bonds (%) (n=1,690) | 9.4 | 29.2 |
| Has shares (%) (n=1,687) | 18.7 | 39.0 |
| Has private pension (%) (n=1,488) | 34.0 | 47.3 |
| Has savings account (%) (n=1,701) | 81.8 | 38.6 |
| Has mortgage (%) (n=968) | 42.7 | 49.5 |
| Has credit card (%) (n=1,703) | 52.0 | 50.0 |
| Has cars (%) (n=1,703) | 82.8 | 37.8 |
| Discretionary saving (%) (n=1,698) | 16.5 | 37.1 |
| ***Respondent-Level (n=1,705)*** | | |
| Age | 53.9 | 16.1 |
| Age <35 (%) | 13.7 | 34.4 |
| Age 35-44 (%) | 15.1 | 35.8 |

| | | |
|---|---|---|
| Age 45-54 (%) | 21.8 | 41.3 |
| Age 55-64 (%) | 19.8 | 39.8 |
| Age 65+ (%) | 29.6 | 45.7 |
| Female (%) | 42.8 | 49.5 |
| Education low (%) | 9.1 | 28.8 |
| Education medium (%) | 51.1 | 50.0 |
| Education high (%) | 39.8 | 48.9 |
| Employment 1: gainfully employed (%) | 45.0 | 49.8 |
| Employment 2: self-employed (%) | 9.4 | 29.1 |
| Employment 3: unemployed, retired, other (%) | 45.6 | 49.8 |
| Born in Germany (%) | 92.4 | 26.5 |
| | | |
| ***Literacy Questions (n=1,735)*** | | |
| FL score 1 (number of correct answers) | 2.57 | 0.705 |
| FL score 2 (all correct) (%) | 67.6 | 46.8 |
| FL: Q1 Interest rate correct (%) | 87.7 | 32.8 |
| FL: Q2 Effects of Inflation correct (%) | 91.9 | 27.2 |
| FL: Q3 Diversification correct (%) | 77.0 | 42.1 |
| | | |
| **Self-Assessments and Expectations** | | |
| Satisfaction with life (n=1,701) | 7.19 | 2.04 |
| Self-assessment: risk (n=1,703) | 3.77 | 2.27 |
| Self-assessment: patience (n=1,703) | 4.68 | 2.53 |
| Visit to religious service (n=1,699) | 2.72 | 1.03 |
| Religious service: 1- Regularly (%) | 13.1 | 33.8 |
| Religious service: 2 – every now and then (%) | 32.3 | 46.8 |
| Religious service: 3 - For feast days and religious festivals (%) | 24.3 | 42.9 |
| Religious service: 4 – Never (%) | 30.3 | 45.9 |
| Price Expectations (n=1,699) | 1.70 | 0.633 |
| 1- strong increase | 38.4 | 48.7 |
| 2- increase | 54.2 | 49.8 |
| 3 – unchanged | 6.5 | 24.6 |
| 4-drecrease | 0.65 | 8.02 |
| 5- strong decrease | 0.18 | 4.20 |

Source: PHF 2010/2011 – Part 1 – unweighted.

**Table A4 – Description of Individual-, Household-, and Interviewer-Level Control Variables**

| Variable | Scale | Construction |
|---|---|---|
| *Reference Person level* | | |
| Born in Germany | Dummy | 1, if reference person was born in Germany |
| Female | Dummy | 1, if reference person is female |
| Age | Categorical | 5 categories of age of the reference person at the time of the interview: below 35, 35-44, 45-54, 54-64, 65 and above |
| Employment | Categorical | 3 categories for the level of education of the reference person at the time of interview: 1 gainfully employed, 2 self-employed, 3 other (includes unemployed and retired) |
| Education | | 3 categories for the education level of the reference person at the time of interview: 1- low, 2-medium, 3-high Only the level of schooling is considered, not the level of professional training. |
| *Household level* | | |
| Gross household income | Quintiles | Quintiles of gross annual household income, calculated based on responses to individual income questions from all members of the household (16+) and household level income questions |
| Household Size | Ordered | Number of household members at the time of interview, includes children. Top coded at 4. |
| Stratum indicator | Categorical | 1- "poor" small municipality (<100 000 inhabitants), 2 – "wealthy" small municipality (<100 000 inhabitants), 3 – wealthy street section in large municipality (>100 000 inhabitants), 4 – poor street section in large municipality (>100 000 inhabitants) |
| *Interviewer level* | | |
| INT: female | Dummy | One, if interviewer is female |
| INT: age | Categorical | 3 categories of age of the interviewer at the start of the field phase: below 44, 45-64, 65 and above |
| INT: Education | | 3 categories for the education level of the reference person at the time of interview: 1- low, 2-medium, 3-high Only the level of schooling is considered, not the level of professional training. |
| INT: Percentage of DK/NA responses to the total number of questions | Percentage | Percentage of don't know and no answer responses for questions asked to a respondent. Averaged over all interviews of one interviewer in part 1 and 2 of the survey |
| INT: Total number of contact attempts | Continuous | Total number of in-person or telephone contact attempt in part 1 and 2 of the survey. |
| INT: Share of interviews in number of addresses received | Percentage | Cooperation Rate at the interview level per address issued to the interviewer in part 1 and 2 of the survey |
| INT: interview-time in seconds per item | Continuous | Average time in seconds the respondent took to answer a "question". Calculated as the total duration of an interview divided by the total number of valid and don't know/ no answer answers. In multiple choice questions each option is one item. Averaged for each interviewer over all interviews conducted by the interviewer and part 1 and 2. |

**Table A5 – Description of Variables used in Substantive Regressions and Variables for which ICCs are Calculated**

| Variable | Scale | Construction |
|---|---|---|
| *Variables used in substantive regressions* | | |
| Has mutual fund | Dummy | One if at least one household member owns a mutual fund. This includes mutual funds investing in stocks, bonds or other securities as well as money market funds or real estate funds. Private pension contracts are not considered. Data not imputed. |
| Has bonds | Dummy | One if at least one household member directly owns bonds. Data not imputed. |
| Has shares | Dummy | One if at least one household member directly owns traded stocks. Untraded shares are not considered. Data not imputed. |
| Has private pension | Dummy | One if the reference person owns a private pension contract. This includes subsidised (Riester/Rürup) and unsubsidised contracts of any type. Data not imputed. |
| *Variables for which ICC are calculated (see Table 1)* | | |
| FL score 1 | Ordered | Sum of correct answers to the three literacy questions, DK/NA treated as wrong answer |
| FL score 2 | Dummy | One, if all three literacy questions are answered correctly |
| Q1 Interest rate | Dummy | One, if correct answer provided for “Let us assume that you have a balance of €100 on your savings account. This balance bears interest at a rate of 2% per year and you leave it for 5 years on this account. How high do you think your balance will be after 5 years? 1 - More than €102, 2 - Exactly €102,<br>3 - Less than €102” |
| Q2 Effect of inflation | Dummy | One, if correct answer provided for “Let us assume that your savings account bears interest at a rate of 1% per year and the rate of inflation is 2% per year. Do you think that in one year's time the balance on your savings account will buy the same as, more than or less than today? 1 – More, 2 - The same, 3 - Less than today” |
| Q3 Diversification | Dummy | One, if correct answer provided for “Do you agree with the following statement: "Investing in shares of one company is less risky than investing in a fund containing shares of similar companies"? 1- Agree, 2 -Disagree”. |
| Has saving account | Dummy | One, if at least one household member has a savings account (includes “Bausparverträge”, but excludes private pension contracts). Data not imputed. |
| Price expectations | Ordered | Respondent’s expectations regarding the development of “prices in general” over the next 12 months: 1- strong increase, 2- increase, 3 – unchanged, 4-drecrease, 5- strong decrease. Data not imputed. |
| Has mortgage | Dummy | One, if at least one household member has a mortgage on the household’s main residence. Missing, if household |

| | | |
|---|---|---|
| | | doesn't own the household main residence. Data not imputed. |
| Has credit card | Dummy | One, if at least one household member owns a credit card. Data not imputed. |
| Has cars | Dummy | One, if at least one household member owns a car (excludes cars used for business purposes). Data not imputed. |
| Discretionary Saving | Dummy | One, if respondent report to have saved some money on an "ad-hoc-basis" during the last 12 months prior to the interview. Data not imputed. |
| Satisfaction with life | Ordered | 11-point Likert scale response to question "How satisfied are you overall with your life at present?" (0- totally dissatisfied to 10 - entirely satisfied). Data not imputed. |
| Self-assessment: risk | Ordered | 11-point Likert scale response to question "Are you in general a risk-taking person or do you try to avoid risks?" (0 - "not at all ready to take risks" to 10 - "very willing to take risks"). Data not imputed. |
| Self-assessment: patience | Ordered | 11-point Likert scale (0 -very patient to 10 - very impatient) response to question "Are you in general a person who is patient or do you tend to be impatient?". Data not imputed. |
| Visit to religious service | Ordered | Response to "How often do you go to church, the synagogue or the mosque? 1- Regularly, 2 – every now and then, when the occasion arises, 3 - For feast days and religious festivals, 4 – Never. Data not imputed. |