Women, confidence, and financial literacy

Tabea Bucher-Koenen
Max-Planck-Institute for Social Law and Social Policy and Netspar
Rob Alessie
University of Groningen and Netspar
Annamaria Lusardi
George Washington School of Business, NBER and Netspar
and
Maarten van Rooij*
De Nederlandsche Bank and Netspar

This version: April 2017

VERY PRELIMINARY AND INCOMPLETE – Please do not cite

Abstract
The literature documents robust evidence of a gender gap in financial literacy: Women consistently show lower levels of financial literacy than men. We have devised two surveys to investigate whether this gender gap is the result of lack of knowledge or lack of confidence. Our findings show that women are less confident in their knowledge than men. They disproportionately answer "do not know" to financial knowledge questions, even if they know the correct answer. We develop an empirical strategy based on a latent class model to consistently estimate whether the respondent knows the correct answer. An important implication of our findings is that traditional financial literacy measures are plagued by confidence bias or measurement error. Using the corrected measures for financial literacy, we show that financial knowledge is important to explain household financial behavior such as stock market participation.

Keywords: financial literacy, gender difference, financial decision-making, measurement error, latent class model
JEL code: C81, D91

* Tabea Bucher-Koenen, Munich Center for the Economics of Aging at the Max-Planck-Institute for Social Law and Social Policy, Amalienstr. 33, 80799 Munich, Germany (bucher-koenen@mea.mpisoc.mpg.de), Rob J.M. Alessie, School of Economics and Business, University of Groningen, P.O. Box 800, 9700 AV, Groningen (R.J.M.Alessie@rug.nl),), Annamaria Lusardi, George Washington School of Business, Duquès Hall, Suite 450E, Washington D.C. (Annamaria.Lusardi@gwu.edu), Maarten C.J. van Rooij, Economics & Research Division, De Nederlandsche Bank, P.O. Box 98, 1000 AB, Amsterdam (M.C.J.van.Rooij@dnb.nl).
The authors wish to thank Martin Brown for feedback. They are also grateful to the participants at the Annual Meeting of the Neuroeconomic Society (Bonn 2013), the EIB meetings (Paris 2013 and 2014), the Annual Conference of the European Economics Association (Toulouse 2014), the second Cintia conference (Turin 2014) and the Annual Meeting of the International Association of Applied Econometrics (Milan 2016) for useful comments. Financial support from the European Investment Bank Institute is gratefully acknowledged. The views expressed in this paper are those of the authors and do not necessarily reflect the views of De Nederlandsche Bank. Any errors are solely the authors' responsibility.

1. Introduction

Women show consistently low levels of financial literacy. They are less likely to answer simple financial knowledge questions correctly, they are more likely to answer "do not know" to those questions, and they rate themselves lower than men in terms of self-assessed financial literacy. This is true across countries and measures of financial knowledge, as well as across socio-demographic characteristics (see, e.g., Bucher-Koenen, Lusardi, Alessie, and Van Rooij, 2016, and OECD, 2013, for overviews). It is particularly striking that financial literacy levels seem to be low among young women who are well educated and have strong labor market attachment. Even women from an elite American college show considerable lack of financial expertise (Mahdavi and Horton, 2014).

The persistent gender gap in financial literacy may be the result of women feeling less confident in their financial knowledge and thus more inclined to answer "do not know." There is ample evidence that women are less confident than men in many situations, in particular in situations related to finance (see, e.g., Beyer, 1990; Barber and Odean, 2001). Some studies indicate that while men appear to be over-confident, women seem under-confident (see Dahlbom et al., 2011). In the context of financial knowledge, Chen and Volpe (2002) find that female college students are less confident and enthusiastic about financial topics. Webster and Ellis (1996) provide evidence that, even among financial experts, women show lower self-confidence in financial analyses compared to men.

This is consistent with the evidence provided by the self-assessed knowledge responses in our surveys, which shows that some of the women who respond with at least one "do not know" give themselves high knowledge assessments (see Bucher-Koenen, Lusardi, Alessie, and Van Rooij, 2016). Thus, irrespective of the fact that they have the inclination not to answer specific financial literacy questions, women still consider themselves financially competent. So the central question is, do those (women) who answer "do not know" know the answer but lack confidence in their knowledge?

In order to investigate this question, we design a simple experiment with the Dutch DNB household Survey (DHS). The objective is to understand what drives the gender gap in financial literacy and in particular what drives the gender difference in the "do not know" responses. Our first hypothesis is that by offering a "do not know" option among the multiple-

choice answers to the financial knowledge questions, we introduce noise in that other characteristics (specifically gender) that affect the propensity to reply with "do not know" enter the literacy measure. Specifically, we ran two surveys among the DHS respondents with a six weeks difference in between. In the first survey, we ask respondents the financial literacy questions with the option (as part of the multiple choice answers) of a "do not know" reply. We then follow these respondents over time and ask the same knowledge questions again, but this time taking away the "do not know" option and adding a follow-up question to assess how confident respondents are in their answers. These new set of data will allow us to dissect the answers to the financial literacy questions and examine the drivers of women's "do not know" responses. Our second hypothesis is that by improving the measurement of financial literacy we can estimate the effect of financial literacy on financial behavior more precisely and eliminate some of the bias plaguing those estimates.

Thus, the central contribution of this paper is that we develop a strategy, based on two survey waves, to consistently estimate whether the respondent truly knows the correct answers to the financial literacy questions. In doing so, we improve the measurement of financial literacy and can solve some of the problems existing in the current literature. Our main result is that women know less than men but they know more than they think they know. That is, if we take away the "do not know" option, women are very likely to give correct responses to the financial literacy questions. At the same time, women appear to be less confident in their answers. Thus, the gender gap in financial literacy is driven by both lower knowledge and lack of confidence. Our results have two implications: First, there should be financial education programs that are tailored to women. They should convey information as well as instill confidence in women of their knowledge and decision-making abilities. The second implication is methodological: when measuring financial literacy in surveys, researchers have to consider systematic bias induced by different response behavior.[1] We suggest alternative strategies to improve financial literacy measurement.

The paper proceeds as follows. In the next section we present the data and the experimental design. In section 3 we show descriptive results. In section 4 we propose a strategy for measuring financial literacy if there are differences in confidence that are heterogeneous

[1] This problem has already been widely discussed in the context of cross-national variation of self-reported health due to different answering scales and reporting styles (see, e.g., Kapteyn et al. 2007; Jürges 2007)

across gender. We explore different financial literacy measures in section 5 and present results for financial behavior in section 6. We conclude with a discussion of our results in section 7.

2. The data

2.1 The CentERpanel

We use data from the CentERpanel to investigate financial literacy and confidence among a representative set of Dutch-speaking households. The CentERpanel is an online household panel run by CentERdata, a survey agency at Tilburg University. Participants without internet connection are provided with the equipment enabling them to participate.[2] We include all panel members who are household heads and their partners in the sample. Respondents are age 18 and older. The data used in our study are collected between May and July 2012. We are able to merge our data with the DNB household survey (DHS). The DHS is an annual survey among the CentERpanel on income, assets and debt, work, health and economic and psychological concepts related to savings behavior.

2.2 The experiment

The experimental design is as follows. We ask the same three quiz-like questions on financial literacy to the same respondents twice (see Appendix A1 for the wording of the questions).[3] When we ask the questions for the first time in May 2012 respondents are offered "do not know" and "refuse to answer" options. When we ask the same questions for the second time about six weeks later at the end of June/beginning of July 2012 those options are deleted and respondents have to guess the answer if they do not know it. In this survey, respondents are required to rate the confidence they have in their answer on a scale from 1 – not confident at all to 7 – completely confident after each question.

2.3 The sample

In the first survey we have 1,748 and in the second survey we have 1,973 participants, including a refresher. For our main analysis we restrict the sample to the respondents who participate in both waves (balanced panel). We allow the household head and their partner to

[2] For more information, see www.centerdata.nl.

[3] These questions, also know as the *Big 3,* have been developed by Annamaria Lusardi and Olivia Mitchell (Lusardi and Mitchell 2011a). Since then they have been used widely to measure financial literacy in surveys around the world (see Lusardi and Mitchell 2011b and 2014 for overviews).

participate, thus for a number of households we have two individual observations (and in the regression analysis we compute standard errors which are clustered at the household level). We drop respondents who did not complete the literacy surveys (30 respondents; 1.35% of the initial raw sample). The reduced sample contains 1,532 respondents for all our analyses; 861 (56.2%) are men and 671 (43.8%) are women.[4]

Before we show our results we would like to make two important points based on the unrestricted, i.e. unbalanced, sample:

1. Attrition: We test for attrition between the waves conditional on financial literacy. Specifically, we look at the average number of correct answers in the first wave and partition the sample into those who participate only in the first wave and those who participate in both waves. We do not find a systematic difference in the average financial literacy of those groups. Thus, we conclude that respondents do not drop out systematically after the first survey because they are uncomfortable with answering the financial literacy questions. The same is true for attrition based on gender. Men and women both drop out after the first wave with equal probability.

2. Learning: Since we ask the same questions twice to the same respondents with only a six weeks difference one might be worried about learning effects. We can test for learning by comparing the probability to give correct answers in the second wave for the refresher sample who participate only in the second wave with the panel cases who participate in both waves. There is no significant difference in the answering behavior of those two groups in the second week. Thus, we feel confident that learning effects due to asking the same questions twice are not confounding our results.

3. Descriptive results

3.1 Comparing answers across waves

In table 1 we present the answers to all financial literacy questions for both the first and the second survey separately for men and women.[5]

[Table 1 - Tabulation of literacy responses in wave 1 and 2 - about here]

[4] The sample used in the regression analyses may vary slightly due to missing values for some control variables, especially when we merge our survey with the information from the DHS.

[5] The statistics presented in this paper are not weighted. We also used sampling weights but found only very small differences.

In the May survey for the interest question, men report more correct answers than women (91.9 % vs. 84.4%, see table 1 panel A). Thus the gender gap in giving the correct answer is around 7.5 percentage points. Women are more often incorrect, but more importantly they report a higher number of do not know (DK) answers. In the July survey we ask the same question without the DK option. The number of correct answers increases significantly to 94.7% for men and 91.2% for women. The number of incorrect answers also increases. However, overall the gender difference decreases to 3.5 percentage points. Note that the number of refusals is very limited. Hence, in the further analysis we lump this category together with the 'do not know' responses. If we condition the answers of wave 2 on the wave 1 responses, it is of particular interest how accurate the wave 2 responses are for those who stated do not know in wave 1 (see table 2). It appears that the majority of this group is able to provide the correct answer when forced to provide an answer, which suggests that they are not simply guessing the answer.[6] Around 70% of both men and women who said "do not know" in the first survey are able to correctly answer the interest question in the second survey.

[Table 2 - Tabulation of wave 2 responses conditional on wave 1 responses - about here]

The inflation question appears to be somewhat more difficult to answer. The number of correct answers is lower and the gender gap is larger at more than 9 percentage points (see table 1 panel B). Two thirds of the gender gap is driven by the DK's although also the number of incorrect answers is somewhat higher among women. When forced to answer, the gender gap diminishes from 9 to 6 percentage points. This is a result from the fact that the group that provides a DK answer is often able to provide the correct answer, when forced to make a choice.[7] Nevertheless, the men within the DK provide more often a correct answer when forced to make a choice (67% for men versus 62% for women; see Table 2 Panel B).

The third question relates to risk diversification. The proportion of DK's is high for both men and women, but especially for the latter group. More than half of the women report they do not know the answer (54.7 %) compared to 30.1% for men (see Table 1 Panel C). As a result, we measure a gender gap of 27.5 percentage points in the probability to give a correct answer

[6] We use a χ^2-test to test for random answering. Random answering is rejected at 0.1% significance.
[7] Random answering is rejected at 0.1% significance.

for this question. Strikingly, when forced to make a choice the gap shrinks to 9 percentage points. Both the majority of women and men who state DK appear able to answer the question correctly.[8] The proportion of correct is higher for men than women (72.6 versus 67.7 %; see Table 2 Panel C).

All in all, the probability to give a correct answer significantly increases for men and women after deleting the DK option. Panel D of table 1 shows the number of correctly answered questions. The probability of giving three correct answers increases from 58.1% to 74.9% for men and from 29.4% to 60.1% between the first and the second survey. The gender gap in financial literacy decreases by about half from almost 29 to around 15 percentage points. Conditional on responding with "do not know" in the first week both men and women are likely to give a correct answer in the second week for the three questions.

We confirm a gender gap for financial literacy. Partly, this is due to the fact that women more often state they do not know when given the option. When men and women are forced to answer, the gender gap decreases (but it does not disappear). This could be due to two reasons. First, those who say they do not know may actually signal that they are not absolutely sure about the correct answer, while at the same time have a high likelihood of being correct. Second, the gender gap may decrease simply because people really do not know but may provide the correct answer by chance. As the group of women stating do not know is larger the gender gap will also decrease because more women than men are forced to guess and thus also the number of additional correct answers will increase more for women than for men. Thus in the next section we would like to understand the relationship between answering behavior and confidence a bit better.

3.2 Confidence in financial literacy

As mentioned in the experimental design in the second survey (without the do not know option) after each of the three questions respondents evaluate how confident they feel about their answer. Evaluations are on a scale from 1 – not confident to 7 – completely confident. We report answers for all three questions separately for men and women in table 3. Overall we confirm that women are significantly less confident in the answers that they give to the financial literacy questions than men (see column "Total" for men and women). While among

[8] Random answering is rejected at 0.1% significance.

men a large fraction is very certain about giving the correct answer (ratings of 6 or 7), this is not true for women. They report much lower levels of confidence. Comparing the ratings for the three questions shows that respondents are fairly certain about their answers to the interest and inflation questions. What is a bit surprising is that ratings for the risk question are relatively low, even though many respondents give the correct response. Overall, the lower confidence ratings of women are consistent with the finding that women provide more often a DK answer.

[Table 3 - Confidence - about here]

We evaluate the confidence levels given after the second survey conditional on a respondent's answers to the same questions in the first survey. This allows us to see if those responding with DK in the first survey are less confident in their answer after the second survey, when they are forced to reply. The results of this exercise can be summarized as follows: Conditional on giving a correct answer in the first survey, women are significantly less confident than men in their answer in the second survey for all three questions. Thus, even when they give the correct answer women are not confident. For the more difficult risk question, conditional on giving an incorrect answer in the first wave women are significantly less confident in their answer in the second wave compared to men. Thus, even when they do not know men are more confident than women. The effect is not significant for the first two questions due to the small number of incorrect answers. Conditional on a DK answer in the first survey, women are much less confident in their reply in the second survey compared to men for the risk question. Again the effect is not significant for the first two questions due to the much lower number of DK responses. Finally, we ran regressions using DK responses to the questions as dependent variables and the confidence rating as well as various background characteristics as controls. There is a high correlation in the probability to answer with "do not know" in the first survey and the level of confidence in ones answer when forced to pick an option in the second survey for all three financial literacy questions.[9]

Summarizing, the financial literacy scores in May reflect both knowledge and confidence in answering. In July, respondents are forced to answer, providing a knowledge measure that is not confounded by confidence. However, at the same time people who do not know the

[9] In addition, lower educated and lower income respondents are more likely to choose the DK option in the third literacy question.

answer are forced to guess an answer, thus the July measure contains measurement error and is upward biased due to the responses of those who guess the correct answer. Thus, in the next section, we use information from both surveys and develop a latent class model in order to estimate 'true financial knowledge'.

4. Modeling true financial knowledge

The descriptive statistics show that the respondents and in particular women are often uncertain about their answer. Respondents seem to pick the 'do not know' option when they are not confident about their answer, even if they actually know the correct answer. This leads to a systematic bias in the measurement of financial literacy. On the other hand, sometimes respondents seem to pick an answer randomly. As these answers may be either correct or incorrect by chance, just counting the number of correct answers creates noisy knowledge measures. We need to disentangle 'true knowledge', 'confidence', and 'guessing' by respondents as to calculate a financial literacy index with minimal measurement error. For this purpose, we propose a measure of 'true financial knowledge' based upon the specific structure of the two surveys using respondents' confidence in their answers to correct for guessing.

First, we define for each of our three financial literacy questions the following latent variable for 'true knowledge':

$\tilde{y}_{ik} = 1$ if respondent i truly 'knows' the correct answer to literacy question k (k=1,2, 3),

$\tilde{y}_{ik} = 0$ otherwise;

Obviously, we do not observe $\tilde{y}_{ik}$, but we do observe some proxies for this variable: let y_{ik}^m be the individual's i answer to literacy question k in May (superindex m). Notice that y_{ik}^m can take on the following three values: 0 (incorrect answer), 1 (correct answer), 2 (do not know/refusal). Since the July questionnaire does not allow for a 'don't know' option, the variable y_{ik}^j (the answer to question k in July) can only take on the values 0 and 1. Instead of the don't know option, the July questionnaire contains for each literacy question a question which measures on a Likert scale (from 1 to 7) how confident the respondent is his/her answer. The variable $conf_{ik}^j$ is the answer to this question. Our goal is to use the information embodied in a vector of background characteristics x_i and in the variables y_{ik}^m, y_{ik}^j and $conf_{ik}^j$ to predict the probability that a respondent truly knows the answer to literacy question k. In

other words, for each respondent in our sample and for each of our three financial literacy questions, we want to estimate the following conditional probability:

$$P(\tilde{y}_{ik} = 1|x_i, y_{ik}^m = l_k, y_{ik}^j = m_k, conf_{ik}^j = z_k), k = 1,2,3 \quad (1)$$

Second, we construct a summary measure of financial literacy by adding up the probabilities of having true knowledge for the three individual financial literacy questions:

$$finlit_i = \sum_{k=1}^{3} P(\tilde{y}_{ik} = 1|x_i, y_{ik}^m = l_k, y_{ik}^j = m_k, conf_{ik}^j = z_k) \quad (2)$$

In the next subsection we present a so-called latent class model which can be used to estimate the probability (cf. equation 1) that the respondent truly knows the answer to literacy question k ($k = 1,2,3$).

4.1 The latent class model

Let $g_{ik} = 3 \cdot y_{ik}^j + y_{ik}^m$. In other words, $g_{ik}$is a random variable that summarizes the answers we observe in the May and July surveys. It can take on six different values: 0,…,5. For example, $g_{ik} = 0$ if a respondent answers incorrectly in both surveys, and $g_{ik} = 4$ if the respondent answers correctly in both surveys The log-likelihood of our latent class model is based on the conditional multinomial density of g_{ik}: $P\left(g_{ik} = g\middle|x_i, conf_{ik}^j = z_k\right)$. This conditional probability can be written as a weighted average of two multinomial probabilities $P\left(g_{ik} = g\middle|\tilde{y}_{ik} = 1, x_i, conf_{ik}^j = z_k\right)$, i.e. the probability to observe answering pattern $g_{ik} = g$ given true knowledge $\tilde{y}_{ik} = 1$, confidence level z_k, and individual socio-demographic characteristics x_i, and $P\left(g_{ik} = g\middle|\tilde{y}_{ik} = 0, x_i, conf_{ik}^j = z_k\right)$, i.e. the probability to observe answering pattern g_{ik} given a lack of true knowledge $\tilde{y}_{ik} = 0$, confidence level z_k, and individual socio-demographic characteristics x_i, where the probabilities for having true knowledge or not, i.e. $P\left(\tilde{y}_{ik} = 1\middle|x_i, conf_{ik}^j = z_k\right)$ and $P\left(\tilde{y}_{ik} = 0\middle|x_i, conf_{ik}^j = z_k\right)$, serve as weights:

$$\begin{aligned}
&P\left(g_{ik} = g\middle|x_i, conf_{ik}^j = z_k\right) = \\
&P\left(g_{ik} = g, \tilde{y}_{ik} = 1\middle|x_i, conf_{ik}^j = z_k\right) + P\left(g_{ik} = g, \tilde{y}_{ik} = 0\middle|x_i, conf_{ik}^j = z_k\right) = \\
&P\left(g_{ik} = g\middle|\tilde{y}_{ik} = 1, x_i, conf_{ik}^j = z_k\right)P\left(\tilde{y}_{ik} = 1\middle|x_i, conf_{ik}^j = z_k\right) + \\
&P\left(g_{ik} = g\middle|\tilde{y}_{ik} = 0, x_i, conf_{ik}^j = z_k\right)P\left(\tilde{y}_{ik} = 0\middle|x_i, conf_{ik}^j = z_k\right) = \\
&\alpha_g^1(x, z_k)P\left(\tilde{y}_i = 1\middle|x_i, conf_{ik}^j = z_k\right) + \alpha_g^0(x, z_k)P\left(\tilde{y}_i = 0\middle|x_i, conf_{ik}^j = z_k\right) \quad (3)
\end{aligned}$$

where the conditional multinomial probabilities are defined as

$\alpha_g^1(x, z_k) = P\left(g_{ik} = g\middle|\tilde{y}_{ik} = 1, x_i, conf_{ik}^j = z_k\right) = P\left(y_{ik}^m = l_{ik}, y_{ik}^j = m_{ik}\middle|\tilde{y}_{ik} = 1, x_i, conf_{ik}^j = z_k\right)$

$$\alpha_g^0(x, z_k) = P\left(g_{ik} = g \middle| \tilde{y}_{ik} = 0, x_i, conf_{ik}^j = z_k\right) = P\left(y_{ik}^m = l_{ik}, y_{ik}^j = m_{ik} \middle| \tilde{y}_{ik} = 0, x_i, conf_{ik}^j = z_k\right)$$

We assume in our econometric model that conditional upon background characteristics x_i true knowledge is independent of 'confidence', i.e.

$$P\left(\tilde{y}_{ik} = 1 \middle| x_i, conf_{ik}^j\right) = P(\tilde{y}_{ik} = 1 | x_i) \tag{4}$$

In other words, only the answers g_{ik} are influenced by confidence, but whether a respondent truly knows the correct answer or not is independent of confidence. In addition, we assume that probability (4) can be modeled by means of a probit specification, so that the conditional probability that respondent *i* truly knows the answer to literacy question *k* is equal to

$$P(\tilde{y}_{ik} = 1 | x_i) = \Phi(x_i'\beta) \tag{5}$$

In the empirical application we also assume that $\alpha_g^1(x_i, z_k) = \alpha_g^1(z_k)$ and $\alpha_g^0(x_i, z_k) = \alpha_g^0(z_k)$, thus the observed answering pattern depends on true knowledge and confidence but not on any additional background characteristics.[10] These two probabilities can be modeled by using a multinomial logit specification, where z_k represents a full set of seven dummy variables:[11]

$$\alpha_g^1(z_k; \gamma^1) = \frac{\exp(\gamma_g^{1\prime} z_k)}{\sum_{h=0}^5 \exp(\gamma_h^{1\prime} z_k)} \tag{6a}$$

$$\alpha_g^0(z_k; \gamma^0) = \frac{\exp(\gamma_g^{0\prime} z_k)}{\sum_{h=0}^5 \exp(\gamma_h^{0\prime} z_k)} \tag{6b}$$

where $\gamma^1 = (\gamma_0^1, \gamma_1^1, \ldots, \gamma_5^1)'$ and $\gamma^0 = (\gamma_0^0, \gamma_1^0, \ldots, \gamma_5^0)'$. Assumptions (4), (5) and (6) imply that the density described in (3) can be rewritten as follows:

$$P\left(g_{ik} = g \middle| x_i, conf_{ik}^j = z_k\right) = \alpha_g^1(z_k; \gamma^1)\Phi(x_i'\beta) + \alpha_g^0(z_k; \gamma^0)\Phi(-x_i'\beta) \tag{7}$$

We base the log-likelihood function on the density function (7). Notice that there is an identification problem: the parameter vector $(\gamma^{1\prime}, \gamma^{0\prime}, \beta')'$ is observationally equivalent with $(\gamma^{0\prime}, \gamma^{1\prime}, -\beta')'$ in the sense that they both result in the same probability distribution of observable data. In order to resolve this problem at the minimum we have to make one

[10] Confidence may depend on individual background characteristics though.

[11] We assume without loss of generality that $\gamma_4^1 = (0,0,0,0,0,0,0)'$ (i.e. for the "$\tilde{y}_{ik} = 1$ multinomial logit model" the reference group consists of those individuals who report a correct answer in both surveys, i.e.for which $y_{ik}^m = y_{ik}^j = 1$, i.e. $g = 1 \cdot 3 + 1 = 4$) and $\gamma_0^0 = (0,0,0,0,0,0,0)'$ (for the "$\tilde{y}_i = 0$ multinomial logit model" the reference group consists of those individuals who report an incorrect answer in both surveys, i.e. for which $y_i^m = y_i^j = 0$, i.e. $g = 0 \cdot 3 + 0 = 0$).

identifying assumption. We assume that a person who is 'fully confident and knowledgeable' about the answer to financial literacy question k ($\tilde{y}_i = 1, conf_{ik}^j = 7$), will not answer this question incorrectly twice, i.e. both in May and July.[12] Thus,

$$P\left(g_{ik} = 0 \middle| \tilde{y}_i = 1, conf_{ik}^j = 7\right) = P\left(y_i^m = 0, y_i^j = 0 \middle| \tilde{y}_i = 1, conf_{ik}^j = 7\right) = 0 \quad (8)$$

In other words, individuals who are financially knowledgeable and confident do not make the same reporting mistake twice.

However, in view of a fast convergence of the numerical likelihood optimization algorithms, we have made the following additional assumptions:

$\alpha_0^1(z_k) = P\left(g_{ik} = 0 \middle| \tilde{y}_{ik} = 1, conf_{ik}^j = z_k\right) = P\left(y_i^m = 0, y_i^j = 0 \middle| \tilde{y}_i = 1, conf_{ik}^j = z_k\right) = 0, z_k = 1,..,7$
$\alpha_1^1(z_k) = P\left(g_{ik} = 1 \middle| \tilde{y}_{ik} = 1, conf_{ik}^j = z_k\right) = P\left(y_i^m = 1, y_i^j = 0 \middle| \tilde{y}_i = 1, conf_{ik}^j = z_k\right) = 0, , z_k = 1,..,7$
$\alpha_2^1(z) = P\left(g_{ik} = 2 \middle| \tilde{y}_{ik} = 1, conf_{ik}^j = z_k\right) = P\left(y_i^m = 2, y_i^j = 0 \middle| \tilde{y}_i = 1, conf_{ik}^j = z_k\right) = 0, , z_k = 1,..,7$
$\alpha_3^1(z) = P\left(g_{ik} = 3 \middle| \tilde{y}_{ik} = 1, conf_{ik}^j = z_k\right) = P\left(y_i^m = 0, y_i^j = 1 \middle| \tilde{y}_i = 1, conf_{ik}^j = z_k\right) = 0, , z_k = 1,..,7$
$\alpha_4^0(z) = P\left(g_{ik} = 4 \middle| \tilde{y}_{ik} = 0, conf_{ik}^j = z_k\right) = P\left(y_i^m = 1, y_i^j = 1 \middle| \tilde{y}_i = 0, conf_{ik}^j = z_k\right) = 0, , z_k = 1, \dots,7$
$\alpha_5^0(z) = P\left(g_{ik} = 5 \middle| \tilde{y}_{ik} = 0, conf_{ik}^j = z_k\right) = P\left(y_i^m = 2, y_i^j = 1 \middle| \tilde{y}_i = 0, conf_{ik}^j = z_k\right) = 0, , z_k = 6,7$

Intuitively, these assumptions can be explained in the following way: First, regardless of the confidence level, if a respondent truly knows the answer to a financial literacy question, he will not pick a wrong answer twice. Second, conditional on true knowledge respondents are not answering inconsistently in both surveys, i.e. correct in May and incorrect in July or vice versa. Moreover, we also exclude the possibility that individuals pick a "do not know" response in May and answer incorrectly in July. Thus, given true knowledge the only possible answer patterns are to provide the correct answer twice or "do not know" in May and the correct answer in July. In other words, respondents who are truly knowledgeable do not randomly pick an answer or make mistakes.

The assumptions in line five and six refer to the structure we impose conditional on the lack of knowledge. Here we impose that given the respondent does not know the answer the probability to guess the correct answer twice (in May and July) is zero. The final assumption

[12] Notice that according to equation (6a)

$$\alpha_0^1(z_k) = \frac{\exp\left(\gamma_0^{1\prime} z_k\right)}{\sum_{h=0}^{5} \exp\left(\gamma_h^{1\prime} z_k\right)}$$

where $\gamma_h^1 = \left(\gamma_{h1}^1, \gamma_{h2}^1, \gamma_{h3}^1, \gamma_{h4}^1, \gamma_{h5}^1, \gamma_{h6}^1, \gamma_{h7}^1\right)'$, $h = 0, \dots,5$. Since the reference group consists of those individuals for which $= 4$, we can impose the condition (8) $P\left(g_{ik} = 0 \middle| \tilde{y}_i = 1, conf_{ik}^j = 7\right) = 0$ by a priori giving the parameter γ_{07}^1 a very small value. In the empirical application we impose the following restriction when we estimate the latent class model: $\gamma_{07}^1 = -22$. Assumptions mentioned in equations (9) are imposed in an analogous way.

is that given lack of knowledge, the probability to answer do not know in May and pick the correct answer in July is zero for those with high levels of confidence.

The estimation results of latent class model (7) are presented in Table … (THIS TABLE IS NOT ADDED YET). WE SHOULD DISCUSS THESE RESULTS

4.2 A summary estimate for respondents‘ financial literacy based on the latent class model

Once we have estimated the parameters we can compute for each financial literacy question the probability $P\left(\tilde{y}_{ik}=1\middle|g_{ik}=g,conf_{ik}^{j}=z_k,x_i\right)$ (cf. equation 1) as follows (see Vermunt, 2015):

$$P\left(\tilde{y}_{ik}=1\middle|g_{ik}=g,conf_{ik}^{j}=z_k,x_i\right)$$
$$=\frac{P\left(g_{ik}=g\middle|\tilde{y}_{ik}=1,conf_{ik}^{j}=z_k,x_i\right)P(\tilde{y}_{ik}=1|x_i)}{P\left(g_{ik}=g\middle|\tilde{y}_{ik}=1,conf_{ik}^{j}=z_k,x_i\right)P(\tilde{y}_{ik}=1|x_i)+P\left(g_{ik}=g\middle|\tilde{y}_{ik}=0,conf_{ik}^{j}=z_k,x_i\right)P(\tilde{y}_{ik}=0|x_i)}$$

$$=\frac{\alpha_g^1(z_k;\gamma^1)\Phi(x_i'\beta)}{\alpha_g^1(z_k;\gamma^1)\Phi(x_i'\beta)+\alpha_g^0(z_k;\gamma^0)\Phi(-x_i'\beta)} \tag{10}$$

Note that this probability can be thought of as the posterior probability of having true knowledge (our latent variable) which results after updating the prior probability using additional information from the two surveys (Bayes' rule).

Thus, for each respondent we can estimate the probability to truly know the correct answer to a given financial literacy question. Notice that this probability depends on the answers to the May and July questions (i.e. $g_{ik}=g$) and on the level of confidence reported by the respondent in July. The higher the estimated posterior probability the more knowledge the individual has. Notice also that the posterior distribution of $\tilde{y}_{ik}$ is degenerate if the following conditions are met:

$P\left(\tilde{y}_{ik}=1\middle|g_{ik}=g,conf_{ik}^{j}=z_k,x_i\right)=1$ if $\alpha_g^0(z_k;\gamma^0)=0$
$P\left(\tilde{y}_{ik}=1\middle|g_{ik}=g,conf_{ik}^{j}=z_k,x_i\right)=0$ if $\alpha_g^1(z_k;\gamma^0)=0$

Due to the 'identifying' assumptions (cf. equations 8 and 9) presented in the previous subsection, the posterior distribution of $\tilde{y}_{ik}$ is degenerate in many cases: irrespective of the confidence levels reported in July, $\tilde{y}_{ik}=0$ if a) respondents answer inconsistently over time (once correct, once incorrect), or b) two times incorrect, c) pick the "don't know" answer in May and incorrect answer in July. $\tilde{y}_{ik}=1$ if the respondents answers the literacy questions two times correctly (irrespective of the confidence level). Only in the case that a don't know answer is given in May and a correct one in July, the latent class model is used to predict

$P\left(\tilde{y}_{ik} = 1 \middle| g_{ik} = g, conf_{ik}^{j} = z_k, x_i\right)$ (cf equation 10). We compute an overall measure of the level of financial literacy for our respondents by summing up the estimated probabilities for the individual questions (cf. equation 2 and equation 10).

In the next section, we will compare the overall financial literacy measures for individual respondents based on the observations in May and July, respectively, and the results from the latent class model. We will then use all three measures to estimate the relationship between stock market participation and financial literacy.

5. Exploring different financial literacy measures

We present the different measures of financial literacy in Table 4. Column 1 presents the probability to observe a correct answer from the May questionnaire for each of the three financial literacy questions. As proposed previously this measure could underestimate financial knowledge since individuals with low confidence pick the do not know response even if they know the correct answer. On the other hand illiterate respondents could abstain from using the "do not know" option and just guess. Overall, the average number of correct answers to the three financial literacy questions is equal to 2.24.

[Table 4 – Alternative financial literacy measures - about here]

In column 2 we present the probability of observing a correct answer in July. Since all respondents have to answer the question there is no confounding with confidence, however there might be some random guessing. Thus, some individuals might guess the right answer without actually having the knowledge. Thus, this financial literacy measure is overestimating levels of 'true financial knowledge'. The comparison of column 1 and 2 has been discussed extensively in section 3. The average number of correct answers is equal to 2.62.

In column 3 we present our measure of 'true financial literacy' based upon the latent class model as defined in the previous section (cf. equation 2). The average value of 'true financial literacy' (2.38) takes on a value the May measure and the July measure.

To further investigate the three financial literacy measures, we run ordinary least squares regressions to display the relation between the different financial literacy measures and a number of background variables including gender, marital status, education and income. All financial literacy variables are standardized so that they have mean 0 and variance 1 which facilitates the comparison of the regression results across specifications. Table 5 reports the results.

Focusing on the gender differences in Table 5 Panel A, we can infer that the raw gender differential is largest for the May measure and smallest for the July measure. As women are less confident than men, they more often use the 'don't know option' than men. According to the July (May) measure, men answered on average 2.71 (2.44) questions correctly (out of 3) and women 2.52 (2.00) questions, i.e. a difference of 0.19 (0.44). Our measure for true knowledge based on the latent class model predicts an average of 2.50 out of three correct answers among men, and 2.22 for women. The resulting gender gap in financial literacy is 0.28. Thus, the gender gap in true knowledge predicted by the latent class model is smaller than the one based in the May questionnaire but larger than the one based on the July questionnaire.

Next, we include personal background variables to explain the variation in the literacy measures (Table 5 Panel B). The R^2 of the regressions range between 0.103 for the July measure and 0.167 for the May measure. The R^2 based on the predicted measure for true financial literacy is 0.154 and close to the May measure. Overall, the impression is that the explained variance in the July measure is lowest, because this measure has the greated measurement error due to random guessing.

With respect to the socio-demographic variables, the correlation patterns found with the different financial literacy measures are very similar: For all literacy measures we find that financial literacy is highest for the middle age categories and lowest for the younger (below 35 years) and older (above 65 years) respondents. While we cannot differentiate time and cohort effects based on our cross-section, this is consistent with a pattern of accumulating knowledge due to schooling and experience when young while the process of declining cognitive abilities start to dominate when old. This hump-shaped pattern is typically found in the empirical literature on age and knowledge accumulation (Agarwal et al., 2009, Lusardi and Mitchell 2011b).

[Table 5 – Multivariate regressions - about here]

Apart from age, marital status, education, income and gender contribute to the explanation of the variation in the measures of financial literacy as well. Singles (without children) and those with higher income and higher education display better scores. Single parents (predominantly divorced female respondents), however, display low literacy and are thus vulnerable to poor financial decision-making. For all measures of literacy, we still find that women score worse than men. However, not surprisingly, the gender differentials have become smaller due to the inclusion of socioeconomic background variables as women on average have lower education and income.

It appears that the education and income gradient are the strongest (weakest) for the May (July) measure. The higher educated/income are more confident/use less often the DK option in May than the lower educated/lower income respondents.[13] This is confirmed by a regression of the difference between the July and May literacy measures on background characteristics (not shown). Women, lower educated and lower income respondents display a larger improvement in literacy scores in July when forced to give an answer. Interestingly, the difference between the "true knowledge" and the May measure only depends on the female dummy: women fare worse in May while this is not the case for lower income and lower educated groups. This suggests that women state DK too frequently (if their knowledge is compared to men), but lower educated and lower income groups are correct to state they do not know.

Summarizing, the financial literacy scores in May reflect both knowledge and confidence in answering. In July, respondents are forced to answer, providing measure of knowledge unconfounded by confidence but plagued by measurement error. The "true financial literacy" measure minimizes both the measurement error and the bias due to confidence which in particular makes a difference for female respondents.

[13] See Footnote 10.

6. Estimating the effect of true knowledge on stock market participation

6.1 Ordinary least squares

The complications in measuring knowledge may not be innocuous for research on household economic decision-making. Our next step is to find out whether the different measures of literacy behave differently in estimating the effect of financial literacy on stock market participation. The objective is to check how our different measures of financial literacy perform in these estimations and what we can learn about the bias plaguing these estimates. The literature firmly documents an effect of financial literacy on economic outcomes. Financial literacy is empirically shown to increase stock market participation, planning for retirement and contribute to wealth accumulation. However, the evidence in this paper suggests that the traditional financial literacy measures employed in these studies jointly measure true knowledge and confidence. Therefore, the coefficients found in previous studies do not necessarily reflect the impact of true knowledge alone.

Below, we will investigate how the use of different measures of literacy impacts the association between financial literacy and stock market participation. We use stock market participation as economic outcome variables as this relationship has been extensively documented in the literature. First, we run a regression using the traditional measure of financial literacy (our May measure) and thereafter we compare the results with regressions based on alternative measures for financial literacy. In discussing the results, we focus on the literacy coefficient as well as the gender coefficient as the error in the traditional measure due to differences in confidence is shown to be related to gender.

Financial literacy has been shown to influence stock market participation previously (see, e.g., Van Rooij et al., 2011). We define a dummy for stock market participation that equals 1 if the respondents hold investments in stocks and/or mutual funds and 0 zero otherwise. There is a strong negative correlation between gender and stock market participation: 33.9% of men in our sample own stocks and 20.3% of women (Table 6; column 1). If we control for the usual background characteristics and the traditional financial literacy measure (May), we find a strong association between financial literacy and stock market participation, while the gender effect becomes much smaller but is still significant (column 2). Compared to men, women have a 4.61 percentage point lower chance to own stocks after controlling for the usual background information including income, education, financial literacy etc. A one standard deviation higher level of literacy results into a 9.01 percentage point higher probability to own

stocks (comparable to the effect found in the literature). While this is a sizeable effect, this coefficient may reflect both confidence as well as knowledge.

Next, we run a regression using the financial literacy measure from July which should be unconfounded by confidence (column 3). While still significant the literacy effect reduces to a 5.49 percentage point higher likelihood of investing in the stock market for a 1 standard deviation higher level of literacy. Note that the female coefficient becomes more negative compared to the effect on column 2 as it is now likely to pick up part of the confidence effect; women being less confident have a lower chance to invest in stocks. The July measure for financial literacy is surrounded with measurement error due to guessing by respondents who are obliged to provide an answer. As a result, the literacy coefficient may be biased towards zero. Indeed, once we use the predicted measure of true financial literacy, the literacy coefficient is somewhat higher (column 4). We estimate a 6.71 percentage point higher likelihood of investing in the stock market for a 1 standard deviation higher level of literacy. The difference with the coefficient for the July literacy measure in column 2 is statistically significant.

6.2 Instruments for financial literacy

Knowledge may increase as a result of investing in the stock market. Investors for instance are likely to gather information before they buy or sell stocks and mutual funds and will more closely follow the stock market than non-investors. Thus, one cannot give a causal interpretation to the positive financial literacy coefficient in the OLS regressions of stock market participation. Below, we report the results of regressions similar to the previous regressions but now based on GMM models using financial education in high school as instruments for financial literacy to identify the causal effect of financial literacy on financial behavior and to reduce measurement error in the literacy variable. The instruments we are using are similar to the instruments used in Van Rooij et al. (2012) and are based upon information on exposure to economic education when young. First, respondents are asked how much attention has been paid to economics during their high school education. The difference to the question used previously is that we specifically refer to high school which makes the instrument more precise. Second, respondents report if economics was part of their final high school exam.

We measure exposure to education before entering the job market using the responses to the questions 'How much of your education in high school was devoted to economic subjects?' with the following answer categories: 'a lot', 'some', 'little', 'hardly at all', 'not applicable, I did not complete high school', 'do not know' or 'refuse to answer'. We distinguish three groups. The first group consists of respondents who did not get economics in high school answering 'hardly at all' or 'not applicable'. This is the reference group in our empirical analysis. Second, based on the 'a lot', 'some' and 'little' responses we create a dummy variable for respondents who were exposed to economics during high school. The third group consists of those who answered with 'don't know' or ' refusal' (very few respondents refused to answer this question). The instruments have high predictive power for financial literacy as shown by the *F*-values in the first stage regression which are mostly above 10 (cf. columns 2, 4 and 6 of Table 7).

Unless respondents indicate they did not complete high school, they receive the next follow-up question: 'Did you have at least one economics subject in your final examination year?' with the response options 'yes', 'no', 'not applicable, I didn't do a final exam', 'do not know' or 'refuse to answer'. We create an additional instrument dummy variable that takes the value 1 for those respondents who answer 'yes' and the value 0 otherwise. When we include this variable in the instrument set, we obtain *F*-values close to 10 for the July measure and in excess of 10 (which serves as the recommended threshold value to avoid weak instruments problems in the literature, see Staiger and Stock (1997)) for the other measures (cf. columns 1, 3 and 5 of Table 7).[14] One may argue, however, that the third instrument dummy is not valid as for some students the economic subject in their final exam may have been a choice variable and thus is likely to be correlated with interest in financial matters (interest in financial matters is an omitted variable in our regression) which in turn may affect financial decision making. Therefore, we present the results including and the results excluding this variable in the information set to instrument financial literacy.

Table 7 presents the GMM results for stock market participation. Both sets of instruments predict the endogenous financial literacy variable reasonably well. We interpret this as another sign that the July measure contains considerable measurement error which makes it more difficult to find valid instruments. The Hansen J test results indicate that the

[14] Table 8 reports the results of the first stage regressions.

overidentifying restrictions cannot be rejected in any of the specifications. The GMM C tests (see Hayashi, 2000) show mixed results for stock market participation. Using the extended set of instruments, the test suggests that financial literacy is endogenous to stock market participation while using the smaller set of instruments it cannot be rejected that financial literacy is exogenous. The latter result is consistent with previous findings (Van Rooij et al., 2011a).

Focusing on the effect of financial literacy on stock market participation, we find that the GMM estimate of the literacy coefficient is statistically significant at the 5 percent level and relatively similar across specifications (around 0.20). This seems a comforting result as apparently the instruments take care of the measurement error and the differences between the literacy measures become less important when good instruments are available. However, finding good instruments is easier for more accurate measures. Note that the predictive value of the instruments is lowest for the July measure which translates into a less precise estimate for the GMM literacy coefficient. The gender effect is insignificant in all specifications, which suggests that once literacy and socio-demographic variables are controlled for females are as likely to invest in stocks as men. We have also run the GMM regressions on the extensions discussed in section 7.2 (not reported). The main conclusions are not affected.

7 How to measure financial literacy?

From a survey methodology point of view, our findings establish that financial literacy is best measured by combining the two surveys and a confidence measure in order to estimate true knowledge. In practical applications, having two surveys among the same group of respondents is not always feasible or is simply less attractive in terms of the available research budget. A possible way out is to combine the two surveys into a single survey. The researcher may for instance first provide the literacy question including the DK option. If the respondent chooses the DK options, the same question is offered but excluding the DK option. After providing an answer, either the first or the second time, the respondent is asked for his or her confidence in the answer chosen. However, when a number of financial literacy questions are subsequently offered according to this scheme respondents will learn that they are required to answer the same question when answering DK. As a behavioral response respondents may want to avoid the DK option and answer the literacy questions right away which takes away the value added of the DK option for the researcher. Moreover, this methodology does not capture the information that was retrieved from the inconsistent answers in the two surveys.

Another disadvantage of this approach is that one needs to include three questions in the survey design to learn about the response on one literacy question, which may be problematic if there are constraints to the length of the survey in terms of costs or the duration of filling in the survey.

Alternatively, one could base a literacy measures solely on the information content as contained in the second questionnaire that we fielded in our experiment. This requires two responses per literacy question. A disadvantage is that we lose the information content of inconsistent answering. However, the number of inconsistent responses is limited and the DK answers show a strong correlation with the confidence questions (see Table 3). The extent to which our combined measure based on two surveys is better able to capture true knowledge than a measure using the questions in the second survey only (both the literacy question plus the confidence question) is an empirical question. Below, we will investigate the difference between these two approaches. First, we construct a measure for literacy based upon the literacy and confidence questions in wave two only (as discussed in Section 4.2). Basically, we assume that respondents indicating they are very unsure about their answers are not knowledgeable even if they guess the answer correctly. The analysis in Table 4 shows that this measure is closely related to the measure based on both survey waves. In addition the regression results for stock market participation show that the literacy coefficient becomes somewhat smaller in the OLS regression (compare column 5 and 4 in Table 6). This is consistent with the fact that the measure employing both surveys is better able to filter out guessing and thus has less measurement error. However, note that the measurement error is much lower than in using only the July literacy questions and not the confidence questions (compare column 5 and 2 in Table 6). The literacy coefficients are economically and statistically significantly higher than for the July measure. Overall the results for the alternative methods of measurement based on only the July answers plus confidence are quite similar to those based upon the measure combining both surveys. These findings show that the alternative measure might provide an adequate proxy for true knowledge as measured by the combined measure while it is much easier to integrate in new research designs and with less costs.[15]

[15] In principle one may consider to include a whole lot of literacy questions as well. The distorting effect of measurement error due to random guessing is likely to diminish when respondents have to guess many questions. However, this strategy would most likely require so many questions that it is either not feasible to implement or very costly in terms of question load.

8. Concluding remarks

The literature has documented large and robust gender differences in financial literacy. For example, Lusardi and Mitchell (2011b) find that 22.5% of female respondents in the US were able to answer three simple questions on inflation, interest and risk diversification correctly versus 38.3% of the male respondents. These findings are robust across different surveys and different countries (Lusardi and Mitchell, 2011a). This is especially worrisome as women who tend to outlive their husband are at risk of being left on their own in managing their financial security after retirement (Lusardi and Mitchell, 2008). We find that the gender gap diminishes once we force women to answer as they are more likely to state they do not know the answer otherwise. The higher propensity to choose DK is related to a lack of confidence in knowledge. Our results show that conditional on their level of knowledge, women are less confident than men. The gender gap diminishes significantly once we correct the traditional financial literacy measures to get improved measures for knowledge, but the gap does not disappear. By and large, we find that half of the gender gap cannot be explained by confidence and other background variables such as income and formal schooling.

Our findings have important implications both methodologically and for the interpretation of economic research and related policy advice. Traditional literacy measures, including a do not know option, capture both confidence and knowledge. This has consequences for the gender coefficient when both gender and literacy are included in regressions explaining financial decisions. Thus one needs to be careful in interpreting literacy and gender effects in this type of regressions. When respondents are forced to answer, literacy measures are not contaminated by confidence and better able to proxy true knowledge. However, this introduces measurement error as respondents who are not knowledgeable are forced to guess the correct answer. We propose an adjusted metric to measure pure knowledge which suffers less from measurement error and show that the different literacy measures have different impact on financial decisions. Nevertheless for economic decision-making both knowledge and confidence are important, so that the traditional measures that combine knowledge and confidence are better able to explain and predict the variation in household financial decisions. In terms of policy interventions it is crucial to disentangle true knowledge from confidence. For example, for an effective design of initiatives raising financial education and increasing awareness it is important to figure out whether limited knowledge or low confidence explains low stock market participation. Using our improved measure, we confirm

that true financial knowledge contributes to explain the observed heterogeneity in important household financial decisions as illustrated for investing in the stock market.

References

Agarwal, S., Driscoll, J., Gabaix, X. and Laibson, D. (2009). The age of reason: Financial decisions over the life cycle and implications for regulation, *Brookings Papers on Economic Activity*, Fall 2009, pp. 51-117.

Alessie, R., Van Rooij, M. and Lusardi, A. (2011). Financial literacy and retirement preparation in the Netherlands, *Journal of Pension Economics and Finance*, vol. 10(4), pp. 527–546.

Ameriks, J., Caplin, A. and Leahy, J. (2003). Wealth accumulation and the propensity to plan, *Quarterly Journal of Economics*, vol. 118(3), pp. 1007–1047.

Barber, B. and Odean, T. (2001), "Boys will be boys: gender, overconfidence and common stock investment", *Quarterly Journal of Economics*, February 2001, pp. 261-292.

Beyer, Sylvia (1990), "Gender differences in the accuracy of self-evaluations of performance", *Journal of Personality and Social Psychology*, 59(5), pp. 960-970.

Bucher-Koenen, T., Lusardi, A., Alessie, R. And Van Rooij, M. (2014), "How financially literate are women? An overview and new insights", NBER Working Paper No. 20793.

Chen, Haiyang and Ronald P. Volpe (2002), "Gender differences in personal financial literacy among college students", *Financial Services Review*, 11, pp. 289-307.

Dahlbom, L., A. Jakobsson, N. Jakobsson, and A. Kotsadam (2011), "Gender and overconfidence: are girls really overconfident?", *Applied Economics Letters*, 18, 325-327.

Kapteyn, A., and Teppa F. (2011). Subjective measures of risk aversion, fixed costs, and portfolio choice, *Journal of Economic Psychology*, vol. 32(4), pp. 564–580.

Lusardi, A. and Mitchell, O. (2007). Baby boomers retirement security: the role of planning, financial literacy and housing wealth, *Journal of Monetary Economics*, vol. 54(1), pp. 205–224.

Lusardi, Annamaria, and Olivia S. Mitchell (2008). Planning and financial literacy: How do women fare?, *American Economic Review*, 98(2), pp. 413-417.

Lusardi, Annamaria and Olivia S. Mitchell (2011a), "Financial literacy and planning: Implications for retirement wellbeing", in Olivia S. Mitchell and Annamaria Lusardi (eds.), *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*. Oxford: Oxford University Press, pp. 17-49.

Lusardi, A. and Mitchell, O. (2011b). Financial literacy around the world: an overview, *Journal of Pension Economics and Finance*, vol. 10(4), pp. 497–508.

Lusardi, A. and Mitchell, O. (2011c). Financial literacy and retirement planning in the United States", *Journal of Pension Economics and Finance*, 10(4), pp. 509-525.

Mahdavi, M. and Horton, N. (2014), Financial Knowledge among Educated Women: Room for Improvement, *Journal of Consumer Affairs*, 48 (2), pp. 403–417.

OECD (2013), Women and Financial Education: Evidence, Policy Responses and Guidance, OECD Publishing. http://dx.doi.org/10.1787/9789264202733-en

Staiger, D. and Stock, J. (1997). Instrumental variables regression with weak instruments, *Econometrica*, vol. 65(3), pp. 557–86.

Van Rooij, M., Lusardi, A. and Alessie, R. (2011a). Financial literacy and stock market participation, *Journal of Financial Economics*, vol. 101(2), pp. 449–472.

Van Rooij, M., Lusardi, A. and Alessie R. (2011b). Financial literacy and retirement planning in the Netherlands, *Journal of Economic Psychology*, 32, 593-608.

Webster, Robert L. and T. Selwyn Ellis (1996), "Men's and women's self-confidence in performing financial analysis," *Psychological Reports*, 79, pp. 1251-1254.

Appendix

A1. Financial Literacy Questions

1. Set Up Week 1 (May 2012):

1. **Interest:** *Suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow? More than $102 / Exactly $102 / Less than $102 / Do not know/ Refuse to answer*
2. **Inflation:** *Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account? More than today / Exactly the same / Less than today / Do not know / Refuse to answer*
3. **Risk:** *Please tell me whether this statement is true or false. "Buying a single company's stock usually provides a safer return than a stock mutual fund." True / False / Do not know / Refuse to answer*

2. Set Up Week 2 (July 2012):

Questions 1 to 3 without the "Do not know" and "refuse to answer" options

After each question – **Confidence**:
On a scale from 1 to 7, How confident are you in this answer? 1-not confident at all ... 7-completely confident

Figures and Tables

Table 1: Descriptive Results (Observations: Men – 861, Women – 671, Total - 1532)

	Survey 1: May 2012			Survey 2: July 2012		
A. Interest:						
	Men	Women	Total	Men	Women	Total
More than 102 euro	91.9	84.4	88.6	94.7	91.2	93.2
Exactly 102 euro	3.0	4.0	3.5	3.7	6.0	4.7
Less than 102 euro	2.0	3.9	2.8	1.6	2.8	2.2
Do not know	2.8	6.7	4.5	-	-	-
Refuse	0.4	1.0	0.7	-	-	-
Total	100	100	100	100	100	100
B. Inflation:						
	Men	Women	Total	Men	Women	Total
More	2.1	2.4	2.2	2.2	2.7	2.4
Exactly the same	3.3	5.4	4.2	4.1	9.8	6.6
Less	89.8	80.6	85.8	93.7	87.5	91.0
Do not know	4.7	10.7	7.3	-	-	-
Refuse	0.2	0.9	0.5	-	-	-
Total	100	100	100	100	100	100
C. Risk Diversification:						
	Men	Women	Total	Men	Women	Total
Incorrect 'right'	7.6	9.7	8.5	17.7	27.0	21.7
Correct 'false'	61.9	34.4	49.9	82.4	73.0	78.3
Do not know	30.1	54.7	40.9	-	-	-
Refuse	0.5	1.2	0.8	-	-	-
Total	100	100	100	100	100	100
D. Overall No. of correct answers:						
	Men	Women	Total	Men	Women	Total
0	3.6	6.6	4.9	0.5	0.8	0.6
1	7.3	16.8	11.5	3.3	6.9	4.8
2	31.0	47.2	38.1	21.4	32.3	26.2
3	58.1	29.4	45.5	74.9	60.1	68.4

Table 2: Answers in July (Wave 2) conditional on answers in May (Wave 1) (Observations: Men – 861, Women – 671, Total – 1532)

A. Interest:		**Men**			**Women**	
Survey May	**incorrect**	**correct**	**don't know**	**incorrect**	**correct**	**don't know**
Survey July						
incorrect	23.26	3.54	29.63	28.3	4.95	30.77
correct	76.74	96.46	70.37	71.7	95.05	69.23
Total	100	100	100	100	100	100
B. Inflation:						
incorrect	41.3	2.72	33.33	30.77	7.02	38.46
correct	58.7	97.28	66.67	69.23	92.98	61.54
Total	100	100	100	100	100	100
C. Risk Diversification:						
incorrect	38.46	10.32	27.38	47.69	12.55	32.27
correct	61.54	89.68	72.62	52.31	87.45	67.73
Total	100	100	100	100	100	100

Table 3: Confidence and answering behavior in May

	Men (N = 861)				Women (N = 671)			
A. Interest	**incorrect**	**correct**	**DK**	**Total**	**Incorrect**	**correct**	**DK**	**Total**
completely unconfident - 1	7.0	1.9	3.7	2.2	3.8	1.4	15.4	2.7
2	2.3	0.9	3.7	1.1	3.8	0.9	5.8	1.5
3	7.0	0.9	7.4	1.4	1.9	1.2	13.5	2.2
4	11.6	1.8	33.3	3.3	18.9	5.5	23.1	7.9
5	4.7	2.3	18.5	2.9	17.0	5.8	11.5	7.2
6	27.9	7.1	11.1	8.3	17.0	19.3	15.4	18.8
completely confident - 7	39.5	85.2	22.2	81.0	37.7	65.9	15.4	59.8
Total	100	100	100	100	100	100	100	100
B. Inflation								
completely unconfident – 1	6.5	1.7	4.8	2.1	7.7	2.2	15.4	4.2
2	10.9	0.7	7.1	1.5	7.7	1.9	15.4	3.9
3	6.5	1.6	9.5	2.2	5.8	4.1	11.5	5.1
4	13.0	2.7	40.5	5.1	17.3	9.8	30.8	12.8
5	15.2	4.8	19.1	6.0	25.0	14.1	15.4	15.1
6	15.2	9.8	4.8	9.9	11.5	19.8	5.1	17.4
completely confident – 7	32.6	78.8	14.3	73.2	25.0	48.2	6.4	41.6
Total	100	100	100	100	100	100	100	100
C. Risk								
completely unconfident – 1	1.5	2.1	6.5	3.4	4.6	3.9	13.9	9.5
2	1.5	0.8	9.1	3.4	6.2	4.8	11.5	8.6
3	0.0	2.6	8.8	4.3	15.4	5.2	13.3	10.7
4	24.6	8.6	34.2	17.7	23.1	23.4	32.3	28.3
5	27.7	20.5	17.1	20.0	33.9	25.1	17.9	21.9
6	20.0	22.7	12.9	19.5	7.7	21.7	7.5	12.4
completely confident – 7	24.6	42.8	11.4	31.8	9.2	16.0	3.7	8.5
Total	100	100	100	100	100	100	100	100

Table 4: Alternative Financial Literacy Measures (N= 1528)

	I Prob (May)	II Prop (July)	III true knowledge
Interest	88.58%	93.15%	87.60%
Inflation	85.77%	90.99%	86.26%
Risk	49.87%	78.26%	64.01%
Financial literacy measure	2.24	2.62	2.38

Col I and II: Fin. literacy measure= no. of correct answers

Col III: Fin.l literacy measure= $\sum_{k=1}^{3} P(\tilde{y}_{ik} = 1 | x_i, y_{ik}^m = l_k, y_{ik}^j = m_k, conf_{ik}^j = z_k)$

Table 5, Panel A: Explaining finanacial literacy (N=1532)

VARIABLES	(1) May	(2) July	(3) „true knowledge“
female	-0.441***	-0.189***	-0.282***
	(0.0386)	(0.0291)	(0.0351)
Constant	2.438***	2.708***	2.504***
	(0.0266)	(0.0187)	(0.0237)
Observations	1,528	1,528	1,528
R-squared	0.068	0.024	0.035

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

may:= may measure financial literacy

july:= july measure financial literacy

ytilde:= financial literacy measure based on the latent class model

Table 5, Panel B: Explaining financial literacy

VARIABLES	(1) May	(2) July	(3) True knowledge
female	-0.359***	-0.146***	-0.221***
	(0.0393)	(0.0301)	(0.0360)
Marital status (ref. Single)			
married, no child	-0.0455	-0.0977**	-0.158***
	(0.0678)	(0.0444)	(0.0537)
married, child	-0.0617	-0.111**	-0.205***
	(0.0751)	(0.0518)	(0.0633)
single parent, other	-0.264**	-0.293***	-0.388***
	(0.131)	(0.106)	(0.113)
Age (ref. <=35)			
36-50	0.229**	0.162**	0.260***
	(0.110)	(0.0748)	(0.0868)
51-65	0.207*	0.174**	0.236***
	(0.110)	(0.0749)	(0.0855)
>65	0.166	0.0999	0.133
	(0.115)	(0.0808)	(0.0908)
Education level (ref. 'primary education)			
lower secondary VMBO	0.255**	0.0278	0.0535
	(0.130)	(0.114)	(0.119)
upper secondary:MBO	0.219	0.113	0.0973
	(0.134)	(0.117)	(0.123)
upper secondary: HAVO/VWO	0.512***	0.253**	0.394***
	(0.133)	(0.119)	(0.122)
Tertiary: HBO	0.435***	0.264**	0.340***
	(0.130)	(0.115)	(0.120)
Tertiary: University	0.662***	0.399***	0.546***
	(0.135)	(0.119)	(0.124)
Income quartiles (reirst quartile)			
1902<x<=2600	0.259***	0.0689	0.170***
	(0.0716)	(0.0509)	(0.0602)
2600<x<=3471	0.353***	0.124**	0.257***
	(0.0759)	(0.0529)	(0.0634)
x>3471	0.444***	0.173***	0.348***
	(0.0776)	(0.0567)	(0.0661)
refuse/dk	0.103	0.145	0.192
	(0.237)	(0.135)	(0.175)
Constant	1.629***	2.367***	1.991***
	(0.170)	(0.135)	(0.145)
Observations	1,528	1,528	1,528
R-squared	0.167	0.103	0.154

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6 Stock market participation (OLS results)

VARIABLES	(1)	(2) May	(3) July	(4) True financial literacy
Financial literacy		0.0901***	0.0549***	0.0671***
		(0.0105)	(0.00970)	(0.0101)
female	-0.136***	-0.0461**	-0.0715***	-0.0646***
	(0.0207)	(0.0212)	(0.0213)	(0.0213)
Marital status (ref. Single)				
married, no child		-0.0943***	-0.0896***	-0.0840***
		(0.0320)	(0.0326)	(0.0324)
married, child		-0.123***	-0.119***	-0.110***
		(0.0371)	(0.0375)	(0.0373)
single parent, other		-0.132**	-0.133**	-0.125**
		(0.0548)	(0.0557)	(0.0556)
Age (ref. <=35)				
36-50		0.139***	0.149***	0.140***
		(0.0475)	(0.0473)	(0.0472)
51-65		0.202***	0.207***	0.203***
		(0.0465)	(0.0460)	(0.0460)
>65		0.201***	0.209***	0.207***
		(0.0496)	(0.0490)	(0.0491)
Education level (ref. 'primary education)				
lower secondary VMBO		-0.0872*	-0.0623	-0.0648
		(0.0510)	(0.0516)	(0.0519)
upper secondary:MBO		-0.0237	-0.0108	-0.00923
		(0.0555)	(0.0564)	(0.0566)
upper secondary: HAVO/VWO		-0.0659	-0.0338	-0.0464
		(0.0589)	(0.0595)	(0.0599)
Tertiary: HBO		-0.00615	0.0166	0.00993
		(0.0557)	(0.0567)	(0.0569)
Tertiary: University		0.141**	0.174***	0.163***
		(0.0618)	(0.0621)	(0.0624)
Income quartiles (reirst quartile)				
1902<x<=2600		0.0470	0.0682**	0.0591*
		(0.0305)	(0.0311)	(0.0309)
2600<x<=3471		0.0958***	0.122***	0.110***
		(0.0352)	(0.0351)	(0.0352)
x>3471		0.183***	0.214***	0.198***
		(0.0381)	(0.0385)	(0.0385)
9.quart_nettohh		0.198**	0.198*	0.192*
		(0.0950)	(0.103)	(0.0997)
Constant	0.339***	0.143*	0.101	0.112
	(0.0161)	(0.0753)	(0.0731)	(0.0741)
Observations	1,532	1,532	1,532	1,532
R-squared	0.023	0.147	0.126	0.132

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 7: Stock market participation: GMM results

	May		July		True knowledge	
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Financial literacy	0.192***	0.166**	0.222***	0.185*	0.204***	0.169**
	(0.0671)	(0.0815)	(0.0842)	(0.0947)	(0.0751)	(0.0801)
female	-0.00335	-0.0142	-0.0310	-0.0399	-0.0235	-0.0342
	(0.0369)	(0.0415)	(0.0308)	(0.0321)	(0.0325)	(0.0331)
Marital status (ref. Single)						
married, no child	-0.0901***	-0.0907***	-0.0627*	-0.0681*	-0.0555	-0.0621*
	(0.0328)	(0.0324)	(0.0364)	(0.0363)	(0.0364)	(0.0364)
married, child	-0.115***	-0.115***	-0.0865**	-0.0919**	-0.0724*	-0.0806*
	(0.0379)	(0.0374)	(0.0416)	(0.0412)	(0.0421)	(0.0420)
single parent, other	-0.0998*	-0.106*	-0.0511	-0.0681	-0.0549	-0.0713
	(0.0601)	(0.0601)	(0.0739)	(0.0740)	(0.0694)	(0.0689)
Age (ref. <=35)						
36-50	0.111**	0.118**	0.106*	0.115**	0.0940	0.105*
	(0.0545)	(0.0549)	(0.0576)	(0.0569)	(0.0577)	(0.0569)
51-65	0.176***	0.183***	0.161***	0.170***	0.163***	0.172***
	(0.0523)	(0.0524)	(0.0563)	(0.0557)	(0.0541)	(0.0532)
>65	0.181***	0.186***	0.183***	0.188***	0.185***	0.189***
	(0.0545)	(0.0539)	(0.0558)	(0.0542)	(0.0537)	(0.0524)
Education level (ref. 'primary education)						
lower secondary VMBO	-0.119**	-0.110*	-0.0699	-0.0664	-0.0762	-0.0716
	(0.0563)	(0.0577)	(0.0587)	(0.0559)	(0.0559)	(0.0542)
upper secondary:MBO	-0.0530	-0.0459	-0.0432	-0.0355	-0.0281	-0.0229
	(0.0593)	(0.0600)	(0.0656)	(0.0635)	(0.0608)	(0.0593)
upper secondary: HAVO/VWO	-0.128*	-0.112	-0.103	-0.0862	-0.120	-0.0996
	(0.0716)	(0.0770)	(0.0741)	(0.0745)	(0.0735)	(0.0742)
Tertiary: HBO	-0.0615	-0.0466	-0.0562	-0.0389	-0.0537	-0.0362
	(0.0669)	(0.0716)	(0.0729)	(0.0736)	(0.0693)	(0.0696)
Tertiary: University	0.0622	0.0826	0.0652	0.0905	0.0654	0.0910
	(0.0812)	(0.0885)	(0.0880)	(0.0912)	(0.0846)	(0.0860)
Income quartiles (reirst quartile)						
1902<x<=2600	0.0145	0.0216	0.0485	0.0518	0.0289	0.0355
	(0.0364)	(0.0382)	(0.0346)	(0.0337)	(0.0356)	(0.0351)
2600<x<=3471	0.0540	0.0632	0.0888**	0.0938**	0.0673	0.0760*
	(0.0442)	(0.0467)	(0.0405)	(0.0399)	(0.0428)	(0.0425)
x>3471	0.130**	0.143**	0.166***	0.176***	0.134***	0.150***
	(0.0508)	(0.0558)	(0.0474)	(0.0479)	(0.0518)	(0.0530)
refuse/dk	0.190**	0.192**	0.179*	0.185*	0.168*	0.174*
	(0.0869)	(0.0881)	(0.0998)	(0.0992)	(0.0907)	(0.0922)
Constant	0.220**	0.200**	0.170*	0.155*	0.181**	0.163*
	(0.0933)	(0.0985)	(0.0875)	(0.0867)	(0.0873)	(0.0865)
Observations	1,532	1,532	1,532	1,532	1,532	1,532
R-squared	0.104	0.123	0.002	0.051	0.053	0.089
F stat first stage	14.19	11.12	9.189	7.694	11.26	12.15
chi2-value Hansen overid test	1.193	0.928	0.800	0.313	0.988	0.0401
p-value Hansen overid test	0.551	0.335	0.670	0.576	0.610	0.841
p-value GMM C exogeneity test	0.111	0.332	0.0255	0.121	0.0476	0.170
pval lm test beta_lit=0	0.00514	0.0533	0.00377	0.0332	0.00401	0.0261
pval ar test beta_lit=0	0.0233	0.0813	0.0233	0.0813	0.0233	0.0813

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 8: First Stage regression

VARIABLES	(1) May	(2) July	(3) True Knowledge
High school education devoted to economics (ref: not applicable, hardly at all)			
little, some, a lot	0.0935*	0.0534	0.0482
	(0.0567)	(0.0442)	(0.0495)
Do not know, refusal	-0.274**	-0.234**	-0.393***
	(0.125)	(0.0999)	(0.107)
Economics in high school exam? 1=yes	0.186***	0.102***	0.109***
	(0.0465)	(0.0321)	(0.0398)
Female	-0.350***	-0.141***	-0.215***
	(0.0394)	(0.0300)	(0.0359)
Marital status (ref. Single)			
married, no child	-0.0379	-0.0931**	-0.152***
	(0.0667)	(0.0438)	(0.0530)
married, child	-0.0781	-0.122**	-0.219***
	(0.0739)	(0.0511)	(0.0624)
single parent, other	-0.249*	-0.283***	-0.374***
	(0.132)	(0.106)	(0.113)
Age (ref. <=35)			
36-50	0.228**	0.163**	0.264***
	(0.106)	(0.0732)	(0.0848)
51-65	0.245**	0.197***	0.261***
	(0.106)	(0.0735)	(0.0832)
>65	0.217*	0.130	0.167*
	(0.112)	(0.0796)	(0.0891)
Education level (ref. 'primary education)			
lower secondary VMBO	0.212	0.00836	0.0425
	(0.130)	(0.116)	(0.120)
upper secondary:MBO	0.152	0.0769	0.0624
	(0.135)	(0.119)	(0.125)
upper secondary: HAVO/VWO	0.403***	0.192	0.332***
	(0.135)	(0.120)	(0.124)
Tertiary: HBO	0.352***	0.218*	0.294**
	(0.131)	(0.118)	(0.122)
Tertiary: University	0.582***	0.354***	0.499***
	(0.136)	(0.120)	(0.125)
Income quartiles (reirst quartile)			
1902<x<=2600	0.252***	0.0658	0.168***
	(0.0701)	(0.0502)	(0.0591)
2600<x<=3471	0.331***	0.111**	0.241***
	(0.0756)	(0.0524)	(0.0632)
x>3471	0.414***	0.155***	0.327***
	(0.0779)	(0.0566)	(0.0661)
refuse/dk	0.113	0.157	0.216
	(0.228)	(0.127)	(0.169)
Constant	1.542***	2.318***	1.946***
	(0.172)	(0.133)	(0.145)
Observations	1,528	1,528	1,528
R-squared	0.191	0.121	0.174

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1