



EUROPEAN CENTRAL BANK  
EUROSYSTEM

## Working Paper Series

Florian Deuffhard, Dimitris Georgarakos and  
Roman Inderst

Financial literacy and savings  
account returns



No 1852 / September 2015



**Note:** This Working Paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the authors and do not necessarily reflect those of the ECB

## **Abstract**

Savings accounts are owned by most households, but little is known about the performance of households' investments. We create a unique dataset by matching information on individual savings accounts from the DNB Household Survey with market data on account-specific interest rates and characteristics. We document heterogeneity in returns across households, which can be partly explained by financial sophistication. A one-standard deviation increase in financial literacy is associated with a 12% increase compared to the median interest rate. We isolate the usage of modern technology (online accounts) as one channel through which financial literacy has a positive association with returns.

*Keywords: Financial literacy; savings accounts; interest rates; household finance*

*JEL Classification: D12, E21, G11, G21*

## Non Technical Summary

Savings accounts are owned by most households, but little is known about the performance of households' investments. We create a unique dataset by matching information on individual savings accounts from the DNB Household Survey (DHS) with market data on account-specific interest rates and characteristics. In particular, we make use of the fact that the DHS reports bank and account *names* for each savings account owned by a household member, as well as the respective invested amount. This information allows us to match the DHS with market data on interest rates and other account characteristics. We can thereby calculate for each household a measure of the average return that is earned across all savings accounts owned by its members.

We document considerable heterogeneity in returns across households for such a widely held and virtually riskless asset. To understand such a difference in performance of what seems to be a relatively simple financial product, our study first points to characteristics of the market and products. There is a wide dispersion of interest rates across products even for the same invested amount. A comparison of individual products is also not straightforward, e.g., as accounts differ in the applicable amount thresholds to earn a higher interest rate as well as in additional restrictions. The difference in account characteristics, for which we can control, and the variety of offers in the market suggest, in particular, a role for financial sophistication as an explanation for the observed heterogeneity in returns.

Our paper is the first, to our knowledge, to show that heterogeneity in returns of a widely held asset such as savings accounts is partly linked to investor financial literacy. We recover measures of financial literacy from a special module of questions that was part of the 2005 wave of the DHS. Even after accounting for a range of socio-economic characteristics, account characteristics, as well as amount invested, we find that financial literacy has a significant relationship with households' individual returns on savings accounts: a one-standard deviation higher advanced financial literacy is associated with an approximately 29 basis points higher interest rate, which represents an increase of 12% compared to the median interest rate of 2.5%. We also calculate the gains from moving a household in the lowest literacy quartile to the highest literacy quartile. Applying the estimated gains of literacy to the average savings volume and projecting this over 10 years, total gains in real terms would accumulate to €838.

Our investigation of products and the market suggests that lack of information may prevent households from securing the highest possible interest rate for the invested amount. Even at a given bank, households may not choose the most preferable offer. In fact, one such channel that we can isolate is the ability and willingness (or the lack of it) to use a higher interest bearing online account. We also find some evidence to suggest that more literate households might be better able to identify accounts across banks that for a given volume and a given set of characteristics offer the highest return.

## 1. Introduction

Savings accounts typically represent the most common vehicle for household financial investment. In the DNB Household Survey (DHS) savings accounts are owned by 82% of all Dutch households and make up the largest part of their financial wealth (with an average share of 43%).<sup>1</sup> This contrasts with much lower ownership rates of funds or directly held stocks.<sup>2</sup> Still, while there exists a large literature documenting how households invest in funds and stocks and how these investments perform, much less is known about savings accounts.

We make use of the fact that the DHS reports bank and account *names* for each savings account owned by a household member, as well as the respective invested amount. This information allows us to match individual accounts held by households in the DHS with market data on interest rates and a set of account characteristics.

We document considerable heterogeneity in returns across households for such a widely held and virtually riskless asset. To understand such a difference in performance of what seems to be a relatively simple financial product, our study first points to characteristics of the market and products. There is a wide dispersion of interest rates across products even for the same invested amount. A comparison of individual products is also not straightforward, e.g., as accounts differ in the applicable amount thresholds to earn a higher interest rate as well as in additional restrictions. Notably, this variation is not due to so-called “teaser rates” that are paid when an account is newly opened or when fresh money is transferred, as these rates are not

<sup>1</sup> The picture is similar for most other Euro area countries according to the recent data from the Household Finance and Consumption Survey (see: [http://www.ecb.europa.eu/home/html/researcher\\_hfcn.en.html](http://www.ecb.europa.eu/home/html/researcher_hfcn.en.html)).

<sup>2</sup> For a comparison, ownership rates (average shares) are 20% (6%) for funds and only 12% (3%) for directly held stocks.

considered in the analysis. The difference in account characteristics, for which we can control, and the variety of offers in the market suggest, in particular, a role for financial sophistication as an explanation for the observed heterogeneity in returns.

This paper is the first, to our knowledge, to show that heterogeneity in returns of a widely held asset such as savings accounts is partly linked to investor financial literacy. We recover measures of financial literacy from a special module of questions that was part of the 2005 wave of the DHS.<sup>3</sup> Even after accounting for a range of socio-economic characteristics, account characteristics, as well as amount invested, we find that financial literacy has a significant relationship with households' individual returns on savings accounts: a one-standard deviation higher advanced financial literacy is associated with an approximately 29 basis points higher interest rate, which represents an increase of 12% compared to the median interest rate of 2.5%. We also calculate the gains from moving a household in the lowest literacy quartile to the highest literacy quartile. Applying the estimated gains of literacy to the average savings volume and projecting this over 10 years, total gains in real terms would accumulate to €838.

Our investigation of products and the market suggests that lack of information may prevent households from securing the highest possible interest rate for the invested amount.<sup>4</sup> Even at a given bank, households may not choose the most preferable offer. In fact, one such channel that we can isolate is the ability and willingness (or the lack of it) to use a higher interest bearing online account. We also find some evidence to suggest that more literate households might be better able to identify accounts across banks that for a given volume and a given set of

<sup>3</sup> These are the same questions as used in van Rooij et al. (2011, 2012).

<sup>4</sup> Several studies cite information/search frictions as a source of price dispersion in retail financial markets net of product differentiation by firms. See, e.g., Hortacsu and Syverson (2004) for S&P 500 index funds and Stango and Zinman (2013) for credit cards. As a result, firms might have an incentive to add complexity to their pricing structures in order to gain market power (Carlin 2009).

characteristics offer the highest return. From banks' perspective, lack of knowledge and sophistication are in fact prerequisites to uphold price dispersion across banks as well as price discrimination across accounts.

A common feature of existing studies on households' investment decisions and on financial literacy, as reviewed below, is the difficulty to both isolate the contribution of financial literacy, which requires specific survey questions, and measure asset returns at the same time. Our combined data contain both pieces of information, allowing us to assess the association of financial literacy with savings returns. Importantly, note that the starting point for our matching process of survey data and administrative market data is a nationally representative survey that contains detailed information on all savings accounts held by household members (some of which are held in different banks), as well as on all other financial assets. Moreover, given that our outcome of interest is the applicable interest rate obtained from administrative market data, it is less likely to correlate with literacy through household unobservables such as knowledge about realized returns or reporting bias. In any case, we take a number of steps in response to endogeneity concerns including a standard IV approach and an alternative identification method recently introduced by Lewbel (2012) that exploits information from the heteroscedastic structure of the data.

A number of studies document significant variation in households' financial literacy in various countries.<sup>5</sup> As savings accounts arguably play an important role in other countries as well, we would suggest that our results are likely to be more widely applicable. In fact, to the

<sup>5</sup> See, for instance, Lusardi and Mitchell (2014) for a recent review of comparable studies in the US as well as Europe, Australia, and Japan. Earlier studies include Bernheim (1998) and Hilgert et al. (2003).

extent that savings accounts represent the most important financial assets, also the respective welfare implications should concern a large fraction of society in many countries.

As noted above, much of the extant literature on investments has focused on the holdings of stocks and other risky assets. In particular, using the same survey, van Rooij et al. (2011) and van Rooij et al. (2012) find that financial literacy induces stockholding and boosts wealth accumulation, respectively.<sup>6</sup> Finally, when households earn higher returns on their investments, this provides another explanation, next to differences in savings rates, for differences in retirement savings, which have been explored widely (e.g., Lusardi and Mitchell 2007a, 2007b, 2008; van Rooij et al. 2012).

The rest of the paper is organized as follows. Section 2 presents the data and matching procedure. Section 3 introduces the empirical specifications to uncover the link between financial literacy and returns from savings accounts. Section 4 presents the empirical results, robustness checks, and evaluates implications for consumer welfare. Section 5 concludes.

## **2. Data**

### *2.1. Household Characteristics and the Use of Savings Accounts*

Our main data source is the DNB Household Survey (DHS) in 2005. The DHS is an annually conducted survey of around 2,000 Dutch households containing extensive information on demographic characteristics, asset and debt holdings, housing, work, health and income, as well

<sup>6</sup> Related work shows that cognitive skills such as numeracy (Christelis et al. 2010) and IQ (Grinblatt et al. 2011) positively associate with stockholding (see also Yoong 2011; Arrondel et al. 2012). Moreover, Banks et al. (2010) find that more numerate individuals save more pre- and dissave more post-retirement. Calvet et al. (2009) construct instead a proxy of financial sophistication based on the relationship between households' financial mistakes and education, income, and wealth. Other studies have analyzed the role of literacy for the choice of debt products (e.g., Lusardi and Tufano 2009, Stango and Zinman 2009). For an overview, see Campbell (2006) and Guiso and Sodini (2013). Moreover, several studies have documented how investment mistakes correlate with proxies for financial knowledge such as education (e.g., Calvet et al. 2007; Biliias et al. 2010).



as economic and psychological concepts. Variables used from the survey are reported below. The survey is representative of the Dutch population and is conducted via the Internet.<sup>7</sup> One key feature of the survey is that it asks detailed information on all savings accounts held by a household, including bank and account name, as well as invested volume on each account.<sup>8</sup> The DHS asks to report invested amounts for each financial asset as of December 31<sup>st</sup> of the year preceding the interviews.

We supplement the DHS data with information from a special module on financial literacy designed by van Rooij et al. (2011) and conducted over a random sub-sample of the 2005 survey. This module contains a series of questions about financial knowledge addressed to the person in charge of household finances.<sup>9</sup> Questions from this module have been used to construct an index of basic and an index of advanced financial literacy.<sup>10</sup> These indices are derived by factor analysis and are normalized to mean zero and standard deviation one (cf. van Rooij et al. 2011). Table 1 presents summary statistics of the main household-level variables including demographics, financial literacy, income, and wealth for the sample later used in the regression analysis.

Households exhibit considerable heterogeneity in more advanced financial knowledge. Instead, basic literacy does not vary over a significant part of the sample, given that 43% of

<sup>7</sup> We use survey weights to make reported statistics representative of the Dutch population. The survey provides equipment to households without Internet access in order to compensate for this form of bias. See Teppa and Vis (2012) for a detailed description of the DHS.

<sup>8</sup> In the regular panel, participants are provided with a list of seven possible answers when asked at which bank they hold each of their savings accounts: ABN Amro, Postbank, Rabobank, ING, Fortis, SNS Bank, and 'Other'. In case participants indicate ownership in the category 'Other', they are further asked to provide the name of the bank. This latter information along with account names is not available in the public version of the dataset, but has been recovered from additional data that were made available to us by CentERdata. Appendix C provides more details.

<sup>9</sup> Smith et al. (2010) have shown that this person is actually the most influential for households' financial decisions. The remaining socio-demographic characteristics that we take into account refer also to this person.

<sup>10</sup> See Appendix D for the exact wording of these questions.

households therein manage to answer all basic literacy questions correctly. Following earlier work using information from the same financial literacy module, we thus use the index of advanced financial literacy as our baseline measure of financial knowledge.<sup>11</sup> As an alternative to this measure we also construct a measure based on correct responses to the ‘Big Three’-questions on financial literacy (see Hastings et al., 2013). This measure draws on three standard questions regarding interest compounding, inflation and risk diversification and has been used in a number of studies examining the role of financial knowledge for financial outcomes. Results on this alternative measure along with different functional forms of the advanced literacy index are presented in the robustness section.

In the DHS, after checking accounts, which are owned by virtually all households, savings accounts represent the second most prevalent financial asset with an ownership rate of 82%. For comparison, only 20% invest in funds and only 12% hold stocks directly. On average, households invest 43% of their financial wealth in savings accounts and hold 21% in checking accounts. Apart from insurances, which account for 12%, all other financial assets have a far lower weight in household portfolios. Thus, in terms of both ownership and financial wealth invested, savings accounts are by far the most important financial asset for Dutch households.

## *2.2. Interest Rate Data on Savings Accounts*

We use data on annual interest rates for savings accounts of all Dutch banks from April 2004 to December 2004 provided by a major Dutch financial institution.<sup>12</sup> The data set covers in total 43

<sup>11</sup> As van Rooij et al. (2011) point out, the basic literacy questions in the DHS special module test for basic numerical skills and are thus more likely to proxy for cognitive abilities that typically depreciate at advanced ages. Our estimates for the advanced literacy index are insensitive to the inclusion of the basic literacy index in the estimated model.

<sup>12</sup> The data contain information on standard interest rates (i.e., not “teaser” ones offered by banks over a short period

banks and 105 savings accounts. For each savings account, it contains the account name, the bank name, and the weekly interest rate for eleven different amount brackets ranging from €0 - €1,000 to €45,000 or more.<sup>13</sup> In addition, using information from the Dutch Internet comparison website ‘SpaarInformatie’, we supplement our data with information on various savings account restrictions, which we use as controls in our empirical specification below.

These account types can be roughly partitioned across two dimensions. First, accounts are either restricted or not. The information from the comparison website allows us to distinguish in total six main restrictions.<sup>14</sup> These restrictions are not exclusive but can coincide for one account. Second, accounts are either internet managed or not. Internet accounts are fully managed online by the depositor and provide very limited face-to-face customer services. In addition, they can be both restricted and unrestricted. In light of the research question of this paper, we thus treat online accounts separately from the remaining set of account restrictions. We relegate an overview of the various characteristics of individual savings accounts to Table A1 in the Appendix.

Table 2 provides summary statistics for the distribution of interest rates across different amount brackets. Statistics in Panel A refer to all accounts offered in the market based on the administrative interest data alone. Statistics in Panel B refer only to accounts actually held by

in order to attract new customers). April 2004 is the first month of the administrative data that we have access to.

<sup>13</sup> The exact amount brackets are €0 - €1,000, €1,000 - €2,500, €2,500 - €3,500, €3,500 - €4,500, €4,500 - €7,000, €7,000 - €8,000, €8,000 - €9,000, €9,000 - €10,000, €10,000 - €25,000, €25,000 - €45,000 and > €45,000.

<sup>14</sup> (1) Accounts with minimum amount requirements offer either very low base rates or zero interest rate up to a certain volume threshold and higher rates above that threshold. (2) Accounts with lowest balance bonus give a bonus rate on the lowest account balance within a year or a quarter and yield a base rate on the remaining balance. (3) Accounts with balance growth bonus yield a bonus rate if the balance grows by a specified percentage amount per quarter or year. (4) Accounts with fixed monthly deposit require a specified absolute deposit automatically withdrawn from the checking account of the consumer each month. (5) Accounts with withdrawal limitations / fees limit the maximum amount that can be withdrawn per month or impose percentage fees for withdrawals (in most cases 1% of the withdrawn amount). (6) Salary accounts are linked to a checking account at the same bank.

households in the survey (as of December 2004).<sup>15</sup> As could be expected, accounts typically pay higher rates for larger volumes. Even for a given volume, dispersion is quite high. For example, for savings accounts actually held by survey respondents, interest rates for volumes from €2,500 to €3,500 range from 1.00% to 4.00% with an interquartile range of 1.45%. Yet, the interquartile range reduces to 1.00% for volumes above €45,000.

### *2.3. Data Matching Procedure*

Interest rate data are matched with DHS data as follows. Given the availability of literacy data in the 2005 wave, which reports the holdings of financial assets as of December 31, 2004, we match interest rates for the last week of December 2004 to the DHS data based on bank and account name as well as account volume.<sup>16</sup> Precisely, based on the volume invested by households in each of their individual accounts, we can assign the respective interest rate for the applicable volume bracket.<sup>17</sup> We achieve a full match for 79% of all accounts held by households in the DHS.<sup>18</sup> For each savings account held by a member of a household, our matched data ultimately contain the invested volume, account name, bank name, and the applicable interest rate.<sup>19</sup>

<sup>15</sup> Thus, the unit of observation in Panel A is an individual account, whereas the unit of observation in Panel B is an individual account held by a household in the survey.

<sup>16</sup> Our results are robust when we use, instead, a (geometrically) weighted interest rate for each account over all weeks in 2004. Precisely, for 2004 we can use interest data from April 2004 to December 2004. Interest rate changes are relatively infrequent in this period.

<sup>17</sup> We recover missing volumes of individual savings accounts following the procedure used by CentERdata for total savings volumes as described in Appendix C.

<sup>18</sup> We do not find evidence that the sample used in the estimation differs in a systematic way from the entire sample of account owners. We provide a comparison of main demographics and bank characteristics for the two samples in Table A2.

<sup>19</sup> In what follows, we present results only from the sample of those accounts with matched interest rate information. We have also imputed missing interest rates utilizing information on the bank names reported by each household. Results from the sample that incorporates these imputed cases along with details on the imputation procedure can be found in Deuflhard et al. (2013). Results from both samples are highly comparable.

For the final estimation sample, we exclude accounts with very low volumes (i.e., below €50), which are quite likely to be inactive.<sup>20</sup> Table 3 shows summary statistics of the account-level APR over the sample used in the estimation and by various socio-economic and account characteristics as well as financial literacy. The mean is 2.31% and the median is 2.50%. Dispersion is quite high given an interquartile range of 1.55% (i.e., 155 basis points). The APR increases considerably with invested volume as well as advanced financial literacy, and decreases in age, while there is no strong association with education and net income.

### **3. Econometric Specification**

The preceding description of the market for savings accounts suggests various channels through which households can fail to obtain the highest possible return on their savings account(s). First, while banks offer different interest rates even for accounts with similar characteristics, lack of information may prevent households from shopping successfully for the highest interest account. Second, even at a given bank, households may not choose the most preferable account for the amount that they save. Finally, even for a given set of own savings accounts, households may fail to allocate their savings to the highest interest account, potentially foregoing higher interest for larger volumes. While we cannot completely disentangle these different channels, we provide some evidence for their relative importance in Section 4.

Our main aim is to provide an estimate of the relationship between financial literacy and savings account returns. To that effect, we estimate the following, account-level specification:<sup>21</sup>

<sup>20</sup> This concerns around 6.8% of the accounts in the account-level sample.

<sup>21</sup> In the welfare section, we also discuss results from a household-level specification. This is possible to estimate by calculating a volume-weighted measure of APR per household and aggregating individual account characteristics at the household level. The results we find are similar.

$$r_{hs} = \beta_1 FinLit_h + \beta_2 X_h + \beta_3 V_{hs} + \beta_4 Z_s + \varepsilon_{hs}, \quad (1)$$

where  $r_{hs}$  represents the interest rate earned on account  $s$  held by household  $h$ .  $FinLit_h$  denotes the advanced financial literacy index of household  $h$  (i.e., the covariate of interest). The vector  $X_h$  contains a set of household demographics including age, gender, marital status, and the number of children as well as occupation status. Furthermore, we include region dummies to take into account any relevant regional disparities, e.g., in density of bank branches or in local employment conditions. In addition, we take into account nine dummies, contained in  $V_{hs}$ , which take the value one if the account volume falls into one of the previously mentioned amount brackets over which interest rates can vary and are zero otherwise.<sup>22</sup> We also include a set of dummies denoting various account restrictions and bank fixed effects specific to each account in  $Z_s$ .

When estimating the baseline specification in Eq. (1) one should take into account the potential endogeneity of financial literacy. This has been a common empirical challenge for studies using survey data to examine the effect of literacy on various economic outcomes. In our set-up, it should be noted that the outcome of interest is the applicable interest rate that is obtained from administrative market data. Thus, it is less likely to correlate with literacy through household-specific unobserved factors such as knowledge about realized returns or reporting bias. Nevertheless, measurement error in the advanced financial literacy index remains a valid

<sup>22</sup> As in Table 2, we group together three amount brackets from €7,000 to €10,000 due to too few observations in these categories and no account reaching a new volume threshold within this range.

concern, given that some of the correct responses are likely to result from guessing (cf. van Rooij et al. 2011), in which case our estimated effect of literacy from OLS will be biased towards zero.

We take a number of steps in order to address the measurement error issue and potential endogeneity concerns. First, we use a standard instrumental variable approach. Second, we utilize an alternative identification approach introduced by Lewbel (2012) that generates instruments using heteroscedasticity in the error structure of a first stage regression. Third, we estimate a specification that is more resilient to the measurement error of literacy. In what follows, we provide details on each of these steps.

In our first approach, we employ the instruments from two earlier studies using the same financial literacy index and data.<sup>23</sup> A valid instrument should exhibit meaningful correlation with advanced financial literacy and affect the interest rate only through the literacy channel and not through other unobserved factors. Building on van Rooij et al. (2011), we use the financial condition of the oldest sibling as an instrument for advanced financial literacy.<sup>24</sup> The financial condition of the oldest sibling is beyond a respondent's immediate control and can thus be seen as relatively exogenous with respect to the savings account choice. Moreover, the authors argue in favor of a learning channel according to which respondents tend to become more interested in learning about financial matters due to the negative financial condition of their siblings. If such a mechanism is at work, one should observe a higher literacy score (on average) among respondents who report their siblings being in worse financial situation than respondents who do not. Results from our first stage regressions show a positive association between the literacy

<sup>23</sup> Summary statistics on both instruments used can be found in Table 1.

<sup>24</sup> Respondents were asked to indicate whether the financial situation of the oldest sibling is better, the same or worse compared to their own financial situation.

score and having siblings in bad financial shape, thus providing support for such a learning channel.<sup>25</sup>

In addition, following van Rooij et al. (2012), we use as a second instrument the economics education of the respondent.<sup>26</sup> Economics education at an early stage is expected to positively affect financial literacy but is also likely to determine a household's current economic situation. We have experimented with specifications that control for contemporaneous household resources (e.g., net income, net financial and net real wealth) in order to take into account a possible channel through which past economics education can influence current investment choices. As we discuss below, our IV estimates are quite comparable across both a parsimonious specification that conditions only on advanced literacy and some very rich ones that take also into account numerous account and household characteristics. The fact that our IV approach does not depend on the various controls included in Eq. (1) provides some indirect support for the exogeneity of our instruments.

Our second approach uses the method recently introduced in Lewbel (2012) and does not rely on the validity of the instruments employed in standard IV. Instead, Lewbel proposes to exploit variation on higher moment conditions of the error distribution from a first stage regression of the likely endogenous covariate on (a sub-set of) other covariates in the model. The method generates a set of instruments that can be used for identification under two assumptions:

(a) The errors from a first stage regression of the endogenous covariate on a (sub-set) of other covariates in the model should be heteroscedastic. In our context:

<sup>25</sup> If the financial condition of the oldest sibling proxies instead for a common set of preferences or a family fixed effect, one would expect a negative correlation between the instrument and financial literacy in the first stage regression.

<sup>26</sup> Specifically, respondents were asked how much of their past education was devoted to economics (i.e., 'a lot', 'some', 'little' and 'hardly at all').



$$FinLit_h = \gamma_2 X'_h + \omega_{hs}, \quad (2)$$

where  $\omega_{hs}$  denotes the error term and  $X'_h$  is a subset of the RHS variables in Eq. (1) including a constant. Natural candidates for  $X'_h$  are variables that are pre-determined relative to the outcomes. We use age, gender and family size indicators. Heteroscedasticity in Eq. (2) implies that  $Cov(X'_h, \omega_{hs}^2) \neq 0$ . This assumption can be tested on the basis of a Breusch-Pagan test for heteroscedasticity and is strongly supported in our data.

(b)  $X'_h$  is assumed to satisfy  $Cov(X'_h, \varepsilon_{hs} \omega_{hs}) = 0$ , where  $\varepsilon_{hs}$  are errors from Eq. (1). This condition holds even when the two error terms share a common unobserved factor component (and are thus correlated), as long as the product of their idiosyncratic error components is uncorrelated with  $X'_h$ .<sup>27</sup> Given that the method generates a number of instruments, one can test for their joint validity using a standard over-identification test, under the assumption that one of them is pre-determined.

Under conditions (a) and (b), Lewbel shows that a set of valid instruments for estimating Eq. (1) can be generated as:  $(X'_h - \overline{X'_h}) \widehat{\omega}_{hs}$ , where  $\overline{X'_h}$  is the mean of  $X'_h$  and  $\widehat{\omega}_{hs}$  are estimated residuals from Eq. (2). The advantage of using this method in our context is twofold: first, it allows to compare our estimates from the standard IV approach with those obtained using the generated instruments from the Lewbel method; second, it makes possible to test for the validity of both external instruments used under the standard IV approach.

<sup>27</sup> This can be seen as a relatively mild assumption compared to the exclusion restriction required under standard IV. For example, in our context it allows for unobserved factors such as general well-being and ability (or intention) to learn about finances to affect both interest rate earned and financial literacy.

Finally, we have estimated a specification that controls for literacy via dummies denoting quartiles of the underlying distribution. Such a specification is likely to be more robust to measurement error, compared to the baseline specification using a continuous variable, as it is resilient to measurement error within each quartile. As we show in the robustness section, the implied effects from the specification that conditions on literacy quartiles are quite comparable to the IV estimates from the baseline model that uses a continuous literacy indicator.

## 4. Results

### 4.1. Baseline Results on Financial Literacy

In what follows, we first discuss the results on financial literacy followed by other covariates and a number of robustness checks we have performed. Subsequently, we discuss the role of online account usage. Finally, we evaluate possible implications for household welfare.

Table 4 presents results from the account-level regressions as in Eq. (1). Given that financial literacy and other background characteristics, used as controls in this specification, do not vary across accounts owned by the same household, we cluster standard errors at the household level.<sup>28</sup>

First, we present results from a parsimonious specification, OLS (1), that conditions only on advanced financial literacy. In the second specification, OLS (2), we consider in addition socio-economic characteristics and account volume dummies, while in the third specification, OLS (3), we add as well account characteristics and bank fixed effects. Next to each of these

<sup>28</sup> Please note that the number of observations slightly changes from the OLS to the IV specifications due to some missing observations in the used instruments.

three OLS specifications, we show results from their IV counterparts (i.e., IV (1), IV (2) and IV (3)).

In all IV specifications, the F-statistics from the first stage regressions are above or slightly below 10 and the two instruments exhibit meaningful correlations with the advanced literacy index (results from the first stage regressions are shown in Table B1, Appendix B). Given that we employ two instruments for one potentially endogenous covariate, one can test for their statistical validity on the basis of a test for over-identifying restrictions. According to the Hansen J-test (reported at the bottom of the table), we fail to reject the null hypothesis that the instruments are jointly valid (p-values: .29, .60 and .45).

Adding account and bank fixed effects into the third specification improves considerably, as expected, the fit of the model. In all three OLS specifications, the coefficient of advanced financial literacy is statistically significant (p-value < .01) and shows a positive association with the APR. The corresponding IV estimates remain statistically significant and suggest a slightly stronger relationship.<sup>29</sup> Notably, the estimated magnitudes are more or less unaffected across all three specifications (i.e., the IV strategy works irrespective of the set of other covariates taken into account).

According to the IV estimates, an assumed one-standard deviation increase in advanced financial literacy implies a roughly 29 basis points increase in the APR. This effect, estimated net of socio-economic characteristics, account restrictions and bank fixed effects, is non-trivial as it corresponds to 12% of the median interest rate in our sample.

<sup>29</sup> Lusardi and Mitchell (2014) review a number of studies that use different instruments to estimate the effect of literacy in various set-ups and point to the fact that the derived IV estimates are typically well above the OLS ones. As they argue, this is likely to be the case due to severe measurement error of the literacy indicators.

As an alternative to the standard IV approach used above, we also apply the identification method of Lewbel (2012) that, as discussed in Section 3, exploits variation from the second moments of the error distribution of the first stage regression in Eq. (2) to generate a set of instruments. For this,  $X'_h$  comprises few pre-determined covariates, namely age, gender and number of children. First, we estimate the first-stage regression in Eq. (2) and test for heteroscedasticity using a Breusch-Pagan test. According to the test results ( $\chi^2=71.1$ ,  $p$ -value=0.00) there is strong evidence for heteroscedasticity in the first stage regression. Following Lewbel, we generate instruments by taking the products of residuals from Eq. (2) with each of the aforementioned covariates, centered at their respective sample means. These generated instruments can be subsequently used either alone or in conjunction with the two external instruments used under the standard IV approach in order to identify Eq. (1).

Table 5 summarizes the relevant results. In particular, IV (1) uses generated instruments from the Lewbel method only, while IV (2) uses the two external instruments employed in the standard IV approach alone. IV (3) uses the two external instruments supplemented with the generated instruments from the Lewbel method resulting in more efficient estimates than in the standard IV specification. Notably, the estimated coefficients on financial literacy suggest qualitatively similar effects and are statistically significant across all three specifications. Moreover, the generated instruments from the Lewbel method meet the exogeneity assumption as Hansen's J-statistic fails to reject the null of exogeneity with high confidence ( $p$ -value: .62).

It is possible to use the instruments from the Lewbel method to test for the joint validity of the instruments employed originally under the standard IV specification. This test is based on the difference in Hansen's J-statistics between the model using the generated instruments according to the Lewbel method only and the full model using the entire set of generated and external

instruments. According to the resulting C-statistic of 7.2 (p-value: .21), one cannot reject the null hypothesis that the two external instruments employed in the standard IV approach are jointly valid. Results from this test lend some further support to the validity of the originally employed instruments.

#### *4.2. Results on other covariates*

With reference to other covariates in the model, the account volume dummies show, as expected, a progressively stronger association with a higher interest rate, consistent with the notion that the benefits from shopping are higher for investors with larger volumes. Given that we control for bank fixed effects and account characteristics, which are highly significant, these differences do not seem to be solely attributable to choices of accounts with more restrictive characteristics.<sup>30</sup>

In addition, we estimate a strong negative association of the APR with age. For example, respondents above sixty earn about 54 basis points less on average as compared to the base category of young adults below thirty. This likely suggests a significant role for age-of-account-effects given that the age of an account and respondents' age should be highly correlated.

Other covariates, such as education and gender (sometimes used as proxies for financial sophistication) and family size and employment status (that are likely to reflect liquidity needs) do not exhibit any significant association with the APR. In addition, we have estimated a richer specification controlling for household net income, net financial wealth (excluding savings accounts) and net real wealth through dummies denoting quartiles of the respective distributions. Notably, these additional controls of household resources are insignificant, while our baseline

<sup>30</sup> The estimated net effect of each account characteristic is hard to interpret, given that many of these restrictions typically co-exist.

estimates of financial literacy and account volume remain unaffected (literacy estimates are .107 and .313, both significant at 1%-level, under the OLS and IV specifications, respectively). This suggests that financial wealth and other household resources do not associate with the APR when we control for invested volume and financial literacy.<sup>31</sup>

### *4.3. Robustness Checks*

In this section, we discuss numerous checks that we have performed in order to verify the robustness of our baseline findings at the account level. Due to space constraints, Table 6 summarizes results from some of these robustness checks, while the entire set of results discussed below is available from the authors upon request.

Panel A shows results from several variations of Eq. (1). OLS (1) and IV (1) exclude volume dummies from the baseline specification, which are potentially endogenous. The derived estimates are highly comparable to the baseline with a financial literacy coefficient of .14 and .32 in the OLS and IV specification, respectively.

OLS (2) and IV (2) use only accounts from the financial respondent, for whom the financial literacy data is available. Recall that in our baseline estimation, we assign to each account held by any member of non-single households the financial literacy of the household's financial respondent. This might be problematic if household members differ significantly in their degree of literacy. Given that a significant fraction of accounts is held by the financial

<sup>31</sup>It is worth noting that most of the empirical household finance literature examines investment decisions in assets that are held by household sub-groups with selected characteristics (e.g., stockholding typically entails high participation and information costs and thus stocks are mostly held by wealthier, better educated and more financially literate investors). We examine instead returns from an asset that has low participation requirements and is held by the vast majority of households in the sample (82.4%). We have estimated a probit model of the probability of owning a savings account and most of the factors that were taken into account (including financial literacy) turn out to be insignificant. Financial wealth was estimated to have a strong positive association with savings account ownership, though such an association is likely to be mechanical.

respondent our sample reduces by only 16%. Our estimates of literacy remain highly significant at 1%-level with a comparable coefficient of .29 in the IV specification.

OLS (3) and IV (3) attach a higher weight to more important accounts by weighting each observation with its relative volume share within the household. OLS and IV estimates from these weighted regressions are .13 and .26, both significant at 1%-level, respectively. Thus, using volume weights leaves our main findings unaffected.

Panel B shows estimates using different financial literacy measures or functional forms. OLS (1) uses financial literacy quartiles instead of a continuous index. As discussed, using quartiles partly accounts for the measurement error in the financial literacy index used in the baseline specification.<sup>32</sup> Households in the top advanced literacy quartile earn on average 29 basis points more compared to the lowest literacy quartile.<sup>33</sup> This effect is highly comparable to the IV estimate from the baseline specification, given that the interquartile range of the literacy index equals roughly one standard deviation. This suggests that the financial literacy index may indeed suffer from measurement error that is taken into account by the standard IV estimation used for the continuous literacy index.

IV (1) uses the continuous financial literacy index but excludes “don’t know”-answers from the instrument denoting economics education. The estimated financial literacy coefficient is .26 (significant at 1%-level). This suggests that correlation between this instrument and the financial literacy index is not just due to a correlation with “don’t know”-responses in the economics education question.

<sup>32</sup> Such a specification is not free of measurement error as it is still subject to misclassification across the four possible quartiles.

<sup>33</sup> Obviously, we cannot easily instrument for advanced financial literacy when using quartiles due to the number of endogenous covariates.

OLS (2) and IV (2) use the standardized number of correct answers to the financial literacy questions instead of constructing an index based on factor analysis.<sup>34</sup> We estimate a financial literacy coefficient of .32 (p-value: <0.01) in the IV-specification, which is highly comparable to our baseline results. OLS (3) and IV (3) use the standardized number of correct answers to the ‘Big-Three’ financial literacy questions. As discussed, information from the three basic questions on interest compounding, inflation and risk diversification, has been collected by various household surveys that, unlike the 2005 DHS, do not have a special literacy module. As a result, this information has been used to measure literacy in a number of studies (see Hastings et al., 2013). Using this measure, we obtain an OLS estimate of 0.06 (p-value < 0.01) and an IV estimate of .43 (p-value: <0.05). According to the Hansen J-test the two instruments used are jointly valid, their F-statistic from the first-stage regression is nevertheless below 10.

We have also accounted for a number of factors that may influence the APR. Given that these additional controls have some missing values that reduce our estimation sample by about 15% to 20% in each case, we add one factor at a time.<sup>35</sup> In a first step, we include a measure of risk aversion from the DHS, as used in a similar robustness check by van Rooij et al. (2011).<sup>36</sup> The inclusion of risk aversion (that is itself insignificant) does not affect our estimate for advanced financial literacy. Second, while we control for employment status in our main specification, households frequently exposed to transitory income shocks might on average hold more liquid accounts with lower APRs. To this end, we include a dummy indicating whether households’ last year’s income was unusually low. The inclusion of this additional variable,

<sup>34</sup> We standardize these measures by their respective mean and standard deviation.

<sup>35</sup> See Appendix D for the exact wording of these questions.

<sup>36</sup> Based on two gambles presented to survey participants in the DHS, this measure can take five possible outcomes from low to high risk aversion (including one category for those who answered ‘don’t know’).



however, leaves our key estimate unaffected. Third, we also added hours worked to our specification to proxy for opportunity costs of shopping for higher rates. This variable has no significant impact on the APR and our estimates for literacy remain unaffected. Last, we include a categorical variable measuring the investment horizon of respondents to account for differences in patience of investors. Again, we find no significant effect of the investment horizon on the APR and our literacy estimate remains unaffected.

#### *4.4. Online Banking Usage*

One possible channel through which literacy could positively associate with APRs is through households' ability to choose the highest interest savings account within a given bank. As discussed in the data section, internet accounts are fully managed online with limited customer services and in return typically offer higher interest rates.

We re-estimate our baseline specification by adding a dummy denoting internet managed accounts.<sup>37</sup> Results are shown in Table 7. The internet account dummy displays a strong positive association with the APR. For example, after accounting for various account restrictions and bank fixed effects, the estimated impact of having an internet-managed account exceeds 130 basis points.

The implied effect of literacy is still statistically significant, albeit quantitatively smaller by around a half. This suggests that a sizable part of the effect of advanced financial literacy on the APR derives from familiarity with new technologies and the willingness and ability to use self-managed online banking.

<sup>37</sup> We obtain similar results when using self-reported online banking use, instead, which is asked in the DHS, as this is highly correlated with having an internet account.

As discussed in Section 3, literacy may also associate with APRs through another two channels: shopping aptitude for the highest interest account across banks and optimal rebalancing among the currently held set of accounts. In the next section we show that the latter channel is of no quantitative importance. Note that the remaining effect of advanced literacy that we estimate is net of various household and account characteristics, internet-managed accounts, as well as fixed differences across banks. That is, there is still room for financial literacy to play a role as more literate households might be better able to identify accounts across banks that for a given volume and a given set of characteristics offer the highest return (i.e., above average differences in returns that are absorbed by bank fixed effects). Thus, these findings appear consistent with the first channel outlined above, namely the limited ability of low literacy households to choose accounts across different banks that - for a given choice of volume and characteristics - offer the highest returns.

#### 4.5. Welfare Implications

In order to assess the welfare implications of limited literacy for a typical household, we first re-estimate our baseline specification at the household level. In particular, we estimate the following counterpart specification to Eq. (1):

$$r_h = \beta_1 FinLit_h + \beta_2 X_h + \beta_3 W_h + \beta_4 Z_h + \varepsilon_h, \quad (3)$$

where  $r_h$  denotes the volume-weighted APR across all accounts that each household  $h$  owns,  $FinLit_h$  represents the advanced financial literacy index for this household,  $X_h$  contains a set of household demographics (including age, education, gender, marital status, and the number of children as well as occupation status) and region fixed-effects, and vector  $W_h$  consists of

dummies denoting quartiles of total savings account volume. In addition, we condition our specification on a set of dummies representing account characteristics and bank fixed effects that are aggregated at the household level.<sup>38</sup>

The specifications (1) – (3) in Table 8 represent the counterparts to those shown in Table 4. In all three cases the estimated effects of financial literacy from both OLS and the corresponding IV regressions are comparable to those derived at the account level. For example, according to the IV estimate in the full specification, IV (3), an assumed one-standard deviation increase in advanced financial literacy implies a 30 basis points increase in the weighted APR. Other covariates in the model display a similar pattern to the one described in Section 4.2. That is, age and savings wealth associate with the interest earned, while other socio-economic characteristics such as gender, family size, employment status and education are statistically insignificant.<sup>39</sup>

It should be noted that the account-level specifications presented in Section 4.1 preclude the possibility to reallocate funds to the highest-interest account within household as a channel for financial literacy to influence the APR, while such a mechanism could be at work in the household-level regressions. The comparable effect of financial literacy in both account- and household-level specifications suggest that this channel is likely to be of limited importance. This is supported by simple data inspection: while the median number of owned accounts is two, most households tend to concentrate their savings in one account that typically earns the highest interest.<sup>40</sup>

<sup>38</sup> Given that the unit of observation is the household, dummies for account characteristics take the value one if at least one of the savings accounts in a household is subject to the restriction in question, while bank fixed effects are volume-weighted.

<sup>39</sup> As is the case with account-level regressions, we find no significant effects of dummies for income, financial and real wealth quartiles of the respective distributions when included in the household-level specifications.

<sup>40</sup> For instance, 66% of households allocate more than 80% to a single account.

Lastly, we attempt to quantify the implications of our key findings for household welfare. To that effect, we estimate how much more a typical household in the lowest literacy quartile could have earned today on its savings accounts when moved to the highest literacy quartile (other things equal).<sup>41</sup> We have to make several assumptions in order to perform such a counterfactual exercise. We consider a 10-year time horizon and suppose that earned returns are reinvested. For each year, we use as the baseline rate the median interest rate of households in the first literacy quartile, which is 2.1%. Using our preferred estimate from the IV(3) specification in Table 8, we calculate that a household, moved from the lowest to the highest literacy quartile, would earn 39 basis points more on average. We then apply this extra return to the average household savings volume.<sup>42</sup> To be conservative, we assume that additional deposits invested by households as a percentage of total savings wealth over one year grow only by the annual inflation rate.<sup>43</sup> In this set-up, losses accumulate to €838 in real terms over 10 years or 4.8% of the initially invested average amount.

## 5. Concluding Remarks

We have constructed a unique data set by matching the 2005 DNB Household Survey, which includes detailed information on individual savings accounts, various socio-economic

<sup>41</sup> For our calculation, we assume more narrowly that such an increase in financial literacy takes place for a single household only. If, for instance, a publicly sponsored program lifts financial literacy for a larger fraction of households, however, our calculation represents only a partial equilibrium analysis in the following sense. Presently, as noted above, price differentiation across banks but also across accounts at a given bank seems to be possible as consumers are sophisticated to a different degree. When more consumers become literate in this sense, there is less scope for such differentiation. In equilibrium, banks would react by adjusting their offers. One possibility, which can be supported by a formal analysis, is that as more households become willing and able to choose the best offer, offers would become more attractive across the board, in which case the general equilibrium effect of a financial literacy program would further enhance the benefits to, in particular, (newly) literate households.

<sup>42</sup> Our calculations use moments on total savings volume and financial literacy from the entire sample of households.

<sup>43</sup> This simplifies matters in the sense that additional deposits and inflation cancel out in the calculation of cumulative losses.

characteristics and financial literacy, with interest rate data on savings accounts from an administrative source, based on bank names, account names and account volume.

While savings accounts represent a relatively simple investment, say compared to direct stock holdings or retirement funds, more financially literate investors earn higher savings returns on average, controlling for demographics, account volume, and various account characteristics. We isolate one channel through which literacy positively associates with interest rates, namely familiarity with new technologies (online banking usage). We also find some evidence to suggest that more literate households are better able to identify higher interest bearing accounts across banks.

Unlike stocks and funds, savings accounts are held by the overwhelming majority of households and have the highest share in household financial wealth on average. Our findings may encourage more research in order to understand household heterogeneity in seemingly ‘simple’ and widely held financial assets. In this respect, it may be worth extending our research to countries with a varying degree of market competitiveness and product complexity, as well as less financially literate populations.

## References

Arrondel, L., Debbich, M., Savignac, F., 2012. Stockholding and financial literacy in the French population. *International Journal of Social Sciences and Humanity Studies* 4, 1309–8063.

Banks, J., O’Dea, C., Oldfield, Z., 2010. Cognitive function, numeracy and retirement saving trajectories. *The Economic Journal* 120, F381–F410.

Bernheim, D. B., 1998. Financial illiteracy, education, and retirement saving. In: Mitchell, O. S., Schieber, S. J. (eds.), *Living with Defined Contribution Pensions*, University of Pennsylvania Press, Philadelphia.

Bilias, Y., Georgarakos, D., Haliassos, M., 2010. Portfolio inertia and stock market fluctuations. *Journal of Money, Credit and Banking* 42, 715–742.

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

Calvet, L. E., Campbell, J. Y., Sodini, P., 2009. Measuring the financial sophistication of households. *The American Economic Review* 99, 393–398.

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

Carlin, B. I., 2009. Strategic price complexity in retail financial markets. *Journal of Financial Economics* 91, 278–287.

Christelis, D., Jappelli, T., Padula, M., 2010. Cognitive abilities and portfolio choice. *European Economic Review* 54, 18–38.

Grinblatt, M., Keloharju, M., Linnainmaa, J., 2011. IQ and stock market participation. *The Journal of Finance* 66, 2121–2164.

Guiso, L., Sodini, P., 2013. Chapter 21 - Household finance: An emerging field. Elsevier, vol. 2, Part B of *Handbook of the Economics of Finance*, pp. 1397 – 1532.

Hastings, J., Madrian, B., Skimmyhorn, W., 2013. Financial Literacy, Financial Education, and Economic Outcomes. *Annual Review of Economics*, Annual Reviews 5(1), pp. 347–373.

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

Hortaçsu, A., Syverson, C., 2004. Product differentiation, search costs, and competition in the mutual fund industry: A case study of s&p 500 index funds. *The Quarterly Journal of Economics* 119, 403–456.

Lewbel, A., 2012. Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics* 30, 67–80.

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

Lusardi, A., Mitchell, O. S., 2007b. Financial literacy and retirement preparedness: Evidence and implications for financial education. *Business Economics* 42, 35–44.

Lusardi, A., Mitchell, O. S., 2008. Planning and financial literacy: How do women fare? *The American Economic Review* 98, 413–417.

Lusardi, A., Mitchell, O. S., 2014. The economic importance of financial literacy: theory and evidence. *Journal of Economic Literature* 52, 5-44.

Lusardi, A., Tufano, P., 2009. Debt literacy, financial experiences, and overindebtedness. Working Paper 14808, National Bureau of Economic Research.

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

Stango, V., Zinman, J., 2009. Exponential growth bias and household finance. *The Journal of Finance* 64, 2807–2849.

Stango, V., Zinman, J., 2013. Borrowing high vs. borrowing higher: Sources and consequences of dispersion in individual borrowing costs. Working Paper 19069, National Bureau of Economic Research.

Teppa, F., Vis, C., 2012. The Centerpanel and the DNB household survey: Methodological aspects. Tech. rep., Netherlands Central Bank, Research Department.

van Rooij, M., Lusardi, A., Alessie, R., 2011. Financial literacy and stock market participation. *Journal of Financial Economics* 101, 449–472.

van Rooij, M., Lusardi, A., Alessie, R., 2012. Financial literacy, retirement planning and household wealth. *The Economic Journal* 122, 449–478.

Yoong, J., 2011. Financial illiteracy and stock market participation: Evidence from the rand American life panel. In: Mitchell, O. S., Lusardi, A. (eds.), *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*, Oxford University Press, Oxford.



**Table 1**

Summary statistics household-level variables.

The sample consists of those households used in the regressions analysis. See Appendix D for details on the construction of all variables. All statistics use sample weights. The data are from the DNB Household Survey 2005.

Variable	Mean	Std. Dev.	25th pct.	Median	75th pct.	N
Number of accounts	1.97	1.26	1.00	2.00	2.00	854
Total account volume	20,008	42,590	1,850	8,000	23,393	854
Net income	29,094	44,247	17,241	25,330	35,959	799
Net financial wealth	45,681	119,568	4,158	17,511	46,195	854
Net real wealth	116,815	205,985	2,200	26,541	182,500	854
Advanced financial literacy						
Index	0.11	0.94	-0.18	0.48	0.78	854
Number of correct answers	6.39	2.83	5.00	7.00	9.00	854
Basic financial literacy						
Index	0.12	1.06	-0.15	0.49	0.79	854
Number of correct answers	4.13	1.05	4.00	4.00	5.00	854
Economics education						
A Lot	0.17	0.38	0.00	0.00	0.00	854
Some	0.36	0.48	0.00	0.00	1.00	854
Little	0.27	0.44	0.00	0.00	1.00	854
Hardly at All/DK/Refusal	0.20	0.40	0.00	0.00	0.00	854
Financial situation oldest sibling						
No Sibling/DK/Refusal	0.14	0.35	0.00	0.00	0.00	787
Worse	0.23	0.42	0.00	0.00	0.00	787
Better/Same	0.62	0.49	0.00	1.00	1.00	787
Age	50.47	15.45	37.00	50.00	62.00	854
Education						
Less than high school	0.27	0.44	0.00	0.00	1.00	854
High school	0.33	0.47	0.00	0.00	1.00	854
College	0.41	0.49	0.00	0.00	1.00	854
Male	0.55	0.50	0.00	1.00	1.00	854
Couple	0.66	0.47	0.00	1.00	1.00	854
Number of Children	0.59	0.99	0.00	0.00	1.00	854
Occupation						
Employed	0.51	0.50	0.00	1.00	1.00	854
Self-employed	0.04	0.19	0.00	0.00	0.00	854
Unemployed	0.02	0.15	0.00	0.00	0.00	854
Other employment	0.20	0.40	0.00	0.00	0.00	854
Retired	0.23	0.42	0.00	0.00	0.00	854

**Table 2**

Distribution of interest rates across amount brackets.

This table shows the distribution of the account-level APR across nine amount brackets. In Panel A the calculation is based on the administrative data set that provides information on the individual accounts offered by banks. In Panel B the calculation is based on the matched household-administrative data that provides information on the accounts actually used by households. We group together three amount brackets from €7,000 to €10,000 due to too few observations in the respective categories for used accounts and no offered account reaching a new volume threshold within this range. All statistics use sample weights. The data are as of the last week of December 2004.

Panel A: Offered accounts									
Volume	Mean	Std. Dev.	Min.	25 <sup>th</sup> pct.	Median	75 <sup>th</sup> pct.	Max.	N	
€0 - €1,000	2.38	0.87	0.25	1.75	2.50	3.05	4.00	92	
€1,000 - €2,500	2.42	0.84	0.25	1.75	2.50	3.10	4.00	95	
€2,500 - €3,500	2.44	0.82	0.25	1.75	2.50	3.10	4.00	95	
€3,500 - €4,500	2.44	0.82	0.25	1.75	2.50	3.10	4.00	95	
€4,500 - €7,000	2.48	0.79	0.25	2.00	2.50	3.10	4.00	95	
€7,000 - €10,000	2.56	0.73	0.50	2.00	2.50	3.10	4.00	97	
€10,000 - €25,000	2.59	0.68	1.00	2.00	2.60	3.10	4.00	98	
€25,000 - €45,000	2.64	0.63	1.00	2.20	2.67	3.10	4.00	98	
> €45,000	2.65	0.63	1.00	2.20	2.70	3.10	4.00	98	

  

Panel B: Used accounts									
Volume	Mean	Std. Dev.	Min.	25 <sup>th</sup> pct.	Median	75 <sup>th</sup> pct.	Max.	N	
€0 - €1,000	2.05	0.94	1.00	1.10	1.55	3.10	4.00	364	
€1,000 - €2,500	2.13	0.91	1.00	1.10	2.40	3.10	4.00	228	
€2,500 - €3,500	2.20	0.86	1.00	1.55	2.40	3.00	4.00	104	
€3,500 - €4,500	2.32	0.84	1.00	1.60	2.40	3.00	4.00	72	
€4,500 - €7,000	2.53	0.79	1.00	2.00	2.50	3.25	4.00	119	
€7,000 - €10,000	2.48	0.76	1.00	2.20	2.50	3.10	4.00	107	
€10,000 - €25,000	2.52	0.75	1.00	1.60	2.70	3.30	3.50	254	
€25,000 - €45,000	2.60	0.61	1.00	2.10	2.50	3.25	3.50	97	
> €45,000	2.83	0.51	1.50	2.30	3.00	3.30	3.50	65	

**Table 3**

Distribution of account-level APR.

This table shows the distribution of the account-level APR across various household- and account-level variables over the full sample used in the regression analysis. All statistics use sample weights. The data are from the matched DNB Household Survey in 2005. Stars indicate whether the bivariate relationship between the APR and the respective variable is statistically significant using heteroscedasticity-robust standard errors. \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

	Mean	Std. Dev.	25th pct.	Median	75th pct.	N
	2.31	0.86	1.55	2.50	3.10	1410
Advanced Literacy Quartiles				Volume Quartiles		
1(low)			2.11 ***	1(low)	2.02 ***	
2			2.23	2	2.11 ***	
3			2.36	3	2.43 ***	
4(high)			2.46 ***	4(high)	2.59 ***	
Age				Education		
18-30 years			2.58 ***	Less than high school	2.27	
31-40 years			2.33	High school	2.34	
41-50 years			2.22 *	College	2.32	
51-60 years			2.37			
61 years and older			2.22 ***			
Gender				Married		
Female			2.30	Single-Person Households	2.25	
Male			2.32	Two-Person Households	2.34	
Internet account				Withdrawal costs / limitations		
No			1.83	No	2.38	
Yes			3.21 ***	Yes	1.67 ***	
Minimum amount				Salary account		
No			2.28	No	2.30	
Yes			2.54 ***	Yes	3.07 ***	
Lowest balance bonus				Individual ownership		
No			2.20	No	2.34	
Yes			2.61 ***	Yes	2.29	
Balance growth bonus				Joint ownership		
No			2.28	No	2.28	
Yes			3.40 ***	Yes	2.37	
Fixed monthly deposit				Third party ownership		
No			2.30	No	2.33	
Yes			4.00 ***	Yes	2.18	

**Table 4**

OLS of account-level APR on financial literacy.

The table reports OLS and IV estimates from regressions of the account-level APR on financial literacy and several other controls. The sample excludes accounts with volume below €50. All IV specifications use economics education and the financial situation of the oldest sibling as an instrument for advanced financial literacy. All IV estimates use two-stage least squares. The data are from the matched DNB Household Survey in 2005. Standard errors are clustered at the household level. \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

	OLS (1)		IV (1)		OLS (2)		IV (2)		OLS (3)		IV (3)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Advanced Financial Literacy	0.116***	0.032	0.304***	0.112	0.128***	0.031	0.275***	0.125	0.127***	0.022	0.288***	0.099
<i>Age dummies</i>												
31-40 years					-0.316***	0.103	-0.246**	0.118	-0.241***	0.079	-0.197**	0.092
41-50 years					-0.474***	0.099	-0.424***	0.106	-0.372***	0.074	-0.357***	0.078
51-60 years					-0.400***	0.100	-0.377***	0.108	-0.285***	0.075	-0.303***	0.081
61 years and older					-0.541***	0.123	-0.564***	0.135	-0.491***	0.090	-0.547***	0.099
<i>Education dummies</i>												
High school					0.039	0.066	0.008	0.081	0.081	0.051	0.050	0.064
College					-0.083	0.064	-0.134	0.085	0.016	0.051	-0.050	0.068
Male					-0.025	0.056	-0.068	0.073	0.011	0.041	-0.047	0.060
Couple					0.014	0.059	-0.023	0.062	0.066	0.053	0.021	0.056
Number of Children					0.021	0.029	0.018	0.030	0.002	0.023	0.001	0.023
<i>Occupation dummies</i>												
Employed					0.061	0.123	0.083	0.134	0.102	0.108	0.088	0.126
Self-employed					0.048	0.163	0.045	0.192	0.063	0.126	0.076	0.152
Unemployed					0.051	0.077	0.098	0.083	0.088	0.060	0.114*	0.065
Retired					-0.021	0.104	0.020	0.110	0.016	0.077	0.052	0.084
<i>Volume dummies</i>												
€1,000 – €2,500					0.108	0.084	0.124	0.088	0.111*	0.065	0.116*	0.069
€2,500 – €3,500					0.185*	0.098	0.185*	0.102	0.078	0.076	0.074	0.077
€3,500 – €4,500					0.271**	0.108	0.249**	0.113	0.219**	0.079	0.216***	0.082
€4,500 – €7,000					0.460***	0.090	0.396***	0.095	0.320***	0.066	0.297***	0.070
€7,000 – €10,000					0.466***	0.095	0.461***	0.098	0.339**	0.078	0.338***	0.083
€10,000 – €25,000					0.503***	0.076	0.447***	0.082	0.474***	0.059	0.434***	0.063
€25,000 – €45,000					0.637***	0.084	0.601***	0.091	0.622***	0.075	0.610***	0.080
€45,000 or more					0.862***	0.088	0.817***	0.093	0.753***	0.081	0.724***	0.083
<i>Account characteristics</i>												
Minimum amount									-0.369***	0.097	-0.399***	0.102
Lowest balance bonus									-0.328***	0.104	-0.345***	0.112
Balance growth bonus									2.121***	0.117	2.169***	0.128
Fixed monthly deposit									1.396***	0.076	1.358***	0.097
Withdrawal costs / limitations									-0.266***	0.060	-0.301***	0.064
Salary account									1.118***	0.152	1.071***	0.150
Joint ownership									0.001	0.048	0.005	0.051
Third party ownership									-0.040	0.063	-0.030	0.071
Constant	2.315***	0.027	2.268***	0.036	2.272***	0.124	2.258***	0.136	2.380***	0.104	2.397***	0.116
Region dummies					yes		yes		yes		yes	
Bank fixed effects									yes		yes	
N	1,410		1,306		1,410		1,306		1,410		1,306	
Adjusted R-squared	0.01		-0.02		0.12		0.10		0.47		0.45	
Hansen J-test p-value			0.29				0.60				0.45	
F-statistic first stage			11.54				9.74				10.44	
Exogeneity test p-value			0.12				0.23				0.07	

**Table 5**

OLS of APR on financial literacy and controls using the Lewbel (2012) method.

The table reports IV estimates from regressions of the account-level APR on financial literacy and age, gender and number of children using: (1) generated instruments from the Lewbel method alone; (2) external instruments (economics education and the financial situation of the oldest sibling) alone; (3) both sets of generated instruments under the Lewbel method and external instruments. The sample excludes accounts with volume below €50. The data are from the matched DNB Household Survey in 2005. Standard errors are clustered at the household level. \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

	IV(1): Lewbel		IV(2): Standard		IV(3): Combined	
	Estimate	SE	Estimate	SE	Estimate	SE
Advanced Financial Literacy	0.218**	0.100	0.300**	0.121	0.228***	0.077
Age dummies						
31-40 years	-0.163	0.116	-0.147	0.123	-0.161	0.117
41-50 years	-0.326***	0.109	-0.321***	0.110	-0.325***	0.109
51-60 years	-0.233**	0.107	-0.246**	0.109	-0.235**	0.107
61 years and older	-0.370***	0.105	-0.378***	0.105	-0.371***	0.105
Male	-0.087	0.072	-0.126	0.078	-0.091	0.065
Number of children	0.024	0.031	0.027	0.031	0.025	0.030
Constant	2.577	0.098	2.581	0.098	2.578	0.098
N	1306		1306		1306	
Hansen J-test	3.51		5.73		10.60	
Hansen J-test p-value	0.62		0.22		0.39	
F-statistic first stage	6.03		11.12		16.75	
Exogeneity test p-value	0.30		0.20		0.16	

**Table 6**

Robustness checks.

The table shows variations of the OLS (3) and IV (3) specifications of Table 4. Panel A: (1) estimates Eq. (1) excluding the nine volume dummies; (2) uses only accounts reported by the financial head; (3) weighs each observation by its volume share within the household. Panel B: OLS (1) uses financial literacy quartiles instead of an index; IV (1) uses the financial literacy index but excludes “don’t know”-answers from the instrument economics education; (2) uses the standardized number of correct answers to the financial literacy questions; (3) uses the standardized number of correct answers to the ‘Big-Three’-questions only. IV uses economics education and the financial situation of the oldest sibling as instruments for advanced financial literacy. All IV estimates use two-stage least squares. The data are from the matched DNB Household Survey in 2005. Standard errors are clustered at the household level. \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

Panel A: Variations of Eq (1)	OLS (1)		IV (1)		OLS (2)		IV (2)		OLS (3)		IV (3)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Advanced Financial Literacy	0.136***	0.025	0.317***	0.104	0.127***	0.024	0.286***	0.102	0.126***	0.022	0.260***	0.09
Demographics and other controls	yes		yes		yes		yes		yes		yes	
Region Dummies	yes		yes		yes		yes		yes		yes	
Account characteristics	yes		yes		yes		yes		yes		yes	
Bank fixed effects	yes		yes		yes		yes		yes		yes	
N	1,410		1,306		1,177		1,091		1,410		1,306	
Adjusted R-squared	0.42		0.39		0.49		0.46		0.46		0.43	
Hansen J-test p-value			0.14				0.32				0.56	
F-statistic first stage			10.43				11.97				13.49	
Exogeneity test p-value			0.05				0.1				0.12	

  

Panel B: Different literacy measures	OLS (1)		IV (1)		OLS (2)		IV (2)		OLS (3)		IV (3)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
<i>Financial literacy quartiles</i>												
Second quartile	0.198***	0.054										
Third quartile	0.209***	0.055										
Fourth quartile	0.288***	0.057										
Advanced Financial Literacy			0.262***	0.094								
Number of correct answers					0.117***	0.022	0.318***	0.116	0.060***	0.023	0.428**	0.196
‘Big-Three’-Index					yes		yes		yes		yes	
Demographics and other controls	yes		yes		yes		yes		yes		yes	
Region Dummies	yes		yes		yes		yes		yes		yes	
Account characteristics	yes		yes		yes		yes		yes		yes	
Bank fixed effects	yes		yes		yes		yes		yes		yes	
N	1,410		1,290		1,410		1,306		1,410		1,306	
Adjusted R-squared	0.47		0.45		0.47		0.43		0.46		0.32	
Hansen J-test p-value			0.48				0.46				0.46	
F-statistic first stage			11.21				7.11				2.68	
Exogeneity test p-value			0.10				0.05				0.02	

**Table 7**

OLS of APR on financial literacy and online banking.

The table reports OLS and IV estimates from regressions of the account-level APR on financial literacy and several controls (used in the third baseline specification shown in Table 4) including in addition a dummy that represents internet-managed accounts. The sample excludes accounts with volume below €50. IV uses economics education and the financial situation of the oldest sibling as instruments for advanced financial literacy. All IV estimates use two-stage least squares. The data are from the matched DNB Household Survey in 2005. Standard errors are clustered at the household level. \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

	OLS		IV	
	Estimate	SE	Estimate	SE
Advanced Financial Literacy	0.041***	0.011	0.124**	0.050
Internet account	1.326***	0.03	1.308***	0.034
Constant	1.765***	0.063	1.776***	0.073
Demographics and other controls	yes		yes	
Region dummies	yes		yes	
Account characteristics	yes		yes	
Bank fixed effects	yes		yes	
<i>N</i>	1,410		1,306	
Adjusted R-squared	0.84		0.83	
Hansen J-test p-value			0.85	
F-statistic first stage			10.10	
Exogeneity test p-value			0.11	

**Table 8**

OLS of household-level APR on financial literacy.

The table reports OLS and IV estimates from regressions of the household-level APR on financial literacy and several controls. The sample excludes accounts with volume below €50. All IV-specifications use economics education and the financial situation of the oldest sibling as instruments for advanced financial literacy. All IV estimates use two-stage least squares. The data are from the matched DNB Household Survey in 2005. Standard errors are clustered at the household level. \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

	OLS (1)		IV (1)		OLS (2)		IV (2)		OLS (3)		IV (3)	
	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
Advanced Financial Literacy	0.148***	0.028	0.317***	0.091	0.129***	0.028	0.286***	0.110	0.123***	0.023	0.297***	0.093
<i>Age dummies</i>												
31-40 years					-0.326***	0.102	-0.264**	0.111	-0.267***	0.084	-0.225**	0.092
41-50 years					-0.542***	0.100	-0.465***	0.109	-0.449***	0.082	-0.408***	0.088
51-60 years					-0.407***	0.102	-0.359***	0.112	-0.301***	0.084	-0.291***	0.092
61 years and older					-0.540***	0.130	-0.541***	0.137	-0.537***	0.107	-0.560***	0.112
<i>Education dummies</i>												
High school					-0.010	0.067	-0.030	0.077	0.028	0.056	0.007	0.064
College					-0.051	0.067	-0.097	0.084	0.006	0.057	-0.060	0.074
Male					-0.023	0.058	-0.023	0.073	0.021	0.048	-0.028	0.062
Couple					0.018	0.061	-0.016	0.065	0.109*	0.066	0.058	0.070
Number of Children					0.008	0.030	0.005	0.031	-0.003	0.025	-0.002	0.026
<i>Occupation dummies</i>												
Employed					0.053	0.130	0.084	0.138	0.081	0.112	0.078	0.128
Self-employed					0.191	0.149	0.212	0.168	0.181*	0.109	0.208*	0.126
Unemployed					-0.071	0.079	-0.036	0.084	0.004	0.066	0.027	0.070
Retired					-0.078	0.111	-0.037	0.113	0.020	0.092	0.045	0.094
<i>Savings wealth quartiles</i>												
Second quartile					0.254***	0.078	0.234***	0.082	0.218***	0.064	0.207***	0.068
Third quartile					0.411***	0.075	0.341***	0.083	0.377***	0.063	0.318***	0.067
Fourth quartile					0.634***	0.077	0.533***	0.093	0.600***	0.069	0.528***	0.079
<i>Account characteristics</i>												
Minimum amount									-0.204***	0.070	-0.251***	0.078
Lowest balance bonus									-0.181**	0.075	-0.207**	0.082
Balance growth bonus									0.779***	0.143	0.843***	0.156
Fixed monthly deposit									0.705***	0.165	0.611***	0.166
Withdrawal costs / limitations									-0.195***	0.067	-0.234***	0.072
Salary account									1.032***	0.120	0.998***	0.120
Joint ownership									-0.071	0.058	-0.055	0.062
Third party ownership									-0.115*	0.069	-0.096	0.073
Constant	2.341***	0.027	2.316***	0.031	2.322***	0.126	2.317***	0.143	2.339***	0.121	2.363***	0.139
Region dummies					yes		yes		yes		yes	
Bank fixed effects									yes		yes	
N	854		787		854		787		854		787	
Adjusted R-squared	0.03		-0.01		0.13		0.10		0.41		0.36	
Hansen J-test p-value			0.14				0.35				0.41	
F-statistic first stage			18.96				14.42				15.30	
Exogeneity test p-value			0.08				0.13				0.05	



## Appendix A. Descriptive Statistics on Savings Accounts

**Table A1**

Summary statistics account-level variables.

The sample consists of those accounts used in the regression analysis. All statistics use sample weights. The data are from the matched DNB Household Survey 2005.

Variable	Mean	Std. Dev.	25th pct.	Median	75th pct.
APR	2.31	0.86	1.55	2.50	3.10
Account volume in €	10,973	26,756	857	3,500	12,750
Bank fixed effects					
ABN AMRO	0.12	0.33	0.00	0.00	0.00
ING Bank	0.37	0.48	0.00	0.00	1.00
Rabobank	0.30	0.46	0.00	0.00	1.00
Fortis Bank	0.03	0.17	0.00	0.00	0.00
SNS Bank	0.03	0.17	0.00	0.00	0.00
Small Banks	0.15	0.36	0.00	0.00	0.00
Ownership					
Individual	0.52	0.50	0.00	1.00	1.00
Joint	0.39	0.49	0.00	0.00	1.00
Third party	0.09	0.29	0.00	0.00	0.00
Account restrictions					
Internet account	0.35	0.48	0.00	0.00	1.00
Minimum amount	0.13	0.33	0.00	0.00	0.00
Lowest balance bonus	0.28	0.45	0.00	0.00	1.00
Balance growth bonus	0.03	0.16	0.00	0.00	0.00
Fixed monthly deposit	0.01	0.10	0.00	0.00	0.00
Withdrawal costs / limitations	0.09	0.29	0.00	0.00	0.00
Salary account	0.02	0.15	0.00	0.00	0.00

**Table A2**

Comparison of full vs. final sample across major demographics and account characteristics.

The full sample contains all accounts held by households in the DNB Household Survey. The final sample contains only those accounts used in the estimation sample. All statistics use sample weights.

Variable	Full sample (N=2337)		Final sample (N=1410)	
	Mean	Std. Dev.	Mean	Std. Dev.
Age	49.94	15.61	50.74	15.33
Male	0.56	0.50	0.58	0.49
Couple	0.71	0.45	0.71	0.46
Number of Children	0.68	1.04	0.65	1.04
Less than high school	0.25	0.43	0.24	0.42
High school	0.35	0.48	0.34	0.47
College	0.41	0.49	0.43	0.49
ABN AMRO	0.13	0.34	0.12	0.33
ING Bank	0.34	0.47	0.37	0.48
Rabobank	0.28	0.45	0.30	0.46
Fortis Bank	0.04	0.19	0.03	0.17
SNS Bank	0.04	0.19	0.03	0.17
Small Banks	0.17	0.38	0.15	0.36
Volume in €	9,636	23,804	10,973	26,756
Individual ownership	0.53	0.50	0.52	0.50
Joint ownership	0.36	0.48	0.39	0.49
Third party ownership	0.11	0.31	0.09	0.29

## Appendix B. First Stage Regression

**Table B1**

First stage regressions, account-level APR.

The table reports estimates from first-stage regressions used to estimate the IV models shown in Table 4. The instruments employed refer to economics education and the financial situation of the oldest sibling. The reference group for the first consists of those with a lot of education in economics, while the base category for the second regards those with no siblings and refusals. The data are from the matched DNB Household Survey 2005. Standard errors are clustered at the household level. \* p<0.1; \*\* p<0.05; \*\*\* p<0.01

	OLS (1)		OLS (2)		OLS (3)	
	Estimate	SE	Estimate	SE	Estimate	SE
<i>Financial situation oldest sibling</i>						
Worse	0.336***	0.115	0.326***	0.104	0.326***	0.102
The same or better	0.107	0.109	0.150	0.097	0.154	0.096
<i>Economics education</i>						
Some	-0.287***	0.082	-0.259***	0.082	-0.261***	0.081
Little	-0.458***	0.088	-0.404***	0.088	-0.414***	0.087
Hardly at all	-0.694***	0.107	-0.579***	0.101	-0.586***	0.098
<i>Age dummies</i>						
31-40 years			-0.199	0.148	-0.189	0.147
41-50 years			0.011	0.123	0.023	0.121
51-60 years			0.264**	0.113	0.265**	0.112
61 years and older			0.399***	0.149	0.420***	0.148
<i>Education dummies</i>						
High school			0.295***	0.093	0.274***	0.091
College			0.383***	0.084	0.363***	0.086
Male			0.316***	0.072	0.323***	0.072
Couple			0.029	0.073	0.064	0.079
Number of Children			-0.019	0.039	-0.019	0.040
Basic financial literacy			0.140	0.213	0.125	0.216
<i>Occupation dummies</i>						
Employed			-0.029	0.194	-0.008	0.192
Self-employed			-0.186*	0.102	-0.188*	0.100
Unemployed			-0.191	0.127	-0.193	0.126
<i>Volume dummies</i>						
€1,000 – €2,500			0.026	0.083	0.015	0.083
€2,500 – €3,500			0.184**	0.092	0.140	0.090
€3,500 – €4,500			0.164	0.108	0.114	0.110
€4,500 – €7,000			0.162	0.099	0.082	0.099
€7,000 – €10,000			0.142	0.107	0.104	0.109
€10,000 – €25,000			0.174**	0.085	0.132	0.082
€25,000 – €45,000			0.139	0.106	0.081	0.109
€45,000 or more			0.271***	0.103	0.213**	0.102
Constant	0.421***	0.106	-0.206	0.171	-0.151	0.188
Region dummies			yes		yes	
Account characteristics					yes	
Bank fixed effects					yes	
N	1,306		1,306		1,306	
Adjusted R-squared	0.07		0.20		0.20	
F-statistic first stage	11.54		9.74		10.44	

## Appendix C. Details on Data Processing

Whereas the majority of survey respondents provide a bank name, the data on the names of savings accounts contain some typos, abbreviations, and few inconsistencies. We process this raw information in the DHS in the following way. Using the bank and account names from the market interest rate data as a reference for the correct spelling, we replace all incorrectly spelled names and abbreviations in the DHS by their proper name. We replace those cases in which participants report outdated names of accounts by the names of their successor accounts. Finally, we set all potential inconsistent cases to missing.<sup>44</sup>

As we later match the DHS and market data based on volume as well, we also recover missing volumes of individual savings accounts following the procedure used by the official provider of the DHS (CentERdata). CentERdata first recovers volumes for individual savings accounts (details follow) and then aggregates over all accounts of each household member yielding total savings volume per household member (i.e., at the individual level). Only the recovered volume of the latter is available in the public version of the dataset. However, we are able to recover the large majority of the inserted values for individual savings accounts by following the same process that CentERdata has applied to calculate total savings account volume per household member.<sup>45</sup>

First, if a respondent does not report the exact amount of a savings account, the respondent is asked to choose from a sequence of follow-up questions in the form of unfolding brackets. In this case, we use the mid-point of the bracketed answer or the lower bound in case of the highest open-ended category (€25,000 or more). This leaves 10.3% of accounts with missing volume. Second, for these missing cases, we use the average amount of this savings account over the last two years. This leaves 8.1% of accounts with missing volume. For the remaining individual household members with at least one account with unreported volume, an imputed value for total savings volume was used by CentERdata. This was derived from a regression of total savings volume on a large set of individual characteristics. We use this imputed value to recover the

<sup>44</sup> For example, some respondents report accounts which are not offered anymore by the reported bank in 2005 but were replaced by an account with a different name.

<sup>45</sup> Details can be found in the documentation of the DHS 2005 wave (available at: <http://cdata3.uvt.nl/dhs/files/CodebookWave2005English.pdf>)

volume of individual savings accounts in the following way. If only one account of a household member is left with missing volume, we use the difference between the total savings volume and the sum of all reported account volumes of that individual to fill in the single missing volume.<sup>46</sup> This still leaves few individual household members with more than one account with missing volume. For those household members, we distribute this difference equally across all savings accounts with remaining missing volume.<sup>47</sup>

<sup>46</sup> 60% of those household members hold only one account and thus total volume and individual account volume are equivalent.

<sup>47</sup> Note that in the last two cases, we only consider accounts that do not exceed the total number of accounts as originally stated by the respondent, for example, we only consider the first three reported accounts of a household that claims to have 3 accounts in total but reports four. The same approach is used in the DHS for the calculation of total savings wealth.

## Appendix D. Details on key questions used

### Account Characteristics

- (1) *Minimum amount*: Minimum amount required to earn full interest rate
- (2) *Lowest balance bonus*: Balance may not fall below specified amount threshold during calendar year/quarter to earn full rate
- (3) *Balance growth bonus*: Balance needs to grow by specified amount per calendar year/quarter
- (4) *Fixed monthly deposit*: Automatic savings option requiring fixed deposit each month
- (5) *Withdrawal costs / limitations*: Additional withdrawal / account fees or limited number of withdrawals
- (6) *Salary account*: Salary account required at the same bank
- (7) *Internet account*: Account is fully managed online

### Account Holder

SPA71 thru SPA77

Who is the account holder of your [1<sup>st</sup> thru 7<sup>th</sup>] account?

- 1 the account is registered in my own name ..... SPA91
- 2 the account is registered in my partner's/spouse's name ..... SPA91
- 3 the account is registered jointly in my own name and someone else's name  
(e.g. partner/spouse)..... SPA91
- 4 the account is registered in (one of) my parents' name ..... SPA91
- 5 other ..... SPA81

#### Note:

We group category 2,4 and 5 into one category 'third party ownership' due to a small number of observations in these categories.

### Account Volume

SPA131 thru SPA137

What was the balance of your [1<sup>st</sup> thru 7<sup>th</sup>] account on 31 December 2004?

- amount: .....BZ04  
don't know ..... SPA141

### Bank Fixed Effects

SPA91 thru SPA97

With which bank or financial institution is your [1<sup>st</sup> thru 7<sup>th</sup>] account registered?

- 1 ABN AMRO ..... SPA121
- 2 Postbank ..... SPA121
- 3 Rabobank ..... SPA121
- 4 ING Bank ..... SPA121
- 5 Fortis Bank ..... SPA121
- 6 SNS Bank ..... SPA121
- 7 other ..... SPA101

#### Note:

We group ING and Postbank together as the former mainly operates through the latter in the market for savings accounts.

### Economics Education

How much of your education was devoted to economics?

- 1 A lot
- 2 Some
- 3 Little
- 4 Hardly at all
- 5 Do not know
- 6 Refusal

#### Note:

We group the last three categories together and use dummies for each category in the regressions analysis, in which 'a lot of education in economics' serves as the base category.

### **Financial Literacy (Advanced)**

- (1) Which statement describes the main function of the stock market? (i) The stock market helps to predict stock earnings; (ii) the stock market results in an increase in the price of stocks; (iii) the stock market brings people who want to buy stocks together with those who want to sell stocks; (iv) none of the above; (v) do not know; (vi) refusal.
- (2) What happens if somebody buys the stock of firm B in the stock market? (i) He owns a part of firm B; (ii) he has lent money to firm B; (iii) he is liable for firm B debt; (iv) none of the above; (v) do not know; (vi) refusal.
- (3) Which statement about mutual funds is correct? (i) Once one invests in a mutual fund, one cannot withdraw the money in the first year; (ii) mutual funds can invest in several assets, for example, invest in both stocks and bonds; (iii) mutual funds pay a guaranteed rate of return which depends on their past performance; (iv) none of the above; (v) do not know; (vi) refusal.
- (4) What happens if somebody buys a bond of firm B? (i) He owns a part of firm B; (ii) he has lent money to firm B; (iii) he is liable for firm B's debts; (iv) none of the above; (v) do not know; (vi) refusal.
- (5) Considering a long time period (e.g. 10 or 20 years), which asset normally gives the highest return? (i) Savings accounts; (ii) bonds; (iii) stocks; (iv) do not know; (v) refusal.
- (6) When an investor spreads his money among different assets, does the risk of losing money (i) increase; (ii) decrease; (iii) stay the same; (iv) do not know; (v) refusal.
- (7) If you buy a 10-year bond, it means you cannot sell it after five years without incurring a major penalty. (i) True; (ii) false; (iii) do not know; (iv) refusal.
- (8) Stocks are normally riskier than bonds. (i) True; (ii) false; (iii) do not know; (iv) refusal.
- (9) Buying a company fund usually provides a safer return than a stock mutual fund. (i) True; (ii) false; (iii) do not know; (iv) refusal.
- (10) If the interest rate falls, what should happen to bond prices? (i) Rise; (ii) fall; (iii) stay the same; (iv) none of the above; (v) do not know; (vi) refusal.
- (11) Normally, which asset displays the highest fluctuations over time? (i) Savings accounts; (ii) bonds; (iii) stocks; (iv) do not know; (v) refusal.

#### Note:

The index for advanced financial literacy is gratefully borrowed from van Rooij et al. (2011) and is the result of a factor analysis on the above set of questions as described in the appendix of their paper.

### **Financial Literacy (Basic)**

- (1) Suppose you had €100 in a savings account and the interest rate was 2% per year. After five years, how much do you think you would have in the account if you left the money to grow? (i) More than €102; (ii) exactly €102; (iii) less than €102; (iv) do not know; (v) refusal.
- (2) 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 five years, how much would you have on this account in total? (i) More than €200; (ii) exactly €200; (iii) less than €200; (iv) do not know; (v) refusal.
- (3) Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After one year, how much would you be able to buy with the money in this account? (i) More than today; (ii) exactly the same; (iii) less than today; (iv) do not know; (v) refusal.
- (4) Assume a friend inherits €10,000 today and his sibling inherits €10,000 three years from now. Who is richer because of the inheritance? (i) My friend; (ii) his sibling; (iii) they are equally rich; (iv) do not know; (v) refusal.
- (5) Suppose that in the year 2010, your income has doubled and prices of all goods have doubled too. In 2010, how much will you be able to buy with your income? (i) More than today; (ii) the same; (iii) less than today; (iv) do not know; (v) refusal.

Note:

The index for basic financial literacy is gratefully borrowed from van Rooij et al. (2011) and is the result of a factor analysis on the above set of questions as described in the appendix of their paper.

**Financial Situation of the Oldest Sibling**

Would you say that your oldest [brother/sister] is in worse, better, or about the same financial condition than you?

- 1 Worse
- 2 Better
- 3 About the same
- 4 Refusal
- 5 Do not know

Note:

We group the last two categories and those with no siblings together and use dummies for each category in the regressions analysis in which Don't know/Refusal/No Sibling serves as the base category.

**Hours Worked**

UURWERK

How many hours per week do/did you on average IN FACT spend on your (last) (most important) job? For this question it doesn't make any difference whether overtime work is/was paid for or not.

- number of hours:.....
- if paid job NOW ..... VAKAN
- if USED TO HAVE a paid job ..... VAKVERL

**Investment Horizon**

PERIODE1

People use different time-horizons when they decide about what part of the income to spend, and what part to save. Which of the time-horizons mentioned below is in your household MOST important with regard to planning expenditures and savings?

- 1 the next couple of months ..... BIJHOUD
- 2 the next year..... BIJHOUD
- 3 the next couple of years ..... BIJHOUD
- 4 the next 5 to 10 years..... BIJHOUD
- 5 more than 10 years from now..... BIJHOUD

**Last Year's Net Income**

INKNORM

Is this income unusually high or low compared to the income you would expect in a 'regular' year, or is it regular?

- 1 unusually low ..... INKROND
- 2 regular..... INKROND
- 3 unusually high ..... INKROND
- 4 don't know ..... INKROND

Note:

We set 'don't know' to missing and group the categories 2 and 3 together. The resulting category serves as the base in the regression analysis.

**Savings Wealth**

SPA3 What was the total balance of your SAVINGS OR DEPOSIT ACCOUNTS on 31 December 2004?

- amount:..... SPA71
- don't know ..... SPA4

**Net Financial Wealth**

Sum of checking, deposit books, savings certificates, bonds, stocks, funds, options, employer-sponsored savings plan, insurances, and other financial wealth minus total financial debt.



Note:

We count financial wealth as missing only if all asset subcomponents are missing.

**Net Income**

Equal to the derived net income on individual level as provided by CentER in the aggregated data on income and aggregate to household level.

Note:

In contrast to the DHS which counts net income as missing if one of the sub components is missing, we count net income as missing only if all asset subcomponents are missing. Bot variables are highly similar in terms of distributional aspects, since often only small subcomponents are not reported.

**Net Real Wealth**

Difference between real assets and real debt. Real assets are the sum of home equity (value of first and second house as well as pieces of real estate), business wealth, and other real wealth (cars, motorbikes, boats and caravans). Real debt is the sum of mortgages on the first and second house as well on other pieces of real estate.

**Risk Aversion**

R1) Suppose that you are the only income earner in the family, and you have a good job guaranteed to give you your current (family) income every year for life. You are given the opportunity to take a new, equally good job, with a 50% chance it will double your (family) income and a 50% chance that it will cut your (family) income by a third. Would you take the new job?

Yes  No  Do not know

[If R1=="yes", then R2]

R2) Suppose the chances were 50% that it would double your (family) income, and 50% that it would cut it in half. Would you take the new job?

Yes  No  Do not know

[If R1=="no" or "do not know", then R3]

R3) Suppose the chances were 50% that it would double your (family) income and 50% that it would cut it by 20%. Would you then take the new job?

Yes  No  Do not know

Note:

We follow the grouping as performed by van Rooij, Lusardi and Alessie (2011). Least risk averse: choosing the most risky option twice; medium risk averse: choosing the most risky option in the first but not in the second question; risk averse: safe option in the first but not in the second question; most risk averse: choosing the safe option twice; do not know: no choice in the first question.

## Household Finance and Consumption Network

This paper contains research conducted within the Household Finance and Consumption Network (HFCN). The HFCN consists of survey specialists, statisticians and economists from the ECB, the national central banks of the Eurosystem and a number of national statistical institutes.

The HFCN is chaired by Oreste Tristani (ECB) and Carlos Sánchez Muñoz (ECB). Michael Haliassos (Goethe University Frankfurt), Tullio Jappelli (University of Naples Federico II), Arthur Kennickell (Federal Reserve Board) and Peter Tufano (University of Oxford) and act as external consultants, and Sébastien Pérez Duarte (ECB) and Jiri Slacalek (ECB) as Secretaries.

The HFCN collects household-level data on households' finances and consumption in the euro area through a harmonised survey. The HFCN aims at studying in depth the micro-level structural information on euro area households' assets and liabilities. The objectives of the network are:

- 1) understanding economic behaviour of individual households, developments in aggregate variables and the interactions between the two;
- 2) evaluating the impact of shocks, policies and institutional changes on household portfolios and other variables;
- 3) understanding the implications of heterogeneity for aggregate variables;
- 4) estimating choices of different households and their reaction to economic shocks;
- 5) building and calibrating realistic economic models incorporating heterogeneous agents;
- 6) gaining insights into issues such as monetary policy transmission and financial stability.

The refereeing process of this paper has been co-ordinated by a team composed of Oreste Tristani (ECB), Pirmin Fessler (Oesterreichische Nationalbank), Michalis Haliassos (Goethe University Frankfurt), Tullio Jappelli (University of Naples Federico II), Sébastien Pérez-Duarte (ECB), Jiri Slacalek (ECB), Federica Teppa (De Nederlandsche Bank), Peter Tufano (Oxford University) and Philip Vermeulen (ECB).

The paper is released in order to make the results of HFCN research generally available, in preliminary form, to encourage comments and suggestions prior to final publication. The views expressed in the paper are the author's own and do not necessarily reflect those of the ESCB.

## Acknowledgements

We would like to thank Tabea Bucher-Koenen, Gabriel Fagan, Luigi Guiso, Arthur Kennickell, Annamaria Lusardi, Maarten van Rooij, and participants at the Annual Congress of the European Economic Association, Toulouse; NETSPAR International Pension Workshop, Amsterdam; Research in Behavioural Finance Conference, Rotterdam; HFCN meeting at the ECB, Frankfurt for helpful comments. We are also grateful to Rob Alessie, Annamaria Lusardi, and Maarten van Rooij for making the data from a special financial literacy module publicly available, to a major Dutch bank for providing interest rate data on savings accounts in the Netherlands, to CentERdata for providing additional information on financial product holdings and bank relationships of DNB survey participants, as well as to SpaarInformatie for providing data on restrictions of Dutch savings accounts. The opinions expressed in the paper are those of the authors and do not reflect the views of the Deutsche Bundesbank or its staff.

## Florian Deuflhard

Goethe University, Frankfurt am Main, Germany; e-mail: [florian.deuflhard@hof.uni-frankfurt.de](mailto:florian.deuflhard@hof.uni-frankfurt.de)

## Dimitris Georgarakos

Deutsche Bundesbank and University of Leicester, Frankfurt am Main, Germany; e-mail: [dimitris.georgarakos@bundesbank.de](mailto:dimitris.georgarakos@bundesbank.de)

## Roman Inderst

Goethe University, Frankfurt am Main, Germany; e-mail: [inderst@finance.uni-frankfurt.de](mailto:inderst@finance.uni-frankfurt.de)

## © European Central Bank, 2015

Postal address 60640 Frankfurt am Main, Germany  
Telephone +49 69 1344 0  
Website [www.ecb.europa.eu](http://www.ecb.europa.eu)

All rights reserved. Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the authors.

This paper can be downloaded without charge from [www.ecb.europa.eu](http://www.ecb.europa.eu), from the Social Science Research Network electronic library at <http://ssrn.com> or from RePEc: Research Papers in Economics at <https://ideas.repec.org/s/ecb/ecbwps.html>.

Information on all of the papers published in the ECB Working Paper Series can be found on the ECB's website, <http://www.ecb.europa.eu/pub/scientific/wps/date/html/index.en.html>.

ISSN 1725-2806 (online)  
ISBN 978-92-899-1665-3  
DOI 10.2866/35825  
EU catalogue No QB-AR-15-092-EN-N