Gender gaps in financial literacy: a multi-arm RCT to break the response bias in surveys*

Laura Hospido
Banco de España, CEMFI and IZA

Nagore Iriberri
University of the Basque Country, UPV/EHU, IKERBASQUE, Basque Foundation for Science, and Lund University

Margarita Machelett
Banco de España

November 3, 2026

Abstract

Women are consistently more likely to respond "I do not know" (IDK) to financial questions, a pattern that may reflect not only differences in knowledge but also response behavior. We study how survey design shapes IDK use and measured financial literacy using a randomized experiment with 6,000 adults in Spain. We compare a standard survey format with three alternatives: removing the IDK option, offering monetary incentives for correct answers, and providing an information nudge that combines information about gender differences in IDK use with encouragement to provide an educated guess. Removing IDK mechanically eliminates such responses, while both incentives and the information nudge substantially reduce them. The information nudge also reduces the gender gap in IDK use by about half and makes the residual gap statistically indistinguishable from zero in specifications with controls. Effects on the measured gender gap in financial literacy are smaller and imprecisely estimated: incentives, if anything, widen the gap, whereas the information nudge provides suggestive evidence of narrowing it.

JEL Classification: C8, C9, D14, D91, G53, I22, J16

Keywords: Financial literacy, gender gaps, survey methods

*We are grateful to Manuel Arellano, Dmitry Arkhangelsky, Andrea Baños, Laia Bosque Mercader, Irma Clots-Figueras, Guillermo Cruces, Sebastian Fanelli, Libertad Gonzalez, Carlos Sanz, Julián Scannone, Angelika Welte, and to the seminar participants at the Banco de España, CEMFI, the Cherry Blossom Financial Education Institute, the COSME Gender Economics workshop, the Diversity and Human Capital Workshop, and the G53 Network Brown Bag Lunch for helpful comments. Nagore Iriberri acknowledges funding by grant PID2023-147817NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe" and Eusko Jaurlaritza (IT2061-26). The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Banco de España or the Eurosystem.

1 Introduction

Financial literacy, the understanding of basic financial concepts such as inflation and risk diversification, impacts important economic decisions, including retirement and saving plans, stock market participation, and, ultimately, households' wealth levels and well-being (IOSCO, 2018; Lusardi and Mitchell, 2014, 2023). Improving the general population's financial literacy, especially for the most vulnerable, has become a major policy goal in many countries (OECD, 2013). For example, the G20 has declared financial literacy an essential skill, supporting the design and evaluation of financial literacy policies (G20, 2021). An appropriate policy response should consider that financial knowledge is not evenly distributed in the population. In particular, women consistently perform worse than men on financial literacy surveys, a fact that is persistent across countries and over time (OECD, 2016; Klapper and Lusardi, 2020). Part of this gap can be explained by differences in observable characteristics such as education, experience in financial decisions, and interest in financial topics.[1] Yet, a considerable part of this gap remains unexplained.

In this paper, we argue that part of the observed gender gap can reflect response behavior in surveys rather than actual differences in knowledge. Standard financial literacy measures are based on the percent of correct answers in surveys that typically allow for "I do not know" (IDK) as a valid option. This creates a challenge: we cannot distinguish whether IDK reflects lack of knowledge or other motivations such as lack of confidence, risk, regret, action aversion, or other behavioral traits. In addition, women are systematically more likely than men to choose IDK (Chen and Garand, 2018; Cziriak et al., 2024; Kim and Mountain, 2019; Hospido et al., 2023; Wilmarth et al., 2023; Conte et al., 2024; Bertola and Lo Prete, 2025; Bucher-Koenen et al., 2024; Tranfaglia et al., 2024). As a result, measured gender gaps may partly capture response bias, a form of measurement error in which the observed distribution of correct answers does not map one-to-one to true knowledge. The implications of this are far from innocuous. To illustrate, consider a simple extreme case: suppose men and women were equally and perfectly knowledgeable, but women always answered IDK while men always provided an answer. Then, conventional measures would show men to be perfectly literate and women entirely illiterate. Such distortions risk misleading policy interventions. More broadly, any relationship between financial literacy and financial behavior should be interpreted with explicit attention to the role of IDK responses.

[1] See, among others, (Chen and Volpe, 2002; Fonseca et al., 2012; Driva et al., 2016; Hsu, 2016; Bucher-Koenen et al., 2017; Zaccaria and Guiso, 2020; Bottazzi and Lusardi, 2021).

We causally test whether the propensity to choose IDK can be shifted and study how these changes affect measured financial literacy by gender. Our interventions alter the conditions under which respondents report their answers rather than their underlying stock of financial knowledge, allowing us to study how response behavior shapes measured financial literacy. Using a sample of 6,000 adults in Spain, we field an online survey based on the OECD International Network of Financial Education (INFE) questionnaire and the Spanish Survey of Financial Competences (OECD, 2022; Hospido et al., 2023). We focus on the key concepts measured by the "Big Five" financial literacy questions (Hastings et al., 2013), originally developed by Annamaria Lusardi and Olivia Mitchell, to measure the basic understanding of inflation, compound interest rate, risk diversification, mortgages, and bond pricing.[2] They are either multiple-choice or true-or-false questions that allow for IDK as a valid answer. We isolate the causal impact of interventions on IDK responses and measured gender gaps by assigning participants to the control or one of the three treatment arms: without IDK, incentives, or information nudge.

Participants in the control group complete the financial literacy section with the IDK option available, as in standard surveys. The first treatment, "without IDK", presents the same financial literacy section but without the IDK option, forcing respondents to provide an answer. The second treatment, "incentives" keeps the IDK option and provides explicit and immediate monetary rewards for each correct answer.[3] Finally, the information nudge also keeps the IDK option while informing all participants about the gender gap in financial literacy and women's propensity to choose IDK relative to men's (the information part) and a sentence prompting participants to provide educated guesses when uncertain (the nudge part).[4] Because the two components are introduced jointly, the experiment identifies only their combined effect and cannot separately identify the effect of the information and the encouragement components.

A conceptual framework helps interpret the effects of our interventions and clarifies the relationship between measured and underlying financial knowledge, which we develop in detail in Section 2. The framework highlights four key determinants of survey responses: underlying financial knowledge, effort, confidence, and the threshold governing respondents' willingness to provide an answer rather than select IDK.

[2]We measure the same five underlying concepts relying on the OECD/INFE Toolkit for Measuring Financial Literacy and Financial Inclusion as our primary reference. As such, the precise wording and format of our questions do not exactly follow the Lusardi–Mitchell formulation.

[3]See similar treatments related to the introduction of incentives, although in different contexts, for example in Fehr and Falk (1999), Nagin et al. (2002), Segal (2012), Krawczyk (2012) and Sittenthaler and Mohnen (2020).

[4]The use of information has been widely used in different settings, as shown by Gabaix (2019), both in non-financial settings (Boring and Philippe, 2021) and in financial settings, for example, to increase loan repayment or savings (Karlan et al., 2015; Haran Rosen and Sade, 2022).

Our interventions affect different elements of this decision process. The without IDK treatment removes the option of selecting IDK and therefore eliminates selection into answering. It thus provides a particularly informative benchmark for response accuracy when everyone is required to answer, although it does not by itself identify underlying knowledge. Under the additional assumption that forced responding does not differentially affect response accuracy by gender, the gender gap observed in this arm is the most informative among our interventions about the underlying gender gap in financial knowledge. By contrast, monetary incentives and the information nudge induce respondents to answer while preserving the IDK option, operating through effort, confidence, and/or response thresholds. Consequently, although both interventions are expected to reduce IDK use, their effects on measured financial literacy and the observed gender gap depend on both how much answering increases among women and men and how accurate the resulting response shifts are, such that the observed gender gap can narrow or widen with respect to the control group.

Overall, we confirm the two key patterns observed in the literature in our control group. First, women score lower on financial literacy and are more likely to select IDK. The difference amounts to roughly 9 percentage points in correct answers and 6 percentage points in IDK responses, while the gap in incorrect answers is small. Controlling for observable characteristics reduces but does not close the gaps in correct and IDK responses. By contrast, the gap in incorrect answers is no longer significantly different from zero. The gender gap in IDK responses accounts for two-thirds of the gender gap in financial literacy.

Interestingly, and the main innovation of this paper, we find that the propensity to provide an answer is highly malleable. First, under all three interventions -removing the IDK option, providing monetary incentives, and giving an information nudge - IDK responses fall sharply. Second, we observe different impacts by treatment and gender. By design, the without IDK reduces the gap in IDK use to zero. Introducing incentives does not significantly reduce the gender gap in IDK answers. By contrast, the information nudge treatment is effective in reducing the gap in response bias, reducing it to half, from over 6 to close to 3 percentage points, and becoming statistically indistinguishable from zero in specifications with controls.

In turn, we test how reductions in IDK affect the gender gaps in financial literacy and incorrect responses. Although IDK use is highly malleable, the implied effects on the gender gap in financial literacy are modest and less precisely estimated.

Eliminating the IDK answer goes in the direction of reducing the gender gap in financial literacy, but the effect is not significant. Women correctly answer 57 percent of the questions (relative to 49 in the control group), while men answer correctly up to 63 percent (relative to 58). In addition, it increases the percent of

incorrect answers significantly more for women. Under the framework assumptions discussed in Section 2, the resulting 6 percentage point gap provides the closest benchmark among our interventions for the underlying gender gap in financial knowledge.

The incentives treatment does not reduce the gender gap in financial literacy; if anything, the point estimates imply a larger gap, but the difference is not statistically significant. Percent correct increases significantly only for men, and percent incorrect significantly increases only for women. Therefore, these response patterns are consistent with a more favorable shift from IDK toward correct rather than incorrect answers for men than for women. By contrast, under the information nudge, the increase in correct answers is larger for women than for men, by about 3 percentage points, a differential effect that is marginally significant at the 10 percent level in specifications with controls. Under this treatment, the gender gap in measured financial literacy decreases, and the remaining difference is not statistically distinguishable from zero in specifications with controls (but with a p-value equal to 0.11). This treatment also shows the smallest widening of the gender gap in incorrect responses. These results show that information nudges help equalize behavior in surveys in terms of the use of IDK across men and women, and they provide suggestive evidence that they narrow gender differences in measured financial literacy. For all interventions, however, the estimated gender-differential treatment effects on correct answers are imprecise and smaller than the minimum detectable effects of our design, so they do not allow us to draw firm conclusions.

A complementary benchmark for interpreting these effects compares the observed increase in correct and incorrect answers with those expected under random guessing. This helps distinguish answering more from answering more accurately. While results regarding correct answers are imprecise, the differential increase in incorrect answers is significantly smaller than what would be expected under random guessing. This suggests that the observed response shifts are not purely consistent with random guessing. This benchmark analysis also reinforces our gender specific results: under monetary incentives, men's response shifts are more favorable relative to the random guessing benchmark than women's. By contrast, the information nudge produces response patterns for women that are more consistent with informed rather than purely random responding, although the evidence is weaker and only marginally statistically significant for incorrect answers.

In summary, we show that IDK use is highly malleable, whereas the effects on gender gaps in measured financial literacy are smaller and less precisely estimated.

Beyond the experimental evidence, we complement our analysis with descriptive correlations between IDK answers and a broad set of financial outcomes that are directly relevant for individual well-being. Using

the 2021 Spanish Financial Competences Survey, we relate the percent of IDK responses in the core financial knowledge questions to standard indices of financial behavior, objective and subjective financial well-being, and financial product awareness and holding as proposed by the 2023 OECD/INFE International Survey on Adult Financial Literacy (OECD, 2023). These indices summarize whether individuals follow recommended financial practices (such as actively saving, keeping track of expenses and paying bills on time), their ability to cope with shocks, their satisfaction with their financial situation, and the range of payment, savings, insurance and credit products they know and use. We find that, even after conditioning on the level of financial knowledge and gender, a higher IDK percent is strongly negatively associated with most of these indicators and with many of their underlying components. This suggests that IDK answers capture systematic differences in financial behavior and market participation, not just measurement noise, and underscores the policy relevance of understanding and potentially shifting response behavior in financial literacy surveys.

This paper contributes to three different strands of the literature. First, we contribute to the survey methodology literature on IDK and no opinion responses and their implications for measurement error (Durand and Lambert, 1988; Schuman and Presser, 1996; Tourangeau et al., 2000). Recent work by Bertola and Lo Prete (2025) brings these measurement concerns directly to financial literacy surveys, showing how heterogeneous response strategies can affect observed financial literacy scores.

Second, we contribute to the literature on gender differences in financial literacy and in the choice of IDK responses. A large literature has documented persistent gender gaps in financial literacy and has focused primarily on observable characteristics that explain part of this gap (see reviews by Lusardi and Mitchell, 2014; Bucher-Koenen et al., 2017). More recently, a smaller but growing body of work has examined how survey design and response behavior shape measured financial literacy and gender differences. In particular, several papers explicitly model the choice of IDK or analyze it in correlational regressions (Chen and Garand, 2018; Kim and Mountain, 2019; Hospido et al., 2023; Wilmarth et al., 2023; Conte et al., 2024). These studies highlight the role of confidence and response behavior in explaining the higher propensity of women to select the IDK option. We add to this literature by providing causal evidence on interventions that impact IDK responses. Later studies by Harrison et al. (2024) and Lee et al. (2025) focus on the role of confidence in the choice of IDK and financial literacy. Harrison et al. (2024) propose an alternative measurement approach in which respondents allocate probability mass across answer options to capture beliefs about financial knowledge. The study by Lee et al. (2025) provides complementary evidence that an intervention with personalized feedback on confidence with information on its relevance for financial abilities reduced

IDK choices and improved decision-making in an incentivized investment task.

More closely related papers are Tranfaglia et al. (2024), Bucher-Koenen et al. (2024). Tranfaglia et al. (2024) and Bucher-Koenen et al. (2024) find that removing the IDK option shrinks gender gaps in financial literacy. Bucher-Koenen et al. (2024) use a within-subject panel design, in which their second-round questions additionally elicit confidence after each item. Tranfaglia et al. (2024) compare financial literacy in the 2021 Survey of Household Economics and Decision Making between respondents who face a survey with an IDK option and those who do not, and find that the gender gap in financial literacy shrinks when IDK is removed. In our experiment, the effect of eliminating the IDK option also goes in this direction but is not statistically significant. Importantly, the without IDK treatment mechanically eliminates the IDK option, so it cannot show whether participants would voluntarily change their IDK response behavior - the channel our remaining treatments can address.

Interestingly, gender gaps in measured political knowledge have also been attributed to gender differences in the choice of IDK; see, for example, (Mondak and Anderson, 2004) and Lizotte and Sidman (2009). As far as we know, no RCT was implemented to test for the effect of survey takers' propensity to choose IDK answer option in the political domain. It remains to be studied whether similar interventions presented in our study are equally effective in reducing the choice of the IDK answer option in surveys measuring political knowledge.

Finally, our work is also related to gender differences in omitting questions in multiple-choice tests used in educational tests. In particular, to a large literature that has studied how gaps vary in multiple-choice tests with and without differential grading for omitting questions and providing wrong answers. Women are found to omit more questions than men (Ben-Shakhar and Sinai, 1991; Baldiga, 2014; Pekkarinen, 2015; Coffman and Klinowski, 2020; Espinosa and Gardeazabal, 2020; Iriberri and Rey-Biel, 2021), which can be related to choosing IDK answers. These studies often find that deleting differential grading of incorrect and omitted answers reduces gender gaps in performance measures. While these studies examine high-stakes testing environments, we analyze anonymous surveys, showing that simple design interventions can reduce gender differences in response behavior even in low-stakes contexts.

The rest of the paper is organized as follows. In Section 2 we describe the survey design, the treatments and a simple conceptual framework. Section 3 presents the results. Section 4 explores the association between IDK and financial outcomes. Finally, Section 5 concludes.

2 Survey Design, Treatments, and Conceptual Framework

We designed an online survey experiment to test the effectiveness of three interventions to reduce the frequency of choosing IDK answers and their effect on financial literacy. The survey design was approved by the Ethics Committee at the Institutional Review Board and the survey design and pre-analysis plan were pre-registered.[5]

The survey consisted of about 40 questions, lasted 15 minutes on average, and was administered by the survey company 40dB in Spain between October 24 and November 18 of 2022. Survey takers received a small fixed payment of 1.20€ for completing the questionnaire. This amount was determined based on the expected time required to complete the survey and is standard compensation in the company.

Figure 1 summarizes the experimental design, which we describe in detail below, and Appendix B reproduces the complete questionnaire. As the figure shows, the survey was divided into three main parts, starting with questions about individual demographic information, family and household background, and the elicitation of behavioral traits and perceptions such as interest and perceived knowledge of financial topics and risk preferences. The inclusion of these questions is motivated by the research showing that these are important variables in explaining gender gaps in financial literacy.[6] We also included questions on family background, such as parental education, to have a richer set of control variables. This section was the same for all survey takers and had no time limit. These questions were included before participants received each treatment and entered the financial literacy section for two reasons. First, this sequence aligns with INFE toolkit guideline and thus follows the standard structure used in international surveys practice measuring financial literacy. Second, this order ensures that participants' responses are not affected by the treatments, preventing potential biases such as shifts in confidence or risk aversion.

Then, the survey included a financial literacy section based on the INFE/OECD questionnaire. The introductory section text is standard and encourages participants to try to avoid choosing the IDK answer option. This text was excluded from the treatment arm without IDK. This section included 10 questions with a total time limit of 7 minutes to complete.[7] These included the "Big Five" financial literacy questions

[5]The University of the Basque Country UPV/EHU Ethics Committee certified the exemption (https://www.ehu.eus/en/web/ceid/) because the researchers received anonymized data from the survey company, 40dB, which has configured and implemented a personal data protection program in accordance with the reforms of the European Data Protection Regulation 2016/679 and the Organic Law on Protection of Personal Data 3/2018. The experiment and the pre-analysis plan are pre-registered in the AEA RCT Registry under the RCT ID AEARCTR-0009896.

[6]See, for instance, (Bucher-Koenen et al., 2017) for an overview of the literature exploring gender gaps in financial literacy.

[7]The purpose of the time limit was to minimize the probability of searching for answers online. Respondents were informed of

that assess the understanding of inflation, compound interest rate, risk diversification, mortgages, and bond pricing concepts. These will be the focus for the main results.[8] The goal of providing a time limit was to minimize the probability of section interruptions and searching for answers. We also kept track of the time spent answering the financial literacy section.

Finally, after the financial literacy section, the survey ended with 7 additional questions about participants' experiences, perceptions, and holdings of financial products. Unlike previous studies, we gathered information on individual's experiences which might help explain decision-making and gender gaps. Based on participants' reported past experiences, we construct a *lean-in index* combining five binary indicators capturing whether the respondent has ever: (i) applied for a promotion at work; (ii) requested a salary increase; (iii) served as a class representative at school or university; (iv) competed in an individual sport; or (v) competed in a team sport.[9] The perceptions refer to behavioral traits such as having the confidence to deal with unexpected events. The financial product questions referred to ownership and means of interacting with banks (e.g., online). Then, the questionnaire ended with a question on perceived survey difficulty. This last part was the same for all survey takers and had no time limit.[10]

We randomly assigned each treatment by gender immediately before entering the financial literacy section. Importantly, we varied the financial literacy section design while keeping the same questions.

The survey design included a control version and three treatments to evaluate as interventions to potentially reduce IDK. Following standard survey practices, the control group included the financial literacy questions with the option IDK among the answer options. Specifically, we follow INFE/OECD Toolkit for Measuring Financial Literacy suggestion (OECD, 2022) and use the verbatim text from the Spanish Financial Competences Survey (Hospido et al., 2023). Related to IDK answers, it states: "If you do not know the answer, just say so. If you think you have the right answer, it is likely that you do." The first treatment arm, without IDK, omitted this option from the possible answers and the standard introductory section text omitted the mention of IDK answers. The second treatment arm, incentives, kept the IDK option in the answers and offered additional

the limit at the start of the section, but the survey platform did not enforce it. As a result, 286 participants (4.77% of the sample) spent more than 7 minutes on this section. Section 3.4 shows that the share exceeding the limit does not differ significantly by treatment or by gender, and that excluding these participants leaves our results unchanged.

[8] In addition to the "Big Five", we included a simple division question, an interest rate question without compounding, a probability question, a question included in the cognitive reflection test, and a simple subtraction exercise that we used to identify those survey takers who did not pay attention.

[9] More in detail, each indicator takes the value one if the respondent reports having undertaken the corresponding activity at least once, and zero otherwise. We standardize each indicator using its sample mean and standard deviation and construct the index as the equally weighted average of the five standardized components. Higher values indicate greater engagement in activities related to self-advocacy, leadership, and competition.

[10] All the questions included in the survey can be found in the Online Appendix B.

monetary incentives for correct answers. Specifically, the incentive consisted of paying 7 additional cents for each correct answer, up to 70 cents if all 10 questions were answered correctly. Although the incentive is low in absolute terms, which may limit its behavioral impact (Gneezy and Rustichini, 2000), it is sizable relative to the fixed participation payment: respondents could earn up to an additional 70 cents, approximately 60 percent of the 1.20€ fixed payment. Finally, the third treatment, referred to as the information nudge, also kept IDK among the answers and included an introductory text before starting the financial literacy section. This information aimed to increase awareness of gender differences in propensities to choose IDK and repeated the instruction to try to avoid choosing IDK answers to nudge them away from choosing this answer. Specifically, all survey takers assigned to this treatment, both men and women, received the following information nudge:

Men typically answer 7 out of 10 financial questions correctly. Women 6 out of 10. This difference is mostly explained (65 percent) because women choose the answer "I do not know" more often than men. Therefore, we ask you to - please - avoid answering "I do not know".

These statistics came from the microdata of the Spanish Survey of Financial Competences implemented by the Bank of Spain and the Comisión Nacional del Mercado de Valores in 2016. Any intervention using informational nudges might be susceptible to potential experimenter demand effects. However, it is reassuring that De Quidt et al. (2018) show that experimenter demand effects tend to be small in online experiments.

The survey was pre-tested in a pilot implemented in September 2022. The sample size was determined by the project's budget constraint rather than by a formal ex ante power calculation. Given the available funding, we maximized the number of completed surveys that could be fielded. This yielded a target sample of 6,000 respondents (3,000 men and 3,000 women), with 2,400 assigned to the control group (1,200 men and 1,200 women) and 1,200 to each of the three treatment arms (600 men and 600 women). To obtain 6,000 completed surveys, the survey company collected 6,879 surveys, of which 879 survey takers abandoned the survey before completing it. In Subsection 3.5, we test whether there are gender differences in the probability of exiting the survey, and, in particular, we analyze whether different treatment arms impact the probability of exiting the survey differently.

To facilitate the interpretation of the three interventions and their effects, it is useful to consider a simple conceptual framework that distinguishes four key elements underlying an observed survey response. First, respondents possess an underlying stock of financial knowledge, K, which is unobserved and cannot be affected by our interventions.[11] Second, respondents exert effort, e, which is not perfectly observed and may

[11]The treatments cannot change respondents' stock of financial knowledge because they are assigned immediately before the

be influenced by the interventions. Both K and e determine the probability that the individual's preferred answer to a question is correct, denoted by $p(K, e)$. When individuals have no knowledge, pure random guessing provides a natural benchmark for this probability. Third, conditional on their knowledge and the effort devoted to answering a question, respondents form a subjective assessment, q, of the probability that their preferred answer is correct, which we interpret as confidence and we assume that q is increasing in p. Fourth, respondents have a threshold, τ, for providing an answer rather than selecting IDK. This threshold may reflect risk preferences, regret, reluctance to guess, and other factors affecting response behavior. All four elements, K, e, q, and τ, may vary across individuals and, importantly, across gender. When the IDK option is available, respondent i provides an answer to question j, A_{ij}, if the perceived probability that the preferred answer is correct exceeds the threshold for providing an answer, that is, $A_{ij} = \mathbf{1}\{q_{ij} \geq \tau_i\}$. Let S_{ij} indicate whether the preferred answer is correct, with its probability determined by knowledge and effort. The observed correct-response indicator can then be expressed as $C_{ij} = A_{ij}S_{ij}$. Thus, measured financial literacy reflects two margins: the extensive margin of providing an answer, which is governed by confidence q and the response threshold τ, with confidence itself potentially reflecting knowledge K and effort e; and the accuracy of the answers provided. At the group level, we define response accuracy as the share of correct answers among the answers provided.

The three treatments affect different elements of the response decision. The without IDK treatment removes the option of selecting IDK and therefore eliminates selection into answering. Although it does not change respondents' underlying stock of knowledge, respondents who would otherwise select IDK must now report their preferred answer, which may reflect an educated or random guess, or other response strategies. The share of correct answers in this treatment arm would equal the share of questions respondents would answer correctly if all of them reported their preferred answer. This arm therefore provides a particularly informative benchmark for what happens when selection into IDK is eliminated. However, the observed correct-response rate in this arm reflects response accuracy under forced responding and does not by itself provide an unbiased measure of underlying financial knowledge.

With respect to the gender gap, under the additional assumption that any changes in effort, guessing, or other response strategies induced by forced responding do not differentially affect the accuracy of men's and women's responses, the without IDK arm provides the closest benchmark among our interventions for the underlying gender gap in financial knowledge. The other two treatments leave the IDK option available and

financial literacy section and do not provide any explanation or new financial information that could increase financial knowledge.

therefore allow us to assess how much respondents are induced to provide an answer and how accurate the resulting response shifts are, including whether these effects differ by gender. Monetary incentives introduce an additional reward for a correct response, which may lower the response threshold (τ) and increase the effort (e) devoted to answering the question. Respondents who seek to maximize their expected monetary payoff should therefore be more likely to provide an answer.

The information nudge may operate through confidence (q), the response threshold (τ), and potentially effort (e). The information about gender differences may change respondents' beliefs about their own performance (q), while the explicit encouragement to avoid IDK may change their willingness to provide an answer (τ) or the effort devoted to doing so (e). Because the information and encouragement components are randomized jointly, we cannot separately identify these channels. Because both treatments are designed to induce respondents to answer, IDK use is expected to decline relative to the control but, unlike in the Without IDK treatment, need not disappear altogether. With respect to the gender gap in measured financial literacy, the effects of incentives and the information nudge depend on both how much answering increases among women and men and how accurate the resulting response shifts are. As a result, these treatments may either widen or narrow the observed gender gap relative to the control group, moving the measured gap closer to or further from the underlying gender gap in financial knowledge.

This framework illustrates what can and cannot be learned from the effects of the interventions on observed responses. Changes in IDK responses identify the effect of each treatment on the propensity to provide an answer, while changes in correct and incorrect answers show how the observed distribution of responses changes. Comparing these changes with those expected under pure random guessing indicates whether the observed response shifts are consistent with random guessing. Interpreting these shifts as reflecting the responses of individuals induced to answer additionally requires that treatments do not induce respondents who would otherwise answer to select IDK and do not alter the accuracy of respondents who would have answered in the control condition.

3 Results

3.1 Descriptive Statistics and Randomization

Table 1 presents summary statistics for the control and the three treated groups. It includes all characteristics later used as control variables and outcome variables in the main analysis and shows the p-value for the F-test

on differences across groups.

First, regarding respondents' main demographics, they are equally represented by gender by design, are, on average, 46 years old, about 92% were born in Spain, and more than 50% live in big cities. More than half of them have completed at most secondary education, and about 70% are currently working. When comparing these main sociodemographic characteristics with the Spanish population aged 25–70 (Appendix Table A1), we find that our experimental sample is somewhat younger (46.1 versus 47.9 years) and more educated (44.5 versus 38.4 percent with post-secondary education), while employment rates are similar (68.5 versus 67.2 percent). Although these differences should be considered when extrapolating our findings to the broader population, the causal identification of treatment effects relies on random assignment within the experimental sample. Reassuringly, the control group also reproduces the two main gender gaps documented in the literature: women answer fewer financial literacy questions correctly and select IDK more frequently than men, as discussed in Section 3.2.

Second, regarding respondents' household characteristics, over 70% live with a partner, have an average of about 3 household members, and are most likely the primary income earners. Survey takers' parents have low education, as more than half of mothers and fathers have completed at most primary education. In line with this, close to 75% recall having fewer than a bookshelf of books when they were 10 years old.

Third, regarding their self-assessments on financial knowledge and risk-taking behavior, we find that most participants perceive they have either neutral (over 40%) or good (close to 40%) knowledge of financial topics. In line with this, they believe they would get close to 6 correct answers in a 10-question financial questionnaire, which is slightly above their observed percent correct answers and their interest in finance is about 6 on a scale between 0 and 10. On average, respondents in the control group expect to answer 5.6 of the ten financial literacy questions correctly, and they actually answer 5.4 correctly (nine knowledge questions, rescaled to 0–10).[12] Both risk aversion measures, one qualitative and one lottery choice using the elicitation method by Eckel and Grossman (2002), show that the survey takers are, on average, slightly risk averse.

Finally, they place themselves close to 4 in their self-efficacy, confidence, and lean-in measures, which come from statements with an agreement scale between 1 and 5. Regarding their management of finances, they have, on average, 2.5 out of 6 saving or investing products, such as saving accounts or pension plans, and

[12]By gender, and in the control group, men expect 5.9 correct answers and women 0.72 fewer. Part of this difference reflects differences in actual performance: men answer 5.9 questions correctly and women 4.9. However, comparing men and women with the same number of correct answers, women still expect 0.51 fewer correct answers than men (p-value < 0.001). This pattern is consistent with lower confidence q among women, even conditional on realized performance.

about 1.4 out of 3 debt products, among loans, mortgages, and credit cards. Most have performed online bank operations over the last 12 months, while about 3% had no bank operations.

The last column reports the *p*-values for the F-test when comparing all control variables across the four treatment groups. Overall, the treatment groups are broadly balanced across observable characteristics, although there are some statistically significant differences in education, self-assessed financial knowledge, and expected correct answers. Finite-sample differences can arise by chance.Reassuringly, our main results are robust to conditioning on these characteristics.

These descriptive statistics are also presented, separately by gender, in Tables A2 and A3 in the Appendix. Consistent with the literature, men and women differ in some characteristics, mainly in behavioral traits and perceptions. They also differ in characteristics related to managing finances. For example, women tend to be more risk-averse (Croson and Gneezy, 2009), less interested in finance (Brown and Graf, 2013), and less confident in their self-perceived financial knowledge than men (Bordalo et al., 2019). They also hold fewer financial products, are less likely to do online bank operations, and are more likely to have no bank operations (Almenberg and Dreber, 2015; Bottazzi and Lusardi, 2021; Balutel et al., 2023). Most importantly, the *p*-values in the final columns overall show that, when split by gender, both men and women have comparable socio-demographics, family background, behavioral traits, and experience managing finances across the four treatment groups. The exceptions are education level (for women), and by the assessment of financial knowledge and the expected number of correct answers (for men). As mentioned above, finite-sample differences can arise by chance, but reassuringly, the results are not sensitive to including these characteristics as controls.

3.2 Main Results: Evaluation of the Three Interventions

We first describe the raw patterns in IDK, correct and incorrect answers, by treatment and by gender (Table 1 and Figures 2 and 3). We then move to the regression framework (Tables 2, 3, and 4). Finally, we carry out a complementary analysis by comparing the observed changes in correct and incorrect answers with those expected under random guessing (Table 5).

The *main outcomes* panel in Table 1 show the mean values for the percent of IDK, correct and incorrect answers for the "Big Five" financial literacy questions, for each of the treatment groups. In the control group, the survey takers choose the IDK answer on average in about 15 percent of the questions, while the three treatment arms reduce this percentage significantly. Mechanically, the biggest decrease is when this

option is eliminated, followed by the information treatment, chosen in 7 percent of the questions, and finally, introducing monetary incentives, chosen in 9 percent of the questions. When participants respond to more questions, the results in financial literacy and percent incorrect are also shifted. On average, in the control group, survey takers answer 53 percent of the questions correctly, and the three treatment arms are effective in increasing this percentage. When deleting the IDK option, survey takers answer correctly 60 percent of the questions, followed by providing information nudges and monetary incentives, with about 57 and 56 percent correct each, respectively. With respect to percent incorrect, they increase from 30 percent in the control to 38, 33 and 34 percent in the without IDK, incentives and information nudge, respectively.[13] As expected by these differences, the *p*-values in the final column show that the different interventions impacted the three outcome variables significantly.

With respect to gender differences, for the control group, we confirm the two main findings on gender gaps in financial literacy, which we show in Figure 2. First, we observe a gender gap of about 9 percentage points in financial literacy (men answer about 58 percent of the questions correctly, while women answer correctly 49 percent of them), in line with the existing literature and comparable to the findings by Hospido et al. (2023), that report a 10 percentage point gender gap among the Spanish adult population, and slightly smaller than the one found by Bucher-Koenen et al. (2024), that report a 14 percentage point gender gap in the Dutch-speaking population in Netherlands. Second, of this roughly 9 percentage point gap, about 6 percentage points correspond to the gender gap in IDK answers. Consistent with Bucher-Koenen et al. (2024) and Hospido et al. (2023), about two-thirds of the gender gap in financial literacy corresponds to different propensities to choose the IDK answer. Specifically, men choose the IDK in close to 12 percent of the questions. In comparison, women choose it at a higher rate (18 percent). Last, when looking at the percent incorrect answers, an outcome not usually studied in the literature, we find that men's and women's performance differs by less than 2 percentage points (significant at the 10% level).

Finally, Figures 3a, 3b and 3c show the raw gender gaps by treatment for percent IDK, percent correct and percent incorrect. Without IDK treatment mechanically eliminates the gender gap in IDK, while decreasing the gap in percent correct and increasing the gap in percent incorrect. The incentives treatment increases the gender gap in both percent correct and incorrect. The information nudge reduces the gap in IDK, reducing also the gap in percent correct and increasing the gap in percent incorrect.

[13] Skipped answers are not considered correct, incorrect or "I do not know" answers. Skipped answers account for 1.2 percent in the control group, 1.7, 1.7, and 2 percent in the without IDK, incentives, and information nudge treatment. Table A11 shows the results do not vary if these are considered as IDK answers.

We next present the regression framework used to evaluate the effects of the three interventions. Formally, for each outcome, - the share of IDK, correct and incorrect answers- we estimate:

$$\begin{aligned} Y_i = {} & \alpha + \beta\,\text{Woman}_i + \theta_1\,\text{WithoutIDK}_i + \theta_2\,\text{Incentives}_i + \theta_3\,\text{Information}_i \\ & + \lambda_1\big(\text{Woman}_i \times \text{WithoutIDK}_i\big) + \lambda_2\big(\text{Woman}_i \times \text{Incentives}_i\big) + \lambda_3\big(\text{Woman}_i \times \text{Information}_i\big) \\ & + X_i'\gamma + \varepsilon_i, \end{aligned} \tag{1}$$

where for each survey respondent i Woman_i is an indicator for female respondents, WithoutIDK_i, Incentives_i, and Information_i are indicators for the three treatment arms, and $\text{Woman}_i \times \text{WithoutIDK}_i$, $\text{Woman}_i \times \text{Incentives}_i$, and $\text{Woman}_i \times \text{Information}_i$ are the respective interactions between being a female respondent and each of the treatments. Finally, X_i is the vector of control variables reported in Table 1, with the corresponding reference categories omitted for categorical variables and ε_i is the error term. All OLS regressions are estimated using robust standard errors.

With this framework, we evaluate the impact of the three interventions on the gender gaps in both, the frequency of IDK, financial literacy (percent correct) and percent incorrect measures. Table 2 shows the results for the choice of IDK, Table 3 shows the results for financial literacy, and finally, Table 4 shows the results for percent incorrect. Tables 2, 3 and 4 show the estimation results without any controls, shown in column (1), with all controls, shown in column (2), and with a lasso-selected set of controls that include the main control variables from column (2) and interactions between these, shown in column (3). Appendix Tables A4, A5 and A6 report all the estimated coefficients of the corresponding Tables 2, 3 and 4, respectively.

First, it is worth noting that the gender gap in IDK answers and financial literacy remain large and significant in the specifications that add control variables. For instance, adding confidence and risk aversion among the controls helps reduce the gender gap in financial literacy from 9 percentage points to 6, and in IDK from 6.5 to 4 percentage points (as shown in Tables 3 and 2). In contrast, the conditional gender gap in incorrect answers is further reduced and no longer significantly different from zero, as shown in Table 4.

Second, regarding the impact on IDK answers, with the extreme intervention of deleting the option of IDK, this is mechanically reduced to 0, and consequently, the gender gap is closed (See Table 2). Offering incentives also reduces significantly this answer option. However, the gender gap is not significantly reduced with this intervention. Finally, the information nudge is also effective in reducing the percent of IDK. In this case, however, the effect is significantly larger for women (reduced to 8.6 percent) than for men (reduced to

5.7 percent), at the 1% significance level, such that the gender gap is reduced. In the specification without controls, in column (1), the response bias is reduced by half, from about 6 to 3 percentage points, and in the specification with controls, in columns (2) and (3), the gender gap in the propensity to choose IDK is closed. As shown in Table 2, we reject the hypothesis that all treatment effects are equal for each gender.

How do these results in the IDK answers translate into the financial literacy and percent incorrect measures? As shown in Tables 3 and 4, respectively, we reject the null hypothesis that all treatment effects in percent correct and incorrect are equal for both men and women.

For women, the 18 percentage point mechanical reduction in the without IDK treatment translates into gains of 8 points in correct and 10 points in incorrect answers. For men, the corresponding 12 percentage point reduction is split evenly into each.[14] Relative to Bucher-Koenen et al. (2024) and Tranfaglia et al. (2024), we also find that eliminating IDK raises correct responses more for women than for men, but, in our case, the implied narrowing of the gender gap in correct answers is not statistically significant. By contrast, the increase in incorrect answers is significantly larger for women.

In contrast to the without IDK treatment—which eliminates IDK by design—the incentives and information nudge treatments operate through adjustments in participants' answer choices. Under incentives, correct answers increase significantly for men, while the corresponding increase for women is smaller. The gender differential is marginally significant in the unadjusted specification but not in specifications with controls. At the same time, incorrect answers increase significantly for women but not for men, widening the gender gap in incorrect answers. Taken together, these response patterns are consistent with a more favorable shift away from IDK for men than for women. This differential behavior is reminiscent of gender differences in reaction to monetary incentives, for example found by Krawczyk (2012) and Sittenthaler and Mohnen (2020). If anything, the point estimates imply a *wider* gender gap in measured financial literacy, although this result is not robust across specifications.

Finally, the information nudge treatment increases the percent of correct answers for both men and women. Contrary to the incentives treatment, the raw increase is larger for women (5 percentage points) than for men (2 percentage points), implying a reduction in the gender gap in correct answers of about 3 percentage points. This differential reduction is marginally statistically significant at the 10 percent level in specifications with controls. As a result, the gender gap in correct answers is reduced by half, from 6 to 3 percentage points in

[14] These effects are obtained from column 1 of Tables 2, 3, and 4, for IDK, correct, and incorrect answers, respectively. For men, they correspond to the Without IDK coefficient; for women, to the sum of the Without IDK and Women × Without IDK coefficients.

the specification with the full set of controls (see Table 3, column 2). The residual gap of 3 percentage points is not statistically distinguishable from zero (p-value = 0.11), suggesting that the information nudge treatment narrows the gender gap in measured financial literacy.[15] In line with this, the treatment does not significantly change the gender gap in incorrect answers (see Table 4). The shift away from IDK under the information nudge is therefore associated with more favorable response patterns for women than for men. If anything, this pattern *narrows* the measured gender gap.[16]

Our estimation results reveal a clear contrast: IDK use is highly sensitive to survey design, with the information nudge significantly narrowing the gender gap in IDK responses, whereas effects on the gender gap in correct answers are smaller and less precisely estimated. In order to characterize the statistical precision of our estimates, we report minimum detectable effects (MDEs) for the gender-by-treatment interactions. We focus on column (1) of Tables 2 and 3, which report the experimental specification without conditioning on control variables, and use a two-sided 5 percent significance level and 80 percent power. The resulting MDEs are reported in Appendix Table A7, with all effects expressed in percentage points. For the change in the gender gap in correct answers, analytical MDEs range from 4.77 to 4.98 percentage points across the three treatments. For the change in the gender gap in IDK responses, they range from 3.75 to 3.90 percentage points across the two treatments in which IDK remains available. Simulation-based MDEs are very similar.

The estimated gender differential effects on correct answers are below their corresponding MDEs for all three treatments. This implies that, if the true effects were of the estimated magnitudes, the realized design would have less than 80 percent power to detect them; it does not imply that such effects cannot be detected. This supports interpreting the evidence on treatment-induced changes in the measured financial literacy gender gap as suggestive and imprecisely estimated.

A complementary benchmark for interpreting these effects is to compare the observed increase in correct and incorrect answers, by treatment and gender, with the increase that would be expected under random guessing. This concern is particularly salient in the without IDK treatment, where participants are forced to provide an answer by design. For example, if an additional response is generated on an item with four answer options, under simple random guessing it would be correct with probability 25 percent and incorrect with probability 75 percent. The differences between the observed increase in percent correct and incorrect and the expected increase under random guessing are provided in Table 5, top panel for correct answers and

[15] Pairwise comparisons by gender show that the treatment effects in financial literacy are significantly different, except when comparing men in the without IDK and incentives group, where the p-value for the last specification is 0.45.

[16] Section 3.3 examines whether these effects vary with individuals' characteristics.

bottom panel for incorrect answers, by treatment and both for all participants and separately by gender.[17]

In the top panel, the first column replicates the overall reduction in IDK answers by each of the interventions, shown in Table 2. The second column, similarly, replicates the effects in percent correct, shown in Table 3. The third column shows the expected change in percent correct within a simple random guessing benchmark. Finally, the fourth column shows the difference between the observed and random-guessing changes, and the last column reports the one-sided p-value for whether the observed response shift is more accurate than implied by random guessing. The bottom panel shows the equivalent results for incorrect answers. In the full sample, no intervention significantly raises correct answers above the random-guessing benchmark. The observed increase in incorrect answers is below the benchmark in all three treatment arms, significantly so for the without IDK treatment and marginally for the information nudge, while the difference for incentives is not statistically significant. The difference in statistical significance across the two outcomes partly reflects differences in statistical precision.[18]

There are some gender-specific differential patterns, in line with Tables 3 and 4. Under incentives, men's response shifts are more favorable relative to the random-guessing benchmark: their increase in correct answers is significantly above the benchmark and their increase in incorrect answers significantly below it, while neither difference is statistically significant for women. Under the information nudge, the pattern is reversed but less precisely estimated. For women, the increase in incorrect answers is below the random-guessing benchmark at the 10 percent level (p-value = 0.09), while the increase in correct answers is above the benchmark but not statistically significant (p-value = 0.15). For men, neither difference is statistically significant.

The random-guessing benchmark is informative about the accuracy of the observed response shifts relative to chance. Response shifts that are more accurate than random guessing indicate that the additional answering is not purely mechanical, but contains information beyond chance. The benchmark does not, however, distinguish whether this reflects changes in confidence, response thresholds, or effort.

[17]To calculate the expected increase in percent correct by random guessing, we first calculate the expected increase question by question, weight it by the increase in provided answers, and then we take the simple average. In particular, as the inflation question has 3 alternative answers, the interest rate question has 4, the risk diversification and mortgages are true/false questions and the bond pricing one has 3 alternative answers, the expected increase for each type of question would differ.

[18]The comparison with the random guessing benchmark is statistically more precise for incorrect answers than for correct answers. Under random guessing, the probability of an incorrect answer is substantially larger than that of a correct one (e.g., 75% vs. 25% with four options), so the observed increase in incorrect answers and the expected increase under random guessing are larger than from our benchmark in correct answers. The difference between these is therefore large enough to be statistically distinguishable. For correct answers, both the observed increase and the random guessing benchmark are small, so their difference is a small number, making it harder to reject the null that the observed response shift is consistent with random guessing.

3.3 Heterogeneity Analysis

Motivated by the possibility that the information nudge may operate differently across respondents with different characteristics and baseline propensities to select IDK, Appendix Table A8 examines treatment effects separately by education, age, and subjective financial knowledge.

The most salient subgroup difference observed is along the subjective financial knowledge dimension.[19] The information nudge produces substantially larger reductions in IDK among respondents with low self-assessed financial knowledge, who also have considerably higher baseline IDK propensities. A similar pattern appears for correct answers. However, for women the reduction in IDK relative to the corresponding control group mean is remarkably stable across subgroups, at approximately 53–54 percent, suggesting that the larger absolute effects largely reflect greater scope for IDK responses to fall.

By contrast, we do not find strong evidence that the gender differential response to the information nudge varies systematically across these dimensions. Tests of equality of the Woman × Information Nudge coefficients across education, age, and subjective knowledge subgroups do not reject equality for either IDK or correct answers. Thus, while the nudge has larger absolute effects among respondents with low subjective financial knowledge, its effect on the gender difference in response behavior does not appear to be concentrated in a particular observable subgroup. These exploratory comparisons suggest that the gender-differential effect of the information nudge does not vary systematically across these observable subgroups.

3.4 Robustness Checks

We perform several robustness checks. First, because our analysis tests multiple hypotheses, we also report false discovery rate (FDR)-adjusted q-values following Anderson (2008). Appendix Table A9 shows that the conclusions remain similar to our main analysis. Using the three main outcomes, we find that the significance remains at similar levels for all the estimates using the FDR-adjusted $q-$values. The exception corresponds to the impact in correct answers for men receiving the information nudge treatment, where the $q-$value increases to 0.115. The gender gap, however, is still significantly reduced at the 10% with this treatment.

Second, the same pattern remains when analyzing the "Big Five" questions individually, as shown in Tables A10 in the Appendix. Although the behavioral patterns go in the same direction and are less precise for

[19] Joint tests reject equality of treatment effects across subjective-financial-knowledge groups for both IDK and correct answers, but do not reject equality across education or age groups.

each of the "Big Five" questions, the strongest effects in closing the gender gap in the choice of IDK answers with the information nudge treatment seem to come from the questions on inflation, risk diversification and bond pricing. In addition, we assess the robustness of the IDK results using alternative sets of financial-literacy questions included in the financial literacy section (see Table A11 in the Appendix). Specifically, we include the "Big Five" but define omitted questions as IDK answers (column 2)[20], the Big Three (column 3), which excludes the questions on mortgages and bond pricing from the "Big Five" questions. We add the simple interest rate question to the "Big Five" questions ('Big Six, column 4), and finally, we include all the questions in the financial literacy section (All, column 5). The IDK results are robust across these alternative question sets.

Third, we examine whether the inclusion of participants who exceeded the seven-minute limit on the financial literacy section affects our results. As noted in Section 2, this limit was announced but not enforced by the survey platform and 286 participants (4.77% of the sample) exceeded it. Columns (1) and (2) of the Appendix Table A14 take an indicator for exceeding this limit as the dependent variable. The share does not differ significantly across treatments, nor across gender-treatment indicators. The largest interaction corresponds to women in the without IDK group, which is consistent with the longer completion time this group takes - as documented in the next Section - though the difference is not statistically significant at conventional levels. Columns (3)-(6) show that our conclusions do not vary when we re-estimate our main IDK and correct-answer specifications when restricting the sample to the 5,714 individuals that completed the section within seven minutes.

To sum up, all three interventions reduce IDK use, but their gender effects differ. By design, removing IDK eliminates the gender gap in IDK use; incentives reduce IDK without significantly narrowing this gap; and the information nudge reduces IDK more strongly for women, significantly narrowing the gender gap in IDK responses. Effects on measured financial literacy are smaller and less precisely estimated. Under incentives, the observed response shifts are more favorable for men than for women, whereas under the information nudge they are more favorable for women than for men. Accordingly, the point estimates widen the measured gender gap under incentives and narrow it under the information nudge, although these gender-differential effects are imprecisely estimated.

[20]There are 140 participants that skip at least one of the "Big Five" questions. They are currently considered non-IDK, non-correct, and non-incorrect answers. In this robustness exercise, they are considered IDK answers

3.5 Further Results: Attrition, Perceived Difficulty, and Completion Time

We also study the impact of the treatments on measures of attrition, perceived difficulty, and completion time of the "Big Five" questions obtained from the survey (Table A15 in the Appendix). Overall, we find no evidence that the treatments affect attrition or perceived difficulty, either overall or differentially by gender. The main gender-specific treatment effect among these auxiliary outcomes concerns completion time.[21]

The probability of abandoning the survey is a policy-relevant outcome, particularly for online surveys. The survey company 40dB collected 6,879 surveys to obtain 6,000 completed surveys. Therefore, 879 survey takers left the survey before completing it, a metric that is expected by 40dB. Of those, 115 exited before answering the first survey question, which records gender. Of the remaining 6,764 participants, 544 exited before treatment randomization, which took place immediately before the financial literacy section, and 220 exited after it. Thus, 659 of the 879 exits (75%) occurred before treatment assignment, most of them during the initial demographic questions (Appendix Table A12 reports exits by survey screen and gender). Among the 6,764 participants whose gender is observed, women are 4.5 percentage points more likely than men to exit the survey (13.5 versus 9.0 percent; column (1) of Table A15), and 4.0 of these 4.5 percentage points arise before randomization. Appendix Table A13 compares participants who exit with those who complete the survey, separately for women and men, using for each characteristic all participants who answered the corresponding question. In both groups, those who exit are somewhat less educated and less interested in finance than those who complete, and these differences are similar for women and men. The only exception is age ($p = 0.02$): men who exit are about two years younger than men who complete, while women who exit are of similar age to women who complete. Financial literacy and IDK answers are elicited only after randomization, so we do not observe them for participants who exit earlier, and we therefore cannot directly assess whether pre-randomization attrition is associated with our outcomes of interest. Among the 6,220 participants who pass the first part of the questionnaire and are assigned to treatments, the gender difference in the probability of exiting is no longer statistically significant. More importantly, we find no evidence that the treatments affect the probability of abandoning the survey, either overall or differently for women and men (see Column (3) of Table A15).

With respect to perceived difficulty, measured on a scale between 0 and 10, survey takers, on average, give a score slightly over 4, as shown at the bottom part of Table 1. Women, on average, find it more difficult than

[21] There is a significant increase in completion time for women in the without IDK group. This pattern is consistent with increased effort under forced responding among women, although completion time is only a noisy proxy for effort.

men. However, we find no evidence that the treatments differentially affect participants' perceived difficulty (columns (5) and (6) in Table A15 in the Appendix).

Finally, regarding completion time for the "Big Five" questions, the bottom panel of Table 1 shows that participants take about a minute and a half to complete the five questions. Interpreting completion time as a noisy proxy for effort, as in Augenblick et al. (2015) and Gneezy et al. (2019), we find little evidence of systematic effort responses under the incentives or information nudge treatments, or of differential responses by gender. The exception is the without IDK treatment, which significantly increases completion time for women relative to men (columns (7) and (8) of Appendix Table A15).

4 Further Discussion: Association between IDK Answers and Financial Outcomes

Our experimental results show that the propensity to choose the IDK option is highly malleable and differs systematically by gender. But does this matter economically? If IDK use were unrelated to financial behavior conditional on measured financial knowledge, then measurement concerns would be limited to survey measurement. We show that this is not the case. We complement the experimental evidence with descriptive correlations between IDK answers and a broad set of financial outcomes related to individual well-being, using the 2023 OECD/INFE International Survey of Adult Financial Literacy (OECD, 2023). The OECD survey combines a battery of financial knowledge questions with rich information on financial behaviors, financial resilience, and ownership and awareness of financial products. We construct the financial indicators[22] and combine them with the microdata from the 2021 Financial Competences Survey, which is representative of the Spanish population, so that our measures are directly comparable to the standard OECD financial literacy and financial well-being indices.

Formally, for each financial outcome Y_i we estimate

$$Y_i = \alpha + \beta_W \, \text{Woman}_i + \beta_K \, \text{Share_Correct}_i + \beta_{IDK} \, \text{Share_IDK}_i + \varepsilon_i, \tag{2}$$

where Woman_i is an indicator for female respondents, Share_Correct_i is the share of correct answers to the seven OECD financial knowledge questions, Share_IDK_i is the share of those questions for which the respondent chose the IDK option, and ε_i is the error term. We estimate Equation (2) by OLS using

[22] The definitions are drawn from Appendix C of the 2023 OECD/INFE International Survey of Adult Financial Literacy (OECD, 2023) and reflect the available data from the Spanish Competences Survey.

heteroskedasticity-robust standard errors.[23] Our coefficient of interest is β_{IDK}, which captures whether a stronger tendency to answer IDK is associated with differences in financial outcomes, conditional on measured financial knowledge and gender.

Table 6 reports the results for five composite indicators: (i) financial behavior index, (ii) objective financial well-being, (iii) subjective financial well-being, (iv) financial product awareness, and (v) financial product holding. The financial behavior index aggregates eight "good practice" behaviors that are keeping track of money, actively saving, not borrowing to make ends meet, seeking advice, monitoring personal finances, setting long-term goals, making considered purchases, and paying bills on time. Objective and subjective financial well-being capture, respectively, the ability to cope with income shocks and self-reported satisfaction with one's financial situation. Finally, the product indices summarize how many broad categories of financial products respondents know and actually hold. These financial product categories involve payment products; savings, investment or retirement products; credit products; insurance products, and crypto-assets.

Two patterns emerge. First, the share of correct answers is strongly and positively associated with all five indicators, as expected: individuals who answer more questions correctly also display more prudent financial behaviors, greater resilience, and broader product awareness and use. Second, and more importantly for our purposes, the share of IDK answers is negatively associated with most of these outcomes, even after conditioning on financial knowledge. A higher propensity to answer IDK is strongly correlated with lower scores in the financial behavior index and with lower objective financial well-being, and it is also associated with knowing and holding fewer categories of financial products. The only time the coefficient on IDK is not statistically significant is for subjective financial well-being, suggesting that self-reported satisfaction is less tightly linked to IDK behavior than objective indicators.

To better understand which dimensions drive these aggregate associations, Online Appendix Tables A16 and A17 decompose the indices into their underlying components. Table A16 reports regressions of each individual behavior (e.g. keeping track of expenses, actively saving, paying bills on time) on the same regressors as in equation (2). Table A17 repeats the exercise at the category level for product awareness (Panel A) and product holding (Panel B). The categories involve knowing or holding at least one product in payment, savings/investment/retirement, insurance, credit, and crypto-asset product categories. This granular evidence allows us to see on which specific behaviors and margins of financial market participation the IDK

[23] Following OECD (2023), the financial knowledge module includes seven questions on basic concepts such as simple and compound interest, inflation, and risk diversification. Share_Correct$_i$ and Share_IDK$_i$ are defined over these seven items. The outcome indices are rescaled so that higher values correspond to better financial behavior or well-being.

propensity is most strongly reflected, and confirms that survey response behavior is strongly associated with economically meaningful outcomes.

Overall, these descriptive correlations indicate that IDK responses carry information about respondents' financial behavior, resilience, and participation in financial markets over and above what is captured by standard measures of financial knowledge. In combination with our experimental results, these associations suggest that policies and survey designs that shift individuals away from IDK—and in particular interventions that mitigate gender differences in the propensity to choose IDK—may affect not only measured financial literacy, but also the identification of groups at risk of poor financial outcomes.

5 Concluding Remarks

This paper shows that, consistent with other studies, there is a gender gap in financial literacy and tests whether it partly reflects differences in response behavior in choosing IDK. As our main contribution, we carried out the first RCT to evaluate how different interventions affect men's and women's propensity to choose IDK answers and, consequently, measured gender gaps in financial literacy.

Our first main finding is that IDK use is highly malleable. All three interventions substantially reduce IDK answers, confirming that the choice of this answer option does not solely reflect lack of knowledge. Importantly, the interventions differ in their effectiveness in reducing the gender gap in IDK use. Removing the IDK option mechanically eliminates this gap. Incentives reduce overall IDK shares but do not significantly narrow the gender difference. By contrast, the information nudge is the only treatment that significantly reduces this gender gap, while preserving the IDK option, reducing it by half in raw specifications and closing it entirely in specifications with controls.

Our second finding is that gender gaps in measured financial literacy and in incorrect answers are more difficult to shift. While the three interventions increase the percentage of correct answers, the effects on gender gaps are smaller and imprecisely estimated. Removing the IDK option tends to reduce the gender gap in correct answers, but at the same time significantly increases the gender gap in incorrect answers. With respect to the incentives treatment, if anything, the point estimates suggest a wider gender gap in measured financial literacy, as the shift away from IDK is associated with more favorable response patterns for men than for women. Finally, the information nudge increases measured financial literacy more for women than for men, with an effect that is only marginally significant at the 10 percent level in specifications with controls.

Taken together, we conclude that, although IDK use is highly malleable, both overall and differentially by

gender, the resulting effects of these interventions on measured financial literacy, and in particular on gender gaps in financial literacy, are smaller in magnitude and less precisely estimated.

These findings have several implications. First, our experiment provides causal evidence on how survey design affects response behavior. However, to assess the broader applicability of these findings, it is important to acknowledge that our sample differs somewhat from the Spanish population in terms of age and educational attainment, and that the experiment was conducted online, with relatively low monetary stakes and an announced seven-minute limit for the financial-literacy section. An interesting avenue for future research would be to examine whether similar interventions produce comparable effects in higher-stakes settings, official surveys with representative samples, different modes of administration, or in other countries.

Second, although the three interventions are highly effective in reducing IDK responses, reducing IDK use does not by itself narrow the measured gender gap in financial literacy. The effect depends on how much men and women shift away from IDK and on how accurate these response shifts are. It is relevant to consider that the interventions present different practical challenges.The without IDK treatment eliminates selection into responding and, under additional assumptions, provides the closest benchmark among our interventions for estimating the underlying gender gap in financial knowledge. However, forced responding does not necessarily provide an unbiased measure of knowledge, it is unlikely to be feasible in voluntary surveys conducted by institutions and may also be undesirable when maintaining comparability with existing financial literacy measures is important. Monetary incentives are costly at scale without significantly reducing gender differences in the propensity to choose IDK at the low stakes we implement. By contrast, the information nudge offers a promising alternative: it is a simple, low-cost intervention that preserves the IDK option while reducing both its overall use and gender differences in the propensity to select it.

More broadly, our findings underscore the importance of reporting correct, incorrect, and IDK responses jointly, rather than focusing exclusively on the percentage of correct answers, particularly when assessing gender gaps in financial literacy.

References

Almenberg, J. and A. Dreber (2015). Gender, stock market participation and financial literacy. *Economics Letters 137*, 140–142.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A

reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association 103*(484), 1481–1495.

Augenblick, N., M. Niederle, and C. Sprenger (2015). Working over time: Dynamic inconsistency in real effort tasks. *The Quarterly Journal of Economics 130*(3), 1067–1115.

Baldiga, K. (2014). Gender differences in willingness to guess. *Management Science 60*(2), 434–448.

Balutel, D., W. Engert, C. Henry, K. P. Huynh, D. Rusu, and M. C. Voia (2023). Crypto and financial literacy of cryptoasset owners versus non-owners: The role of gender differences. *Journal of Financial Literacy and Wellbeing 1*(3), 514–540.

Ben-Shakhar, G. and Y. Sinai (1991). Gender differences in multiple-choice tests: The role of differential guessing tendencies. *Journal of Educational Measurement 28*(1), 23–35.

Bertola, G. and A. Lo Prete (2025). Who prefers guessing to admitting they don't know? measurement error in financial literacy surveys. *Journal of Economic Behavior & Organization 233*, 107003.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2019). Beliefs about gender. *American Economic Review 109*(3), 739–773.

Boring, A. and A. Philippe (2021). Reducing discrimination in the field: Evidence from an awareness raising intervention targeting gender biases in student evaluations of teaching. *Journal of Public Economics 193*, 104323.

Bottazzi, L. and A. Lusardi (2021). Stereotypes in financial literacy: Evidence from pisa. *Journal of Corporate Finance 71*, 101831.

Brown, M. and R. Graf (2013). Financial literacy and retirement planning in switzerland. *Numeracy 6*(2), 6.

Bucher-Koenen, T., R. Alessie, A. Lusardi, and M. van Rooij (2024). Fearless woman: Financial literacy, confidence, and stock market participation. *Management Science*.

Bucher-Koenen, T., A. Lusardi, R. Alessie, and M. V. Rooij (2017). How financially literate are women? an overview and new insights. *Journal of Consumer Affairs 51*(2), 255–283.

Chen, H. and R. P. Volpe (2002). Gender differences in personal financial literacy among college students. *Financial Services Review 11*(3), 289–307.

Chen, Z. and J. C. Garand (2018). On the gender gap in financial knowledge: Decomposing the effects of don't know and incorrect responses. *Social Science Quarterly 99*(5), 1551–1571.

Coffman, K. B. and D. Klinowski (2020). The impact of penalties for wrong answers on the gender gap in test scores. *Proceedings of the National Academy of Sciences 117*(16), 8794–8803.

Conte, A., P. Paiardini, and J. Temperini (2024). Dealing with "do not know" responses in the assessment of financial literacy: The use of a sample selection model. *International Journal of Financial Studies 12*(3), 76.

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

Cziriak, M., T. Bucher-Koenen, and R. Alessie (2024). Beyond knowledge: Confidence and the gender gap in financial literacy. Technical report, ZEW Discussion Papers.

De Quidt, J., J. Haushofer, and C. Roth (2018). Measuring and bounding experimenter demand. *American Economic Review 108*(11), 3266–3302.

Driva, A., M. Lührmann, and J. Winter (2016). Gender differences and stereotypes in financial literacy: Off to an early start. *Economics Letters 146*, 143–146.

Durand, R. M. and Z. V. Lambert (1988). Don't know responses in surveys: analyses and interpretational consequences. *Journal of Business Research 16*(2), 169–188.

Eckel, C. C. and P. J. Grossman (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and human behavior 23*(4), 281–295.

Espinosa, M. and J. Gardeazabal (2020). The gender-bias effect of test scoring and framing: A concern for personnel selection and college admission. *The B.E. Journal of Economic Analysis & Policy 20*(3), 20190316.

Fehr, E. and A. Falk (1999). Wage rigidity in a competitive incomplete contract market. *Journal of political Economy 107*(1), 106–134.

Fonseca, R., K. J. Mullen, G. Zamarro, and J. Zissimopoulos (2012). What explains the gender gap in financial literacy? the role of household decision making. *Journal of Consumer Affairs 46*(1), 90–106.

G20 (2021). G20 finance ministers and central bank governors' meeting 9 -10 july 2021 – 2 communiquÉ.

Gabaix, X. (2019). Behavioral inattention. In *Handbook of behavioral economics: Applications and foundations 1*, Volume 2, pp. 261–343. Elsevier.

Gneezy, U., J. A. List, J. A. Livingston, X. Qin, S. Sadoff, and Y. Xu (2019). Measuring success in education: The role of effort on the test itself. *American Economic Review: Insights 1*(3), 291–308.

Gneezy, U. and A. Rustichini (2000). Pay enough or don't pay at all. *The Quarterly journal of economics 115*(3), 791–810.

Haran Rosen, M. and O. Sade (2022). The disparate effect of nudges on minority groups. *The Review of Corporate Finance Studies 11*(3), 605–643.

Harrison, G., D. Ross, and J. T. Swarthout (2024). Gender, confidence, and the mismeasure of intelligence, competitiveness and literacy.

Hastings, J. S., B. C. Madrian, and W. L. Skimmyhorn (2013). Financial literacy, financial education, and economic outcomes. *Annu. Rev. Econ. 5*(1), 347–373.

Hospido, L., S. Izquierdo, and M. Machelett (2023). The gender gap in financial competences. *Banco de España Analytical Article March*.

Hospido, L., M. Machelett, M. Pidkuyko, and E. Villanueva (2023). Survey of financial competences 2021. *Survey of Financial Competences/Banco de España*.

Hsu, J. W. (2016). Aging and strategic learning: The impact of spousal incentives on financial literacy. *Journal of Human Resources 51*(4), 1036–1067.

IOSCO, O. (2018). The application of behavioural insights to financial literacy and investor education programmes and initiatives.

Iriberri, N. and P. Rey-Biel (2021). Brave boys and play-it-safe girls: Gender differences in willingness to guess in a large scale natural field experiment. *European Economic Review 131*, 103603.

Karlan, D., M. Morten, and o. Zinman (2015). A personal touch in text messaging can improve microloan repayment. *Behavioral Science & Policy 1*(2), 25–31.

Kim, N. and T. P. Mountain (2019). Financial knowledge and "don't know" response. *Journal of Consumer Affairs 53*(4), 1948–1969.

Klapper, L. and A. Lusardi (2020). Financial literacy and financial resilience: Evidence from around the world. *Financial Management 49*(3), 589–614.

Krawczyk, M. (2012). Incentives and timing in relative performance judgments: A field experiment. *Journal of Economic Psychology 33*(6), 1240–1246.

Lee, I., A. Lleó-Bono, E. Tipoe, and C. Rauh (2025). The causal effects of confidence awareness on financial literacy and behaviour. Technical report, Barcelona School of Economics.

Lizotte, M.-K. and A. H. Sidman (2009). Explaining the gender gap in political knowledge. *Politics & Gender 5*(2), 127–151.

Lusardi, A. and O. S. Mitchell (2014). The economic importance of financial literacy: Theory and evidence. *American Economic Journal: Journal of Economic Literature 52*(1), 5–44.

Lusardi, A. and O. S. Mitchell (2023). The importance of financial literacy: Opening a new field. Technical report, National Bureau of Economic Research.

Mondak, J. J. and M. R. Anderson (2004). The knowledge gap: A reexamination of gender-based differences in political knowledge. *The Journal of Politics 66*(2), 492–512.

Nagin, D. S., J. B. Rebitzer, S. Sanders, and L. J. Taylor (2002). Monitoring, motivation, and management: The determinants of opportunistic behavior in a field experiment. *American Economic Review 92*(4), 850–873.

OECD (2013). Evaluating financial education programs: Survey, evidence, policy, instruments and guidance.

OECD (2016). Oecd/infe international survey of adult financial literacy competencies. *OECD*.

OECD (2022). Oecd/infe toolkit for measuring financial literacy and financial inclusion 2022.

OECD (2023). Oecd/infe 2023 international survey of adult financial literacy. *OECD Business and Finance Policy Papers 39*.

Pekkarinen, T. (2015). Gender differences in behaviour under competitive pressure: Evidence on omission patterns in university entrance examinations. *Journal of Economic Behavior & Organization 115*, 94–110. Behavioral Economics of Education.

Schuman, H. and S. Presser (1996). *Questions and answers in attitude surveys: Experiments on question form, wording, and context*. Sage.

Segal, C. (2012). Working when no one is watching: Motivation, test scores, and economic success. *Management Science 58*(8), 1438–1457.

Sittenthaler, H. M. and A. Mohnen (2020). Cash, non-cash, or mix? gender matters! the impact of monetary, non-monetary, and mixed incentives on performance: Hm sittenthaler, a. mohnen. *Journal of Business Economics 90*(8), 1253–1284.

Tourangeau, R., L. J. Rips, and K. Rasinski (2000). *The psychology of survey response*, Volume 595. Cambridge New York.

Tranfaglia, A., A. Lloro, and E. Merry (2024). Question design and the gender gap in financial literacy.

Wilmarth, M. J., K. T. Kim, and T.-Y. Pak (2023). What do we really know about "don't know"? re-assessing the measurement of financial knowledge. *Journal of Consumer Affairs 57*(4), 1623–1649.

Zaccaria, L. and L. Guiso (2020). Patriarchy to partnership: Gender equality and household finance. *mimeo*.

Figures and Tables

Figure 1: Diagram of the Experimental Design

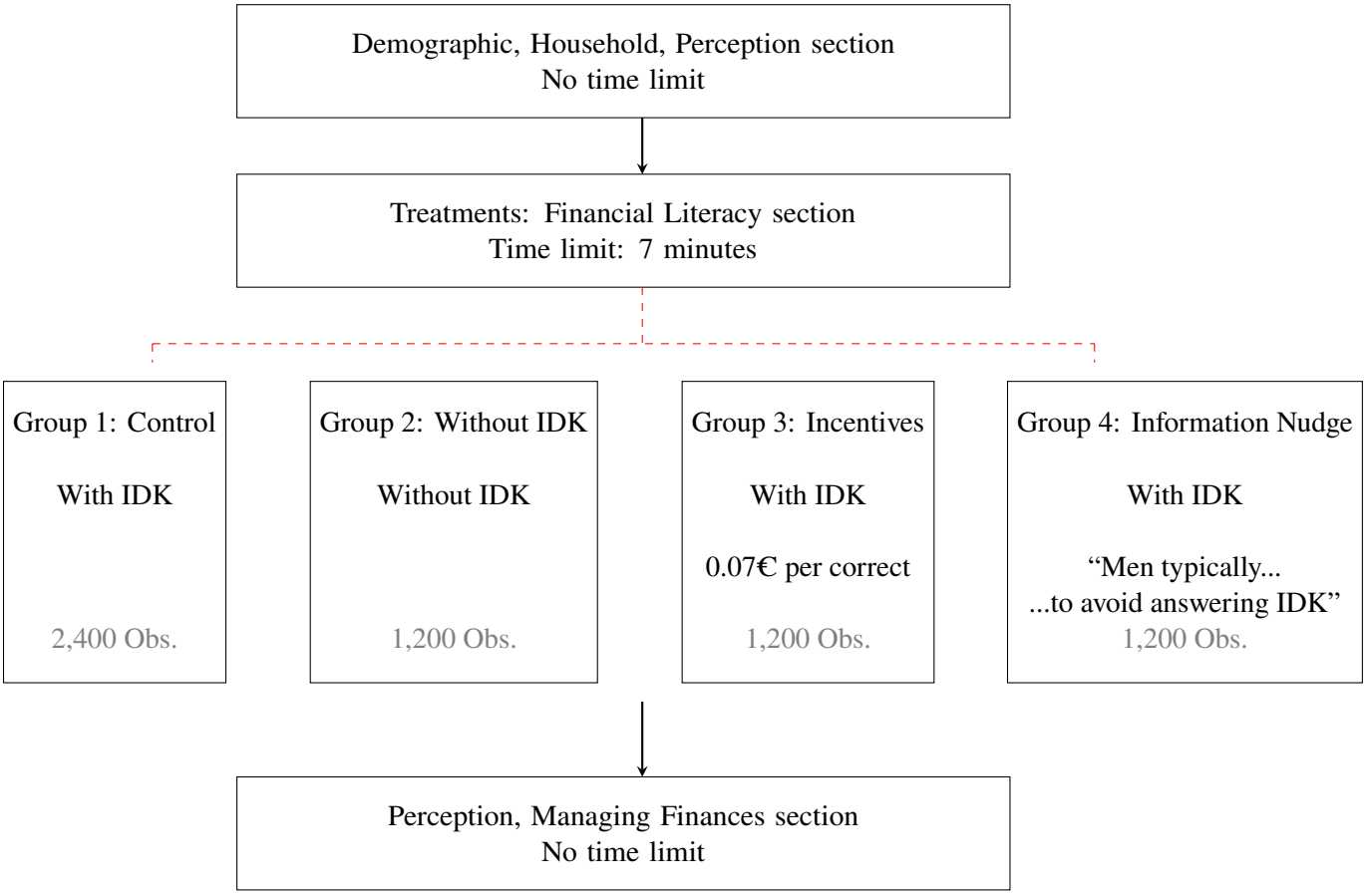
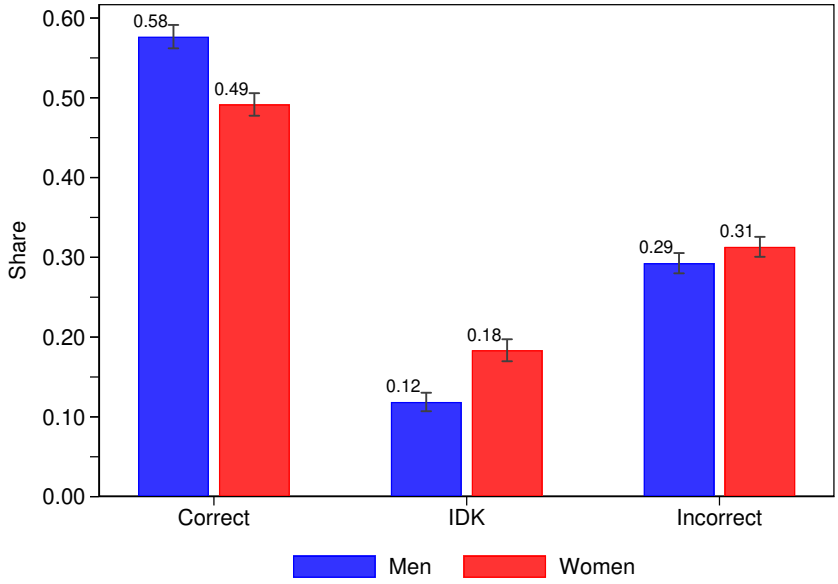
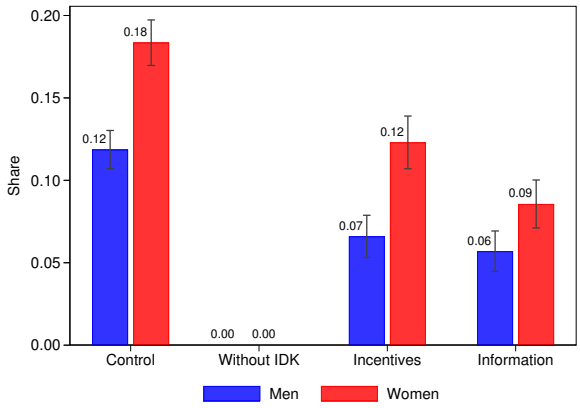


Figure 2: Main Outcomes in the "Big Five" Questions in the Control Group by Gender

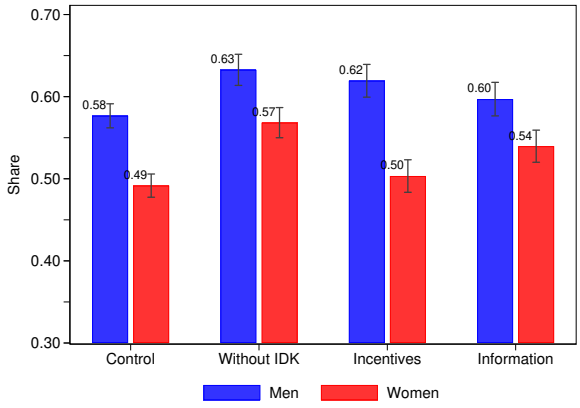


Notes: Raw average shares of correct answers, IDK answer option, and incorrect answers, along with their 95% confidence intervals, for the "Big Five" questions in the control group, by gender. Average shares do not add up to 1 because 1.2% of men and 1.2% of women skipped at least one question. There are 140 participants that skipped at least one of the "Big Five" questions. There are currently considered non-IDK, non-correct, and non-incorrect answers as in the Spanish Financial Competences Survey.

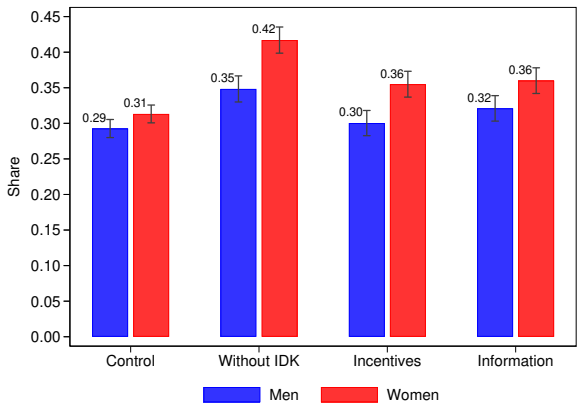
Figure 3: Shares of IDK, Correct and Incorrect Answers by Treatment Group and Gender



(a) "I do not know" answers



(b) Correct answers



(c) Incorrect answers

Notes: Panel (a) plots raw average shares of IDK answer option, panels (b) and (c) plot raw average shares of correct and incorrect answers, along with their 95% confidence intervals, for the "Big Five" questions by treatment group and gender. Average shares do not add up to 1 because 1.2% of men and 1.2% of women skipped at least one question. There are 140 participants that skipped at least one of the "Big Five" questions. They are currently considered non-IDK, non-correct, and non-incorrect answers as in the Spanish Financial Competences Survey.

Table 1: Descriptive Statistics (mean values) and Randomization

		(1) Control	(2) Without IDK	(3) Incentives	(4) Information Nudge	(5) p-value
Demographics	Woman	0.50	0.50	0.50	0.50	1.00
	Age 25-34	0.19	0.19	0.17	0.18	0.70
	Age 35-44	0.26	0.27	0.26	0.27	0.87
	Age 45-54	0.30	0.30	0.30	0.30	0.97
	Age 55-70	0.25	0.24	0.27	0.26	0.47
	Spaniard	0.91	0.93	0.93	0.91	0.16
	Pop. size 0-20th	0.21	0.18	0.19	0.20	0.16
	Pop. size 20th-100th	0.27	0.26	0.28	0.27	0.70
	Pop. size 100th+	0.52	0.56	0.53	0.53	0.16
	Primary education	0.20	0.17	0.25	0.21	0.00
	Secondary education	0.34	0.37	0.35	0.35	0.36
	University education	0.35	0.35	0.30	0.34	0.02
	Master, PhD education	0.11	0.11	0.10	0.10	0.76
	Working	0.70	0.69	0.67	0.67	0.12
	Retired	0.10	0.10	0.12	0.11	0.37
	Unemployed	0.19	0.19	0.20	0.21	0.31
Household	< 1 bookshelf at age 10	0.75	0.74	0.73	0.73	0.74
	> 2 bookshelves at age 10	0.25	0.26	0.27	0.27	0.74
	Household size	2.99	2.97	3.04	2.97	0.43
	Primary earner	0.67	0.66	0.69	0.66	0.44
	Lives with partner	0.73	0.71	0.72	0.73	0.59
	Mother: Primary education	0.60	0.57	0.59	0.59	0.60
	Mother: Secondary education	0.19	0.21	0.19	0.20	0.43
	Mother: Post-secondary education	0.20	0.19	0.19	0.19	0.80
	Father: Primary education	0.53	0.53	0.55	0.56	0.38
	Father: Secondary education	0.20	0.20	0.20	0.19	0.68
	Father: Post-secondary education	0.23	0.23	0.20	0.22	0.42
	Partner: Primary education	0.17	0.15	0.18	0.18	0.12
	Partner: Secondary education	0.24	0.25	0.24	0.23	0.78
	Partner: Post-secondary education	0.33	0.32	0.30	0.32	0.50
Assessments	Very low financial knowledge	0.02	0.02	0.02	0.03	0.70
	Low financial knowledge	0.12	0.12	0.12	0.12	0.94
	Neutral financial knowledge	0.42	0.41	0.40	0.44	0.20
	Good financial knowledge	0.38	0.41	0.40	0.37	0.09
	Very good financial knowledge	0.06	0.04	0.06	0.05	0.03
	Expected correct answers (0-10)	5.58	5.73	5.79	5.52	0.00
	Interest in finance (0-10)	6.10	6.13	6.14	5.97	0.35
	Risk willingness (0-10)	4.65	4.77	4.74	4.62	0.48
	Lottery choice (1-8)	3.62	3.58	3.67	3.65	0.77
Perceptions	Lean-in index (mean of five z-scores)	0.01	0.01	-0.03	0.00	0.39
	Perceived self-efficacy (1-5)	3.96	4.00	4.00	4.02	0.20
	Perceived confidence (1-5)	3.80	3.87	3.83	3.83	0.18
	Perceived lean-in (1-5)	3.65	3.67	3.64	3.63	0.84
Managing finances	Saving products (0-6)	2.53	2.60	2.46	2.56	0.07
	Debt products (0-3)	1.39	1.39	1.36	1.42	0.48
	Online bank operations (0/1)	0.80	0.82	0.80	0.81	0.66
	No bank operations (0/1)	0.03	0.03	0.03	0.04	0.34
Main outcomes	"Big Five": IDK answers (share)	0.15	0.00	0.09	0.07	0.00
	"Big Five": Correct answers (share)	0.53	0.60	0.56	0.57	0.00
	"Big Five": Incorrect answers (share)	0.30	0.38	0.33	0.34	0.00
Other outcomes	Perceived survey difficulty (0-10)	4.23	4.15	4.17	4.13	0.76
	"Big Five": Completion time (sec.)	99.57	102.34	100.77	102.59	0.49
Observations		2,400	1,200	1,200	1,200	

Table 2: Share of "I do not know" Answers: "Big Five" Questions

	(1)	(2)	(3)
Women	0.065***	0.041***	0.041***
	(0.009)	(0.009)	(0.009)
Without IDK	-0.119***	-0.115***	-0.114***
	(0.006)	(0.006)	(0.006)
Incentives	-0.053***	-0.049***	-0.050***
	(0.009)	(0.008)	(0.008)
Information Nudge	-0.062***	-0.063***	-0.060***
	(0.009)	(0.008)	(0.008)
Women x Without IDK	-0.065***	-0.067***	-0.071***
	(0.009)	(0.009)	(0.009)
Women x Incentives	-0.008	-0.015	-0.015
	(0.014)	(0.013)	(0.012)
Women x Information Nudge	-0.036***	-0.038***	-0.041***
	(0.013)	(0.012)	(0.012)
Av outcome for control group men	0.119	0.119	0.119
Controls	No	All	Selected
P-value Test: treatments equal for men	0.000	0.000	0.000
P-value Test: treatments equal for women	0.000	0.000	0.000
Observations	6000	6000	6000
R-squared	0.105	0.239	0.263

Notes: OLS regression of the outcome share of IDK answers in the "Big Five" questions. The first column includes no control variables, the second column includes all control variables specified in Table 1 and the third column includes a lasso-selected set of control variables from a regression of the outcome variable on these control variables and the interactions between them. Table A4 in the Appendix reports the estimated values of the baseline coefficients in each of the columns. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Share of Correct Answers: "Big Five" Questions

	(1)	(2)	(3)
Women	-0.085***	-0.056***	-0.076***
	(0.010)	(0.010)	(0.012)
Without IDK	0.056***	0.052***	0.052***
	(0.012)	(0.011)	(0.011)
Incentives	0.043***	0.040***	0.043***
	(0.013)	(0.012)	(0.011)
Information Nudge	0.020	0.021*	0.021*
	(0.013)	(0.012)	(0.011)
Women x Without IDK	0.021	0.021	0.019
	(0.017)	(0.016)	(0.016)
Women x Incentives	-0.031*	-0.021	-0.024
	(0.018)	(0.016)	(0.016)
Women x Information Nudge	0.028	0.028*	0.029*
	(0.018)	(0.017)	(0.016)
Av outcome for control group men	0.577	0.577	0.577
Controls	No	All	Selected
P-value Test: treatments equal for men	0.042	0.077	0.056
P-value Test: treatments equal for women	0.000	0.000	0.000
Observations	6,000	6,000	6,000
R-squared	0.037	0.176	0.208

Notes: OLS regression of the outcome share of correct answers in the "Big Five" questions. The first column includes no control variables, the second column includes all control variables specified in Table 1 and the third column includes a lasso-selected set of control variables from a regression of the outcome variable on these control variables and the interactions between them. Table A5 in the Appendix reports all the estimated values of the baseline coefficients in each of the columns. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Share of Incorrect Answers: "Big Five" Questions

	(1)	(2)	(3)
Women	0.021**	0.017*	0.007
	(0.009)	(0.009)	(0.011)
Without IDK	0.056***	0.056***	0.054***
	(0.011)	(0.011)	(0.011)
Incentives	0.008	0.006	0.008
	(0.011)	(0.011)	(0.011)
Information Nudge	0.028**	0.029***	0.027**
	(0.011)	(0.011)	(0.011)
Women x Without IDK	0.048***	0.051***	0.056***
	(0.016)	(0.016)	(0.016)
Women x Incentives	0.034**	0.030*	0.030**
	(0.016)	(0.015)	(0.015)
Women x Information Nudge	0.018	0.019	0.019
	(0.016)	(0.016)	(0.015)
Av outcome for control group men	0.293	0.293	0.293
Controls	No	All	Selected
P-value Test: treatments equal for men	0.001	0.000	0.001
P-value Test: treatments equal for women	0.000	0.000	0.000
Observations	6000	6000	6000
R-squared	0.026	0.080	0.117

Notes: OLS regression of the outcome share incorrect answers in the "Big Five" questions. The first column includes no control variables, the second column includes all control variables specified in Table 1 and the third column includes a lasso-selected set of control variables from a regression of the outcome variable on these control variables and the interactions between them. Table A6 in the Appendix reports all the estimated values of the baseline coefficients in each of the columns. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: "Big Five": Random Guessing Benchmark

	(1) Δ Answers	(2) Δ Correct	(3) Δ Correct by Random Guessing	(4) Difference	(5) p-value
			All		
Without IDK	0.15	0.066	0.057	0.010	0.13
Incentives	0.06	0.027	0.022	0.005	0.25
Information Nudge	0.08	0.034	0.030	0.004	0.30
			Women		
Without IDK	0.18	0.077	0.069	0.008	0.23
Incentives	0.06	0.012	0.023	-0.011	0.85
Information Nudge	0.10	0.048	0.037	0.01	0.15
			Men		
Without IDK	0.12	0.056	0.045	0.011	0.18
Incentives	0.05	0.043	0.021	0.022	0.03
Information Nudge	0.06	0.020	0.024	-0.003	0.60
	Δ Answers	Δ Incorrect	Δ Incorrect by Random Guessing	Difference	p-value
			All		
Without IDK	0.15	0.080	0.094	-0.015	0.04
Incentives	0.06	0.025	0.035	-0.009	0.11
Information Nudge	0.08	0.038	0.049	-0.012	0.07
			Women		
Without IDK	0.18	0.104	0.115	-0.011	0.15
Incentives	0.06	0.043	0.038	-0.005	0.67
Information Nudge	0.10	0.046	0.061	-0.015	0.09
			Men		
Without IDK	0.12	0.056	0.074	-0.018	0.06
Incentives	0.05	0.008	0.032	-0.024	0.02
Information Nudge	0.06	0.029	0.038	-0.009	0.20

Notes: The first column reports the treatment-induced increase in the share of provided answers relative to the control group. The second column reports the observed increase in the share of correct answers (top panel) or incorrect answers (bottom panel). The third column reports the corresponding expected increase under pure random guessing, calculated question by question using the number of available answer options. The fourth column reports the difference between the observed (column (2)) and random-guessing changes (column (3)). Column (5) reports one-sided bootstrap p-values for the hypothesis that the observed response shift is no more accurate than random guessing: in the correct-answer panel, $p = \Pr(\widehat{diff}^{\,b} \leq 0)$; in the incorrect-answer panel, $p = \Pr(\widehat{diff}^{\,b} \geq 0)$, where $\widehat{diff}^{\,b}$ denotes the bootstrapped difference.

Table 6: Association: IDK & Financial Indicators

	(1) Financial behavior	(2) Objective financial well-being	(3) Subjective financial well-being	(4) Financial products awareness	(5) Financial products holding
Women	0.997**	-0.839*	0.903**	0.175***	-0.043
	(0.476)	(0.499)	(0.460)	(0.049)	(0.049)
Correct (share)	17.138***	17.498***	12.535***	2.849***	2.707***
	(1.504)	(1.560)	(1.457)	(0.163)	(0.159)
IDK (share)	-15.259***	-4.341*	-0.075	-2.483***	-1.089***
	(1.871)	(2.243)	(1.829)	(0.258)	(0.181)
Constant	51.376***	25.170***	17.948***	7.851***	1.627***
	(1.256)	(1.334)	(1.204)	(0.142)	(0.130)
Observations	7764	7754	7764	7764	7764

Notes: OLS regression of the each financial indicator on indicator for women participants, the share of correct financial knowledge questions (out of seven) following the OECD measures, and the share of IDK anwers to these same financial knowledge questions. The first column reports the correlation of IDK to financial behavior index with a score equal to 100. This behavior score is built on 8 questions on the following savvy behaviors: keeping track of money, actively saving, not borrowing to make ends meet, seeking advise, watching personal financial affairs, setting long-term financial goals, making considered purchases, and paying bills on time. The second column outcome is objective financial well-being as measured by having income to cover expenses if the main source of income had been lost in the last 12 months, and being able to cover living expenses without borrowing or moving homes for at least three months if the main source of income was lost. The third column outcome is subjective financial well-being as measured by being satisfied with their present financial situation, disagreeing that the financial situation limits their ability, disagreeing with having too much debt, never tending to worry about paying normal living expenses, and not being concerned that money won t last. The forth column outcome is financial product awareness as measured by the sum of being aware of at least one financial product on the categories of payment products, saving-investment-retirement product, insurance products, credit products (including a mortgage), or crypto products. The fifth and last column outcome is financial product holding as measured by the sum of holding at least 1 financial product on each category as in product awareness. In this case, product holding can be shared with other household members. See OECD (2023) Appendix C for a detailed definition and construction of each measurement. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

A Additional Figures and Tables

Table A1: Sample Composition Relative to the Spanish Population (Ages 25–70)

	(1) Experimental sample	(2) ECF 2021 (weighted)	(3) Difference (1)–(2)	(4) p-value
Women (%)	50.0	50.5	-0.5	0.55
Age				
Mean age	46.1	47.9	-1.7	< 0.01
25–34 (%)	18.3	17.7	0.6	0.39
35–54 (%)	56.5	49.4	7.1	< 0.01
55–70 (%)	25.2	32.9	-7.7	< 0.01
Education				
Primary or less (%)	6.7	13.2	-6.5	< 0.01
Secondary (%)	48.8	48.4	0.4	0.66
Post-secondary (%)	44.5	38.4	6.1	< 0.01
Employment status				
Employed (%)	68.5	67.2	1.3	0.13

Notes: The experimental sample includes all 6,000 respondents who completed the online survey, without reweighting. The population benchmark is the 2021 Spanish Survey of Financial Competences (ECF), restricted to respondents aged 25–70 and weighted using its survey weights. Differences are expressed in percentage points, except for mean age (years), and are computed using unrounded values. The final column reports two-sided p-values for tests of equality of means or proportions, assuming independent samples. The experimental sample was stratified by gender by design.

Table A2: Descriptive Statistics (mean values) and Randomization: Women Sample

		(1) Control	(2) Without IDK	(3) Incentives	(4) Information Nudge	(5) p-value
Demographics	Age 25-34	0.20	0.21	0.20	0.17	0.35
	Age 35-44	0.28	0.30	0.28	0.27	0.77
	Age 45-54	0.31	0.28	0.29	0.33	0.39
	Age 55-70	0.21	0.21	0.23	0.23	0.63
	Spaniard	0.90	0.91	0.92	0.90	0.56
	Pop. size 0-20th	0.22	0.17	0.20	0.20	0.04
	Pop. size 20th-100th	0.26	0.28	0.28	0.29	0.49
	Pop. size 100th+	0.52	0.55	0.52	0.51	0.37
	Primary education	0.21	0.15	0.26	0.20	0.00
	Secondary education	0.31	0.36	0.36	0.35	0.06
	University education	0.36	0.37	0.28	0.36	0.00
	Master, PhD education	0.12	0.13	0.10	0.08	0.02
	Working	0.67	0.66	0.63	0.61	0.08
	Retired	0.07	0.07	0.08	0.09	0.45
	Unemployed	0.25	0.24	0.28	0.29	0.18
Household	< 1 bookshelf at age 10	0.71	0.73	0.70	0.73	0.64
	> 2 bookshelves at age 10	0.29	0.27	0.30	0.27	0.64
	Household size	3.03	2.98	3.02	3.02	0.85
	Primary earner	0.51	0.50	0.54	0.49	0.35
	Lives with partner	0.72	0.70	0.69	0.70	0.61
	Mother: Primary education	0.59	0.57	0.61	0.60	0.40
	Mother: Secondary education	0.19	0.22	0.20	0.19	0.52
	Mother: Post-secondary education	0.20	0.20	0.17	0.19	0.29
	Father: Primary education	0.54	0.51	0.56	0.54	0.39
	Father: Secondary education	0.20	0.20	0.20	0.22	0.89
	Father: Post-secondary education	0.21	0.25	0.19	0.20	0.10
	Partner: Primary education	0.19	0.16	0.19	0.18	0.38
	Partner: Secondary education	0.23	0.24	0.26	0.25	0.50
	Partner: Post-secondary education	0.30	0.30	0.25	0.27	0.06
Assessments	Very low financial knowledge	0.03	0.02	0.03	0.02	0.29
	Low financial knowledge	0.15	0.16	0.16	0.14	0.84
	Neutral financial knowledge	0.44	0.43	0.39	0.47	0.07
	Good financial knowledge	0.33	0.37	0.38	0.34	0.11
	Very good financial knowledge	0.05	0.03	0.04	0.03	0.03
	Expected correct answers (0-10)	5.22	5.37	5.35	5.17	0.28
	Interest in finance (0-10)	5.88	5.73	5.80	5.70	0.50
	Risk willingness (0-10)	4.24	4.34	4.32	4.18	0.75
	Lottery choice (1-8)	3.59	3.48	3.48	3.58	0.69
Perceptions	Lean-in index (mean of five z-scores)	-0.12	-0.11	-0.16	-0.13	0.41
	Perceived self-efficacy (1-5)	3.96	3.98	4.00	4.02	0.56
	Perceived confidence (1-5)	3.82	3.85	3.79	3.83	0.73
	Perceived lean-in (1-5)	3.64	3.67	3.62	3.65	0.88
Managing finances	Saving products (0-6)	2.35	2.42	2.21	2.33	0.04
	Debt products (0-3)	1.35	1.39	1.32	1.34	0.51
	Online bank operations (0/1)	0.79	0.81	0.78	0.80	0.71
	No bank operations (0/1)	0.04	0.04	0.04	0.06	0.27
Main outcomes	"Big Five": IDK answers (share)	0.18	0.00	0.12	0.09	0.00
	"Big Five": Correct answers (share)	0.49	0.57	0.50	0.54	0.00
	"Big Five": Incorrect answers (share)	0.31	0.42	0.36	0.36	0.00
Other outcomes	Perceived survey difficulty (0-10)	4.41	4.45	4.36	4.38	0.94
	"Big Five": Completion time (sec.)	99.23	108.92	103.41	104.36	0.02
Observations		1,200	600	600	600	

Table A3: Descriptive Statistics (mean values) and Randomization: Men Sample

		(1) Control	(2) Without IDK	(3) Incentives	(4) Information Nudge	(5) p-value
Demographics	Age 25-34	0.17	0.17	0.15	0.18	0.49
	Age 35-44	0.25	0.25	0.24	0.26	0.87
	Age 45-54	0.30	0.31	0.30	0.28	0.58
	Age 55-70	0.29	0.27	0.30	0.28	0.63
	Spaniard	0.92	0.94	0.94	0.92	0.25
	Pop. size 0-20th	0.20	0.20	0.17	0.20	0.74
	Pop. size 20th-100th	0.28	0.24	0.28	0.25	0.15
	Pop. size 100th+	0.52	0.56	0.54	0.56	0.32
	Primary education	0.20	0.20	0.23	0.21	0.22
	Secondary education	0.36	0.38	0.34	0.35	0.61
	University education	0.34	0.34	0.32	0.31	0.61
	Master, PhD education	0.11	0.10	0.11	0.13	0.30
	Working	0.73	0.72	0.71	0.72	0.77
	Retired	0.14	0.12	0.15	0.13	0.45
	Unemployed	0.12	0.14	0.13	0.13	0.72
Household	< 1 bookshelf at age 10	0.78	0.76	0.76	0.73	0.19
	> 2 bookshelves at age 10	0.22	0.24	0.24	0.27	0.19
	Household size	2.95	2.97	3.06	2.92	0.15
	Primary earner	0.83	0.81	0.83	0.83	0.76
	Lives with partner	0.75	0.73	0.75	0.76	0.69
	Mother: Primary education	0.61	0.58	0.57	0.59	0.60
	Mother: Secondary education	0.18	0.20	0.19	0.20	0.75
	Mother: Post-secondary education	0.19	0.18	0.20	0.18	0.71
	Father: Primary education	0.52	0.54	0.54	0.57	0.18
	Father: Secondary education	0.21	0.21	0.20	0.16	0.09
	Father: Post-secondary education	0.24	0.21	0.22	0.23	0.62
	Partner: Primary education	0.15	0.14	0.17	0.17	0.24
	Partner: Secondary education	0.24	0.26	0.23	0.21	0.31
	Partner: Post-secondary education	0.35	0.34	0.35	0.37	0.71
Assessments	Very low financial knowledge	0.01	0.02	0.01	0.03	0.01
	Low financial knowledge	0.09	0.09	0.09	0.09	0.99
	Neutral financial knowledge	0.40	0.38	0.40	0.41	0.84
	Good financial knowledge	0.43	0.46	0.42	0.40	0.25
	Very good financial knowledge	0.07	0.05	0.08	0.07	0.19
	Expected correct answers (0-10)	5.94	6.09	6.23	5.87	0.01
	Interest in finance (0-10)	6.32	6.53	6.48	6.25	0.15
	Risk willingness (0-10)	5.05	5.19	5.17	5.06	0.69
	Lottery choice (1-8)	3.66	3.68	3.87	3.72	0.35
Perceptions	Lean-in index (mean of five z-scores)	0.13	0.12	0.10	0.13	0.80
	Perceived self-efficacy (1-5)	3.96	4.01	4.01	4.02	0.41
	Perceived confidence (1-5)	3.79	3.90	3.87	3.82	0.06
	Perceived lean-in (1-5)	3.65	3.67	3.65	3.62	0.87
Managing finances	Saving products (0-6)	2.70	2.79	2.71	2.78	0.52
	Debt products (0-3)	1.44	1.38	1.41	1.50	0.09
	Online bank operations (0/1)	0.82	0.83	0.82	0.82	0.93
	No bank operations (0/1)	0.03	0.01	0.02	0.02	0.38
Main outcomes	"Big Five": IDK answers (share)	0.12	0.00	0.07	0.06	0.00
	"Big Five": Correct answers (share)	0.58	0.63	0.62	0.60	0.00
	"Big Five": Incorrect answers (share	0.29	0.35	0.30	0.32	0.00
Other outcomes	Perceived survey difficulty (0-10)	4.04	3.86	3.99	3.88	0.50
	"Big Five": Completion time (sec.)	99.92	95.72	98.16	100.79	0.51
Observations		1,200	600	600	600	

Table A4: Share of IDK Answers: "Big Five" Questions

	(1)	(2)	(3)
Women	0.065***	0.041***	0.041***
	(0.009)	(0.009)	(0.009)
Without IDK	-0.119***	-0.115***	-0.114***
	(0.006)	(0.006)	(0.006)
Incentives	-0.053***	-0.049***	-0.050***
	(0.009)	(0.008)	(0.008)
Information Nudge	-0.062***	-0.063***	-0.060***
	(0.009)	(0.008)	(0.008)
Women x Without IDK	-0.065***	-0.067***	-0.071***
	(0.009)	(0.009)	(0.009)
Women x Incentives	-0.008	-0.015	-0.015
	(0.014)	(0.013)	(0.012)
Women x Information Nudge	-0.036***	-0.038***	-0.041***
	(0.013)	(0.012)	(0.012)
Age 35-44		-0.001	
		(0.007)	
Age 45-54		-0.007	
		(0.007)	
Age 55-70		-0.018**	
		(0.007)	
Spaniard		-0.002	
		(0.008)	
Pop. size 20th-100th		-0.002	
		(0.007)	
Pop. size 100th+		-0.004	
		(0.006)	
Secondary education		-0.005	
		(0.008)	
University education		-0.016*	
		(0.008)	
Master, PhD education		-0.014	
		(0.010)	
Working		-0.002	
		(0.018)	
Retired		-0.003	
		(0.019)	
Unemployed		0.002	
		(0.018)	
> 2 bookshelves at age 10		0.002	
		(0.005)	
Household size		0.002	
		(0.002)	
Primary earner		-0.008	
		(0.006)	
Lives with partner		-0.029***	
		(0.008)	
Mother: Secondary education		0.009	
		(0.006)	
Mother: Post-secondary education		0.001	
		(0.007)	
Father: Secondary education		-0.007	
		(0.006)	
Father: Post-secondary education		-0.001	
		(0.007)	
Partner: Secondary education		0.010	
		(0.007)	
Partner: Post-secondary education		0.014*	
		(0.008)	
Low financial knowledge		-0.056**	
		(0.026)	
Neutral financial knowledge		-0.082***	
		(0.025)	
Good financial knowledge		-0.106***	-0.011
		(0.026)	(0.007)
Very good financial knowledge		-0.113***	-0.010
		(0.027)	(0.011)
Interest in finance		-0.008***	-0.008***
		(0.001)	(0.002)
Expected correct answers		-0.001	-0.001
		(0.002)	(0.002)
Risk willingness		-0.002**	-0.002**
		(0.001)	(0.001)
Lottery choice		-0.001	
		(0.001)	
Saving products (N)		-0.004**	-0.005***
		(0.002)	(0.002)
Debt products (N)		-0.006**	
		(0.003)	
Online bank operations		-0.006	
		(0.007)	
No bank operations		0.104***	
		(0.024)	
Lean-in index		-0.016***	0.005
		(0.003)	(0.010)
Perceived self-efficacy		-0.004	
		(0.003)	
Perceived confidence		-0.013***	-0.014***
		(0.004)	(0.003)
Perceived lean-in		-0.006**	-0.005**
Constant	0.119***	0.433***	0.299***
	(0.006)	(0.035)	(0.016)
Controls	No	All	Selected
Observations	6000	6000	6000
R2	0.105	0.239	0.263

Notes: OLS regression of the outcome share of IDK answers in the "Big Five" questions. The first column includes no control variables, the second column includes all control variables, and the third column includes the coefficients from the set of a lasso-selected control variables, excluding the selected interaction terms. The reference categories for categorical controls are: age 25–34, population size 0–20th, primary education, other labor-market status, fewer than one bookshelf at age 10, primary education for mother, father, and partner, and very low self-assessed financial knowledge. Men and the control group are the reference categories for gender and treatment, respectively; for binary indicators, the reference category is zero. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Share of Correct Answers: "Big Five" Questions

	(1)	(2)	(3)
Women	-0.085***	-0.056***	-0.076***
	(0.010)	(0.010)	(0.012)
Without IDK	0.056***	0.052***	0.052***
	(0.012)	(0.011)	(0.011)
Incentives	0.043***	0.040***	0.043***
	(0.013)	(0.012)	(0.011)
Information Nudge	0.020	0.021*	0.021*
	(0.013)	(0.012)	(0.012)
Women x Without IDK	0.021	0.021	0.020
	(0.017)	(0.016)	(0.016)
Women x Incentives	-0.031*	-0.021	-0.024
	(0.018)	(0.016)	(0.016)
Women x Information Nudge	0.028	0.028*	0.030*
	(0.018)	(0.017)	(0.016)
Age 35-44		0.013	
		(0.009)	
Age 45-54		0.036***	
		(0.009)	
Age 55-70		0.071***	
		(0.010)	
Spaniard		0.008	
		(0.011)	
Pop. size 20th-100th		0.007	
		(0.009)	
Pop. size 100th+		-0.000	
		(0.008)	
Secondary education		0.033***	
		(0.009)	
University education		0.079***	0.027
		(0.011)	(0.042)
Master, PhD education		0.074***	
		(0.014)	
Working		-0.002	
		(0.025)	
Retired		0.003	
		(0.027)	
Unemployed		-0.010	
		(0.025)	
> 2 bookshelves at age 10		0.009	
		(0.007)	
Household size		-0.005*	-0.006*
		(0.003)	(0.003)
Primary earner		0.004	
		(0.008)	
Lives with partner		0.012	
		(0.010)	
Mother: Secondary education		-0.021**	
		(0.009)	
Mother: Post-secondary education		-0.018*	
		(0.011)	
Father: Secondary education		0.001	
		(0.008)	
Father: Post-secondary education		-0.011	
		(0.010)	
Partner: Secondary education		-0.000	
		(0.010)	
Partner: Post-secondary education		0.010	
		(0.011)	
Low financial knowledge		0.031	
		(0.023)	
Neutral financial knowledge		0.030	
		(0.023)	
Good financial knowledge		0.033	
		(0.024)	
Very good financial knowledge		0.011	
		(0.028)	
Interest in finance		0.011***	0.012***
		(0.002)	(0.003)
Expected correct answers		0.008***	
		(0.002)	
Risk willingness		-0.002*	
		(0.001)	
Lottery choice		-0.002	
		(0.001)	
Saving products (N)		0.011***	0.006*
		(0.003)	(0.004)
Debt products (N)		0.005	
		(0.004)	
Online bank operations		0.062***	
		(0.009)	
No bank operations		-0.050**	
		(0.021)	
Lean-in index		0.026***	
		(0.005)	
Perceived self-efficacy		0.011**	0.011**
		(0.004)	(0.005)
Perceived confidence		0.007	0.006
		(0.005)	(0.005)
Perceived lean-in		0.005*	0.004
		(0.003)	(0.003)
Constant	0.577***	0.189***	0.305***
	(0.007)	(0.039)	(0.024)
Controls	No	All	Selected
Observations	6000	6000	6000
R2	0.037	0.176	0.208

Notes: OLS regression of the outcome share of correct answers in the "Big Five" questions. The first column includes no control variables, the second column includes all control variables, and the third column includes a lasso-selected set of control variables. The reference categories for categorical controls are: age 25–34, population size 0–20th, primary education, other labor-market status, fewer than one bookshelf at age 10, primary education for mother, father, and partner, and very low self-assessed financial knowledge. Men and the control group are the reference categories for gender and treatment, respectively; for binary indicators, the reference category is zero. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Share of Incorrect Answers: "Big Five" Questions

	(1)	(2)	(3)
Women	0.021**	0.017*	0.007
	(0.009)	(0.009)	(0.011)
Without IDK	0.056***	0.056***	0.054***
	(0.011)	(0.011)	(0.011)
Incentives	0.008	0.006	0.008
	(0.011)	(0.011)	(0.011)
Information Nudge	0.028**	0.029***	0.027**
	(0.011)	(0.011)	(0.011)
Women x Without IDK	0.048***	0.051***	0.056***
	(0.016)	(0.016)	(0.016)
Women x Incentives	0.034**	0.030*	0.030*
	(0.016)	(0.015)	(0.015)
Women x Information Nudge	0.018	0.019	0.020
	(0.016)	(0.016)	(0.015)
Age 35-44		-0.017*	
		(0.009)	
Age 45-54		-0.040***	
		(0.009)	
Age 55-70		-0.057***	
		(0.010)	
Spaniard		-0.004	
		(0.011)	
Pop. size 20th-100th		-0.006	
		(0.008)	
Pop. size 100th+		0.000	
		(0.008)	
Secondary education		-0.031***	
		(0.009)	
University education		-0.069***	-0.014
		(0.010)	(0.033)
Master, PhD education		-0.078***	
		(0.013)	
Working		0.010	
		(0.023)	
Retired		0.011	
		(0.026)	
Unemployed		0.015	
		(0.024)	
> 2 bookshelves at age 10		-0.008	
		(0.007)	
Household size		0.002	
		(0.003)	
Primary earner		0.002	
		(0.007)	
Lives with partner		0.013	
		(0.010)	
Mother: Secondary education		0.003	
		(0.008)	
Mother: Post-secondary education		0.014	
		(0.010)	
Father: Secondary education		0.008	
		(0.008)	
Father: Post-secondary education		0.016	
		(0.010)	
Partner: Secondary education		-0.011	
		(0.010)	
Partner: Post-secondary education		-0.018*	
		(0.010)	
Low financial knowledge		0.013	
		(0.024)	
Neutral financial knowledge		0.040*	
		(0.023)	
Good financial knowledge		0.063***	
		(0.024)	
Very good financial knowledge		0.086***	
		(0.028)	
Interest in finance		-0.002	
		(0.002)	
Expected correct answers		-0.007***	
		(0.002)	
Risk willingness		0.005***	0.002
		(0.001)	(0.002)
Lottery choice		0.003***	0.003
		(0.001)	(0.004)
Saving products (N)		-0.006**	0.005
		(0.002)	(0.008)
Debt products (N)		-0.002	
		(0.003)	
Online bank operations		-0.054***	-0.003
		(0.008)	(0.020)
No bank operations		-0.051***	
		(0.020)	
Lean-in index		-0.008	
		(0.005)	
Perceived self-efficacy		-0.006	
		(0.004)	
Perceived confidence		0.008*	
		(0.005)	
Perceived lean-in		0.001	
		(0.003)	
Constant	0.293***	0.369***	0.358***
	(0.006)	(0.038)	(0.017)
Controls	No	All	Selected
Observations	6000	6000	6000
R2	0.026	0.080	0.117

Notes: OLS regression of the outcome Share of incorrect answers in the "Big Five" questions. The first column includes no control variables, the second column includes all control variables, and the third column includes a lasso-selected set of control variables. The reference categories for categorical controls are: age 25–34, population size 0–20th, primary education, other labor-market status, fewer than one bookshelf at age 10, primary education for mother, father, and partner, and very low self-assessed financial knowledge. Men and the control group are the reference categories for gender and treatment, respectively; for binary indicators, the reference category is zero. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Minimum Detectable Effects for Gender-by-Treatment Interactions (percentage points)

Outcome	Treatment	Estimate	Robust SE	Analytical MDE	Simulated MDE	Simulated power
IDK	Incentives	−0.78	1.39	3.90	3.91	0.808
IDK	Information Nudge	−3.62	1.34	3.75	3.71	0.805
Correct	Without IDK	2.07	1.70	4.77	4.73	0.805
Correct	Incentives	−3.10	1.77	4.96	4.96	0.805
Correct	Information Nudge	2.77	1.78	4.98	4.84	0.804

Notes: This table reports minimum detectable effects (MDEs) for the gender-by-treatment interactions in column (1) of Tables 2 and 3. The dependent variables are defined as shares (ranging from 0 to 1) in the main regression tables. For ease of interpretation, the estimated interactions, robust standard errors, and MDEs reported here are multiplied by 100 and thus expressed in percentage points. Analytical MDEs use a two-sided 5 percent significance level and 80 percent power. Simulation-based MDEs use 2,000 Monte Carlo replications and correspond to the smallest imposed absolute interaction effect for which the null hypothesis of no gender-differential treatment effect is rejected in at least 80 percent of replications. Simulated power reports the corresponding rejection frequency.

Table A8: Heterogeneous Treatment Effects by Education, Age and Subjective Financial Knowledge

		Education		Age		Subj. fin. knowledge	
	Full sample	≤ Secondary	Post-sec.	25–45	46–70	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A. Share of IDK answers						
Women	0.065***	0.066***	0.069***	0.045***	0.084***	0.048***	0.050***
	(0.009)	(0.014)	(0.011)	(0.013)	(0.013)	(0.014)	(0.010)
Without IDK	-0.119***	-0.150***	-0.079***	-0.130***	-0.109***	-0.185***	-0.053***
	(0.006)	(0.009)	(0.007)	(0.009)	(0.008)	(0.010)	(0.005)
Incentives	-0.053***	-0.066***	-0.037***	-0.056***	-0.050***	-0.077***	-0.029***
	(0.009)	(0.013)	(0.011)	(0.014)	(0.011)	(0.015)	(0.007)
Information Nudge	-0.062***	-0.066***	-0.056***	-0.057***	-0.067***	-0.103***	-0.024***
	(0.009)	(0.013)	(0.009)	(0.014)	(0.011)	(0.014)	(0.008)
Women × Without IDK	-0.065***	-0.066***	-0.069***	-0.045***	-0.084***	-0.048***	-0.050***
	(0.009)	(0.014)	(0.011)	(0.013)	(0.013)	(0.014)	(0.010)
Women × Incentives	-0.008	-0.003	-0.027	-0.004	-0.012	0.017	-0.021
	(0.014)	(0.020)	(0.018)	(0.020)	(0.019)	(0.022)	(0.014)
Women × Information Nudge	-0.036***	-0.051**	-0.023	-0.038*	-0.035*	-0.022	-0.031**
	(0.013)	(0.021)	(0.015)	(0.020)	(0.018)	(0.020)	(0.014)
Av. outcome, control-group men	0.119	0.150	0.079	0.130	0.109	0.185	0.053
Av. outcome, control-group women	0.184	0.216	0.149	0.175	0.193	0.233	0.103
p-value: level effects equal (Incentives, Nudge)		0.226	0.226	0.680	0.680	0.000	0.000
p-value: Women × treatment equal (Incentives, Nudge)		0.154	0.154	0.935	0.935	0.311	0.311
p-value: Women × Nudge equal		0.266	0.266	0.919	0.919	0.712	0.712
	Panel B. Share of correct answers						
Women	-0.085***	-0.077***	-0.104***	-0.057***	-0.108***	-0.061***	-0.088***
	(0.010)	(0.014)	(0.015)	(0.015)	(0.015)	(0.014)	(0.015)
Without IDK	0.056***	0.080***	0.027	0.051***	0.061***	0.085***	0.026
	(0.012)	(0.016)	(0.018)	(0.018)	(0.017)	(0.018)	(0.016)
Incentives	0.043***	0.051***	0.035*	0.036*	0.045***	0.050***	0.035**
	(0.013)	(0.017)	(0.019)	(0.019)	(0.017)	(0.018)	(0.017)
Information Nudge	0.020	0.028	0.012	0.020	0.023	0.050***	-0.006
	(0.013)	(0.017)	(0.018)	(0.019)	(0.017)	(0.018)	(0.017)
Women × Without IDK	0.021	0.002	0.042*	0.023	0.020	0.012	0.016
	(0.017)	(0.023)	(0.024)	(0.024)	(0.024)	(0.023)	(0.025)
Women × Incentives	-0.031*	-0.045*	0.008	-0.040	-0.016	-0.035	-0.037
	(0.018)	(0.023)	(0.027)	(0.025)	(0.025)	(0.024)	(0.025)
Women × Information Nudge	0.028	0.025	0.037	0.027	0.026	0.012	0.032
	(0.018)	(0.024)	(0.025)	(0.025)	(0.025)	(0.024)	(0.026)
Av. outcome, control-group men	0.577	0.527	0.640	0.544	0.604	0.520	0.633
Av. outcome, control-group women	0.492	0.450	0.536	0.488	0.496	0.459	0.545
p-value: level effects equal (all treatments)		0.194	0.194	0.972	0.972	0.036	0.036
p-value: Women × treatment equal (all treatments)		0.401	0.401	0.896	0.896	0.943	0.943
p-value: Women × Nudge equal		0.743	0.743	0.989	0.989	0.575	0.575
Observations	6,000	3,329	2,671	2,910	3,090	3,346	2,654

Notes: Each column reports a separate OLS regression of the outcome on the three treatment indicators, a Women indicator, and their interactions, estimated on the indicated subsample; robust standard errors in parentheses. Column (1) uses the full sample. Education splits at completed secondary, age at 46, and subjective financial knowledge at "neutral" or below versus "good"/"very good"; all three are measured before the treatments are administered. The reported p-values are for joint Wald tests of the null that the indicated coefficients are equal across the two subsamples of each dimension, computed in the corresponding fully interacted specification. In Panel A the Without IDK terms are mechanical, since the treatment sets IDK use to zero by construction, and they are excluded from the joint tests; in Panel B they are behavioural and are included. The number of observations is reported once, at the foot of Panel B. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Adjusted p-values for the Effect of Interventions on the Share of "I do not know", Correct and Incorrect for the "Big Five" Questions

	"I do not know"		Correct		Incorrect	
	(1) Estimate	(2) FDR $q-$value	(3) Estimate	(4) FDR $q-$value	(5) Estimate	(6) FDR $q-$value
Women	0.041***	0.001	-0.076***	0.002	0.007	0.511
	(0.009)	[0.000]	(0.012)	[0.000]	(0.011)	[0.051]
Without IDK	-0.114***	0.001	0.052***	0.001	0.054***	0.001
	(0.006)	[0.000]	(0.011)	[0.000]	(0.011)	[0.000]
Incentives	-0.050***	0.001	0.043***	0.001	0.008	0.493
	(0.008)	[0.000]	(0.011)	[0.000]	(0.011)	[0.469]
Information Nudge	-0.060***	0.001	0.021*	0.115	0.027**	0.025
	(0.008)	[0.000]	(0.011)	[0.082]	(0.011)	[0.014]
Women x Without IDK	-0.071***	0.001	0.019	0.228	0.056***	0.001
	(0.009)	[0.000]	(0.016)	[0.194]	(0.016)	[0.000]
Women x Incentives	-0.015	0.262	-0.024	0.207	0.030**	0.083
	(0.012)	[0.237]	(0.016)	[0.157]	(0.015)	[0.051]
Women x Information Nudge	-0.041***	0.002	0.029*	0.087	0.019	0.228
	(0.012)	[0.001]	(0.016)	[0.058]	(0.015)	[0.195]

Notes: This table reports the adjusted $p-$values for the impact of the interventions on the "Big Five" share of IDK, correct, and incorrect answers. Columns (1), (3), and (5) reproduce the results for each outcome from Column (3) of tables 2, 3 and 4. Columns (2), (4) and (6) report the false discovery rate (FDR) adjusted $p-$values ("$q-$values"), standard $p-$values are reported below each adjusted $p-$value in brackets. The FDR adjusted $p-$values control for the number of false positives when multiple hypotheses are tested following Anderson (2008).

Table A10: Share of IDK Answers Question by Question: "Big Five" Questions

	(1) Inflation	(2) Compound Interest Rate	(3) Risk Diversification	(4) Mortgages	(5) Bond Pricing
Women	0.027***	0.023**	0.064***	0.010	0.079***
	(0.010)	(0.011)	(0.017)	(0.013)	(0.017)
Without IDK	-0.046***	-0.051***	-0.196***	-0.090***	-0.193***
	(0.006)	(0.006)	(0.011)	(0.008)	(0.012)
Incentives	-0.013	-0.017*	-0.100***	-0.040***	-0.073***
	(0.009)	(0.010)	(0.016)	(0.012)	(0.017)
Information Nudge	-0.024***	-0.031***	-0.123***	-0.039***	-0.098***
	(0.009)	(0.009)	(0.015)	(0.012)	(0.017)
Women x Without IDK	-0.045***	-0.039***	-0.104***	-0.030**	-0.118***
	(0.011)	(0.011)	(0.018)	(0.013)	(0.018)
Women x Incentives	-0.017	-0.012	-0.003	-0.012	-0.029
	(0.016)	(0.016)	(0.026)	(0.018)	(0.026)
Women x Information Nudge	-0.029**	-0.010	-0.051**	-0.022	-0.079***
	(0.014)	(0.015)	(0.024)	(0.018)	(0.025)
Constant	0.317***	0.265***	0.622***	0.353***	0.609***
	(0.045)	(0.044)	(0.063)	(0.051)	(0.063)
Controls	Yes	Yes	Yes	Yes	Yes
Observations	6000	6000	6000	6000	6000
R2	0.101	0.091	0.163	0.102	0.149

Notes: OLS regression of the outcome share of IDK answers in each of the "Big Five" questions with all control variables included. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A11: Share of IDK Answers with "Big Five", "Big Three", "Big Six" and All Nine Questions

	(1) "Big Five"	(2) "Big Five"	(3) "Big Three"	(4) "Big Six"	(5) All Nine
Women	0.041***	0.040***	0.038***	0.040***	0.037***
	(0.009)	(0.009)	(0.009)	(0.008)	(0.007)
Without IDK	-0.115***	-0.108***	-0.098***	-0.113***	-0.108***
	(0.006)	(0.008)	(0.006)	(0.006)	(0.005)
Incentives	-0.049***	-0.046***	-0.043***	-0.047***	-0.044***
	(0.008)	(0.009)	(0.008)	(0.008)	(0.007)
Information Nudge	-0.063***	-0.050***	-0.059***	-0.061***	-0.056***
	(0.008)	(0.010)	(0.008)	(0.008)	(0.007)
Women x Without IDK	-0.067***	-0.072***	-0.063***	-0.064***	-0.057***
	(0.009)	(0.012)	(0.009)	(0.009)	(0.008)
Women x Incentives	-0.015	-0.009	-0.011	-0.012	-0.005
	(0.013)	(0.014)	(0.013)	(0.012)	(0.011)
Women x Information Nudge	-0.038***	-0.048***	-0.030**	-0.032***	-0.027**
	(0.012)	(0.014)	(0.013)	(0.012)	(0.011)
Constant	0.433***	0.442***	0.401***	0.429***	0.400***
	(0.035)	(0.037)	(0.037)	(0.033)	(0.029)
Controls	All	All	All	All	All
Observations	6000	6000	6000	6000	6000
R2	0.239	0.191	0.204	0.249	0.251

Notes: OLS regression of the outcome share of IDK answers in different set of question with all control variables included. Column (1) reproduces our main results for the "Big Five" answers, Column (2) shows the results for the "Big Five" answers redefining the IDK to include the skipped answers, Column (3) uses the "Big Three" answers, corresponding to inflation, compound interest rate and risk diversification, Column (4) adds the simple interest rate answer to the "Big Five", and Column (5) includes all questions of the section. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A12: Exit by Survey Screen

		Exits	Percent	Cumulative	Men	Women
		By screen exit				
Demographics	Gender	115	13.08	13.08	–	–
+ household	Age	51	5.80	18.89	19	32
	Country of birth	10	1.14	20.02	4	6
	Books at home at age 10	39	4.44	24.46	15	24
	Place of residence	33	3.75	28.21	15	18
	Living with a partner	21	2.39	30.60	12	9
	Marital status	6	0.68	31.29	4	2
	Household size	3	0.34	31.63	0	3
	Household income earners	9	1.02	32.65	3	6
	Main household earner	8	0.91	33.56	4	4
	Own education[a]	116	13.20	46.76	28	88
	Labour-market status	12	1.37	48.12	2	10
	Parental education[b]	169	19.23	67.35	63	106
Assessments	Self-assessed financial knowledge	17	1.93	69.28	7	10
	Expected correct answers	11	1.25	70.53	3	8
	Interest in finance	8	0.91	71.44	5	3
	Risk willingness	7	0.80	72.24	2	5
	Lottery choice	24	2.73	74.97	12	12
Randomization	Instructions	61	6.94	81.91	29	32
Financial knowledge	Inflation	26	2.96	84.87	11	15
	Simple interest	30	3.41	88.28	12	18
	Compound interest	15	1.71	89.99	8	7
	Risk diversification	4	0.46	90.44	1	3
	Mortgages	3	0.34	90.78	1	2
	Bond pricing	3	0.34	91.13	1	2
	Lottery probability	12	1.37	92.49	5	7
	Cognitive reflection	8	0.91	93.40	3	5
	Attention check	5	0.57	93.97	3	2
Perceptions +	Expected correct answers (post)	2	0.23	94.20	1	1
	Beliefs about scores of others	7	0.80	94.99	2	5
Managing finances	Financial products held	11	1.25	96.25	8	3
	Banking channels	10	1.14	97.38	6	4
	Lean-in experiences	5	0.57	97.95	1	4
	Self-perception	5	0.57	98.52	4	1
	Difficulty of the survey	2	0.23	98.75	1	1
	Final open comment	11	1.25	100.00	1	10
Before randomization		659	74.97		198	346
After randomization		220	25.03		98	122
Total		879	100.00		296	468

Notes: The table shows in which screen individuals dropped out among those that did not complete the survey. Each exit is assigned to the screen after the last one with a recorded answer. We only consider questions asked to all respondents. Items on FL section could be skipped, so respondent may have left on a later screen than the one shown. Exits column show total individuals leaving on that screen, Percent and Cumulative are shares of all exits. Men and Women columns show total exits by each gender. Men and women do not sum to the total because the 115 respondents left before gender was answered. The split before and after randomization depends only on whether a treatment assignment was recorded. The treatment assignment happened on the Instructions screen.

Table A13: Characteristics of Respondents who Exit and who Complete the Survey, by Gender

	Men			Women			Difference
	Completed	Exited	*N* exited	Completed	Exited	*N* exited	*p*-value
Panel A. Demographics							
Age	47.05	44.93	277	45.18	45.24	436	0.02
Spaniard	0.93	0.91	273	0.91	0.87	430	0.53
University education	0.33	0.30	192	0.35	0.28	276	0.38
Master, PhD education	0.11	0.05	192	0.11	0.04	276	0.75
Working	0.72	0.69	190	0.65	0.63	266	0.80
> 2 bookshelves at age 10	0.24	0.18	258	0.28	0.21	406	0.71
Panel B. Pre-treatment financial variables							
Self-assessed financial knowledge	3.44	3.33	120	3.22	3.17	150	0.61
Expected correct answers	6.01	5.60	117	5.27	4.92	142	0.85
Interest in finance	6.38	5.60	112	5.80	4.78	139	0.53
Risk willingness	5.11	4.98	110	4.27	3.62	134	0.20
Lottery choice	3.72	4.18	98	3.54	3.87	122	0.68

Notes: The table reports means by gender and survey completion status. Completed refers to the 3,000 men and 3,000 women who finished the survey; Exited to respondents who abandoned it. Each row uses every respondent for whom the characteristic is observed, so the number of exits falls for questions asked later in the survey; *N* exited reports it. The *p*-value tests whether the difference between exits and completers is the same for women and men: it is the *p*-value of the interaction term in an OLS regression of the characteristic on a female indicator, an exit indicator and their interaction, with robust standard errors. Self-assessed financial knowledge ranges from 1 (very low) to 5 (very good); expected correct answers (out of ten), interest in finance and risk willingness range from 0 to 10; the lottery choice ranges from 1 (safest) to 8 (riskiest).

Table A14: Robustness: Exceeding 7 Minutes Time Limit and Sub-Sample Effects

	(1)	(2)	(3)	(4)	(5)	(6)
	> 7 min	> 7 min	IDK (share)	IDK (share)	Correct (share)	Correct (share)
Women	0.007	0.009	0.068***	0.044***	-0.086***	-0.057***
	(0.008)	(0.009)	(0.010)	(0.009)	(0.011)	(0.010)
Without IDK	-0.008	-0.006	-0.120***	-0.115***	0.058***	0.052***
	(0.009)	(0.009)	(0.006)	(0.006)	(0.012)	(0.012)
Incentives	0.009	0.009	-0.053***	-0.047***	0.047***	0.040***
	(0.011)	(0.011)	(0.009)	(0.008)	(0.013)	(0.012)
Information nudge	-0.004	-0.002	-0.061***	-0.062***	0.023*	0.023*
	(0.010)	(0.010)	(0.009)	(0.008)	(0.013)	(0.012)
Women x Without IDK	0.024	0.022	-0.068***	-0.072***	0.020	0.021
	(0.015)	(0.015)	(0.010)	(0.010)	(0.017)	(0.016)
Women x Incentives	-0.001	-0.003	-0.011	-0.018	-0.036**	-0.026
	(0.015)	(0.015)	(0.014)	(0.013)	(0.018)	(0.017)
Women x Information nudge	0.013	0.008	-0.041***	-0.044***	0.029	0.028*
	(0.015)	(0.015)	(0.014)	(0.013)	(0.018)	(0.017)
P-value: T equal 0	0.515	0.605	0.000	0.000	0.000	0.000
P-value: T by female equal 0	0.368	0.462	0.000	0.000	0.008	0.022
Av. outcome for control group men	0.041	0.041	0.120	0.120	0.578	0.578
Sample	All	All	Within limit	Within limit	Within limit	Within limit
Controls	No	Yes	No	Yes	No	Yes
Observations	6000	6000	5714	5714	5714	5714
R-squared	0.002	0.014	0.107	0.243	0.040	0.182

Notes: In columns (1) and (2) the dependent variable is an indicator equal to one if the participant spent more than seven minutes on the financial literacy section. Columns (3)–(6) re-estimate our main specifications on the subsample of respondents who completed the section within the seven-minute limit (it excludes 286 individuals that hit the time limit). The dependent variable is the share of "I don't know" answers across the five financial literacy in columns (3) and (4), and the share of correct answers in columns (5) and (6). Robust standard errors are in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A15: Further Results: Prob. of Exiting, Perceived Difficulty and Completion Time

	(1) Exit	(2) Exit Q N.	(3) Exit	(4) Exit Q N.	(5) Difficulty	(6) Difficulty	(7) Time	(8) Time
Women	0.045***	0.520***	0.006	0.170	0.368***	0.249**	1.754	0.715
	(0.008)	(0.126)	(0.007)	(0.175)	(0.113)	(0.116)	(3.513)	(3.502)
Without IDK			0.002	0.007	-0.183	-0.148	-3.824	-3.884
			(0.008)	(0.197)	(0.137)	(0.132)	(3.431)	(3.436)
Incentives			0.005	0.151	-0.052	-0.029	-1.321	-1.363
			(0.009)	(0.217)	(0.139)	(0.133)	(3.755)	(3.812)
Information nudge			0.006	0.248	-0.162	-0.176	6.144	7.444
			(0.009)	(0.231)	(0.145)	(0.140)	(5.194)	(5.202)
Women x Without IDK			-0.000	-0.085	0.222	0.217	11.591**	11.359**
			(0.012)	(0.287)	(0.196)	(0.191)	(5.201)	(5.155)
Women x Incentives			0.011	0.192	-0.002	-0.010	8.674	8.192
			(0.013)	(0.332)	(0.195)	(0.190)	(8.324)	(8.318)
Women x Information nudge			-0.005	-0.265	0.130	0.195	-0.742	-4.123
			(0.013)	(0.319)	(0.200)	(0.196)	(6.540)	(6.596)
Men control	0.089	1.326	0.029	0.681	4.042	4.042	100.48	100.48
Sample	All	All	Treated	Treated	Treated	Treated	Treated	Treated
Controls	No	No	No	No	No	Yes	No	Yes
Observations	6764	6764	6220	6220	6000	6000	6000	6000
R2	0.000	0.001	0.001	0.001	0.007	0.067	0.002	0.030

Notes: Columns (1) and (3) show the probability abandoning or exiting the survey. Columns (2) and (4) show the number of question in which the survey taker abandons the survey. Columns (5) and (6) show the survey perceived difficulty measured in a scale between 0 and 10. Columns (7) and (8) show the time spent in responding the "Big Five" questions in seconds. If a participant skipped a question, the time recorded in that question is imputed as 0. Men control row refers to the mean value of each outcomes variable for men in the control group, except for Columns (1) and (2) that show this value for men. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A16: Association: IDK & Financial Behavior

	(1) Keep track of money	(2) Actively saving	(3) Not borrow to make ends meet	(4) Seek financial advise	(5) Watch personal financial affairs	(6) Set long-term financial goals	(7) Makes considered purchases	(8) Pays bills on time
Women	0.076***	-0.012	-0.025**	-0.063***	0.050***	0.024*	0.027***	0.013*
	(0.014)	(0.013)	(0.010)	(0.018)	(0.011)	(0.014)	(0.010)	(0.007)
Correct	0.122***	0.263***	0.219***	0.451***	0.129***	-0.008	0.245***	0.121***
	(0.043)	(0.040)	(0.033)	(0.059)	(0.036)	(0.044)	(0.032)	(0.022)
IDK	-0.262***	-0.376***	0.066	-0.428***	-0.097*	-0.295***	0.081*	-0.063*
	(0.056)	(0.056)	(0.045)	(0.058)	(0.050)	(0.058)	(0.045)	(0.034)
Constant	0.456***	0.563***	0.712***	0.187***	0.690***	0.496***	0.670***	0.849***
	(0.036)	(0.034)	(0.028)	(0.048)	(0.031)	(0.036)	(0.028)	(0.020)
Observations	7764	7764	7764	7764	7764	7764	7764	7764

Notes: OLS regression of the each financial savvy behavior indicator on women participants, the share of correct financial knowledge questions (out of seven), and the share of IDK anwers to these same financial knowledge questions. This table reports the individual financial behavior measures that integrate the financial behavior score presented in table 6. Each column represents the following behaviors: (1) keeping track of money, (2) actively saving, (3) not borrowing to make ends meet, (4) seeking advise, (5) watching personal financial affairs, (6) setting long-term financial goals, (7) making considered purchases, and (8) paying bills on time. See OECD (2023) Appendix C for a detailed definition and construction of each measurement. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A17: Association: IDK & Financial Product Awaraness and Holding

	(1) Payment product	(2) Savings, investment, or retirement product	(3) Insurance product	(4) Credit product	(5) Crypto assets
			Panel A. Financial Product Awareness		
Women	0.002	0.006**	0.009***	-0.001	-0.034***
	(0.004)	(0.003)	(0.003)	(0.002)	(0.010)
Correct	0.048***	0.037***	0.048***	0.022***	0.333***
	(0.014)	(0.009)	(0.011)	(0.007)	(0.032)
IDK	-0.124***	-0.044***	-0.068**	-0.016	-0.478***
	(0.031)	(0.016)	(0.027)	(0.011)	(0.049)
Constant	0.956***	0.967***	0.957***	0.983***	0.681***
	(0.012)	(0.008)	(0.010)	(0.006)	(0.028)
			Panel B. Financial Product Holding		
Women	-0.007	0.008**	-0.014	0.025**	-0.036***
	(0.013)	(0.004)	(0.013)	(0.013)	(0.005)
Correct	0.673***	0.057***	0.324***	0.335***	0.088***
	(0.040)	(0.014)	(0.043)	(0.041)	(0.019)
IDK	-0.113**	-0.074**	-0.266***	-0.296***	-0.040**
	(0.049)	(0.029)	(0.051)	(0.055)	(0.017)
Constant	-0.038	0.946***	0.242***	0.484***	0.010
	(0.033)	(0.013)	(0.035)	(0.035)	(0.015)
Observations	7764	7764	7764	7764	7764

Notes: OLS regression of the each product category integrating the financial product awareness (Panel A) and financial product holding (Panel B) indexes on indicator for women participants, the share of correct financial knowledge questions (out of seven) following the OECD measures, and the share of IDK anwers to these same financial knowledge questions. Each category is an indicator equal to one if at least 1 financial product on each category is known or hold for Panel A and B, respectively. In this case, product holding can be shared with other household members. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

B Survey Questions in the Four Different Versions

The text in squared brackets [] is not shown to the survey respondents but we include it here to ease the understanding of the survey design to readers.

[The survey starts here:]

This Bank of Spain survey aims to measure the familiarity of the Spanish population with basic economic and financial concepts. Its duration is approximately 15 minutes. The survey is carried out in accordance with the applicable regulations on the protection of personal data, which guarantees that your data will be processed solely for statistical purposes and for quality control of the survey, guaranteeing their due integrity and confidentiality. We inform you that both your personal identification and contact data provided by 40db, as well as the academic-professional, economic-financial and related to your personal characteristics that you provide us, are processed by the Bank of Spain exclusively for (i) measure for statistical purposes the familiarity of the Spanish population with basic economic and financial concepts; and (ii) supervise and control the quality of the survey. You can withdraw your consent by sending an email to micro@bde.es and exercise your rights regarding the protection of personal data, proving your identity, either in person, by postal mail to C/Alcalá 48, 28014, Madrid (A/A Data Protection Officer) or electronically through the procedure indicated in the Virtual Office of the Bank of Spain, available at: link

For more information, you can consult the Record of Processing Activities available at:link

Q1. Do you agree to participate and the processing of your data for the purposes indicated?

a. Yes

b. No

[First part of the questionnaire: questions include socio-demographic variables, family background variables and variables measuring behavioral traits: Q2-Q29]

Q2. Are you a ...?

a. Man

b. Woman

Q3. How old were you on your last birthday?

Q4. In which country were you born?

a. Spain

b. Other, which one?

Q5. About how many books were in your home when you were 10 years old? (Do not include magazines, newspapers, or textbooks)

a. None or very few (between 0 and 10 books)

b. Enough to fill a shelf (between 11 and 25 books)

c. Enough to fill a bookshelf (between 26 and 100 books)

d. Enough to fill two bookshelves (between 101 and 200 books)

e. Enough to fill more than two bookshelves (more than 200 books)

Q6. Zip Code

Q7. What is your current relationship status?

a. I live with a partner

b. I do not live with a partner

Q8. What is your current legal marital status?

a. Single (never married or domestic partner before)

b. Married or common-law partner

c. Divorced or separated

d. Widower

e. Other, which one?

Q9. Including yourself, how many people live in your household?

a. 1 person

b. 2 people

c. 3 people

d. 4 people

e. 5 people

f. 6 or more people

Q10. Including yourself, how many people receive some type of income?

a. 0

b. 1

c. 2

d. 3

e. 4 or more

Q11. Are you the person who contributes the most income to the household?

a. Yes

b. No

[If Q11 is answered with “a” then jump to Q17]

Q12. What is your relationship with the person who contributes the most income to the household?

a. It is my partner

b. It is my father

c. It is my mother

d. He is my son

e. She is my daughter

f. Other, which one?

Q13. What are the highest level official studies that the main breadwinner of the household has completed? (We understand the main breadwinner or head of the family to be the person from whom the basic income of the family comes).

a. Can’t read or write

b. Without studies or with unfinished primary studies

c. First Grade (School certificate, 1st stage of EGB, more or less up to 10 years old)

d. Second Grade - 1st Cycle (School graduate, or EGB 2nd stage, or 1st and 2nd ESO, up to 14 years old)

e. Second Grade - 2nd Cycle (FP I and II, High School, BUP, ESO 3 and 4, COU, PREU, 1 and 2 Baccalaureate, up to 18 years old)

f. Third Degree - 1st Cycle (Equivalent to Technical Engineer, 3 years, University Schools, Technical Engineers, Technical Architects, Experts, Teaching, ATS, University Diplomas, 3-year degree, Social Graduates, Social Assistants, etc.)

g. Degree, Degree - 2nd Cycle (Universities, Higher Graduates, Faculties, Higher Technical Schools, etc.)

h. Third Degree (Master)

i. Third Degree (PhD)

j. Others

Q14. What is the employment status of the main breadwinner in the household?

a. Currently working

b. Retired/pensioner/disabled

c. Unemployed, have previously worked

d. Unemployed, has not previously worked

e. Student

f. Unpaid household work

Q15. What is the current employment status of the main breadwinner? (If he/she is not currently working, please indicate the status of the last job he/she had)

a. Employee account (eg: employee)

b. Own account (eg: self-employed or entrepreneur)

Q16. What is the profession of the main breadwinner? (If not currently working, please indicate the last job held)

a. Agricultural entrepreneur with 6 or more employees

b. Agricultural entrepreneur with 1 to 5 employees

c. Farm owner with no employees

d. Member of agricultural cooperatives

e. Entrepreneur/Businessperson with 6 or more employees

f. Entrepreneur/Trader with 1 to 5 employees

g. Businessman/Trader without employees

h. Non-agricultural cooperative member

i. Self-employed Professional or Technician (Doctor, Lawyer, etc.)

j. Self-employed manual worker and Craftsman (Bricklayer, Painter, Plumber, Electrician, Upholsterer, etc.)

k. None of the above

l. Director of a Public or Private Company with 25 or more workers

m. Director of a Public or Private Company with less than 25 workers

n. Senior Management/Employee at a higher level of Companies, Public Administration or Army Chiefs (Occupations associated with 2nd and 3rd cycle university degrees)

o. Intermediate Management/Employee at the medium level of Companies, Public Administration or Army Officers (Occupations associated with 1st cycle university degrees, diplomas, etc...)

p. Foreman, Supervisor, Warrant Officer Army

q. Commercial Agent, Representative, etc...

r. Administrative

s. Specialized worker, Civil Guard and Police number

t. Seller, Clerk, etc...

u. Junior Clerk (Janitor, etc.)

v. Unskilled worker, Laborer, Domestic Service

x. Farm laborer

y. Other unqualified personnel

z. None of the above

Q17. What are the highest level official studies that you have completed? (obtaining the corresponding official degree)

[Answers as in Q13]

Q18. In which of the following situations do you currently find yourself?

a. I currently work

b. Retired/pensioner/disabled

c. Unemployed, I have worked before

d. Unemployed, has not previously worked

e. Student

f. Unpaid household work

[If Q18 answered with (“d”, “e”, or “f”) then Q21]

Q19. What is the current labor regime in which you find yourself? (If you are not currently working, please indicate the status of the last job you had)

a. Employee account (eg: employee)

b. Own account (eg: self-employed or entrepreneur)

Q20. What is your profession? (If you are not currently working, please indicate the last job you had)

[Answers as in Q16]

[Do not show if Q8 answered with “a” or Q7 answered with “b”]

Q21. What are the highest level official studies that your partner completed? (obtaining the corresponding official degree)

[Answers as in Q13.]

[Do not show if Q8 answered with “a” or Q7 answered with “b”]

Q22. In which of the following situations is your partner currently?

a. I currently work

b. Retired/pensioner/disabled

c. Unemployed, I have worked before

d. Unemployed, has not previously worked

e. Student

f. Unpaid household work

[Show only if (Q11=b and Q12!=c) or Q11=a)]

Q23. What are the highest level official studies that your mother has completed? (obtaining the corresponding official degree)
[Answers as in Q13]

[Show only if (Q11=b and Q12!=b) or Q11=a]

Q24. What are the highest level official studies that your father has completed? (obtaining the corresponding official degree)
[Answers as in Q13.]

Q25. How would you rate your level of general knowledge on financial topics?
a. Very good
b. good
c. Neutral
d. Bad
e. Very bad

Q26. How many correct answers do you think you could have in 10 questions about basic financial topics? Use a scale of 0 to 10, where 0 means “none correct” and 10 “all correct”

Q27. What is your interest in financial matters? (We refer to the management of personal finances) Use a scale from 0 to 10, where 0 indicates “No interest” and 10 “Maximum interest”

Q28. Are you generally willing to take risks? Use a scale from 0 to 10, where 0 indicates “I am not willing to take any risk” and 10 “I am totally willing to take risk”

Q29. Choose which of the following 8 lotteries you would prefer to participate in. Each lottery has two

possible payouts, each with a 50% probability:

a. Lottery 1 1.1€ with 50% and 1.1€ with 50%

b. Lottery 2 1.0€ with 50% and 1.2€ with 50%

c. Lottery 3 0.7€ with 50% and 1.6€ with 50%

d. Lottery 4 0.6€ with 50% and 1.8€ with 50%

e. Lottery 5 0.5€ with 50% and 1.9€ with 50%

f. Lottery 6 0.3€ with 50% and 2.0€ with 50%

g. Lottery 7 0.1€ with 50% and 2.1€ with 50%

h. Lottery 8 0€ with 50% and 2.2€ with 50%

[Middle part of the questionnaire: Financial Literacy questions, FQ1-FQ10. We will also underline the questions included in the big-five, which will be the main focus of our main results. Define 4 groups. Group 1: *Control*, group 2: *Without IDK*, group 3: *Incentives*, and group 4: *Information*]

The next 10 questions include various exercises. It is okay if you can not answer them all, but it is important that you try to answer each one.

If you do not know the answer, just say so. If you think you have the right answer, it is likely that you do.
[Filter show if Group = 1, 3 or 4]

[If *Incentives* treatment only:]
You will earn an additional 7 cents for each correct answer. If all 10 answers are correct, you can earn 70 more cents, increasing your payment for participating by more than 60%.

[If *Information* treatment only:]
Men typically answer 7 out of 10 financial questions correctly. Women 6 out of 10. This difference is explained mostly (65%) because women choose the answer "I do not know" more often than men. Therefore, we ask you - please - to avoid answering "I do not know".

The section must be completed in a maximum of 7 minutes. Once started, you will not be able to interrupt

it. If you exceed this time, the screen will take you to the next section and you will not be able to go back. When you are ready to start, click “next”.

FQ1: Imagine that 5 brothers receive a gift of 1,000 euros in total. If they share the money equally, how much will each get?

a.

b. I do not know [Filter show if Group = 1, 3 or 4]

FQ2 [Big Five.1: Inflation]: Now imagine that the 5 brothers had to wait a year to get their share of the 1,000 euros, and that inflation for that year was 8%. With that money and within a year they will be able to buy:

a. More than they could buy today with their share of the money

b. The same amount

c. Less than they could buy today

d. I do not know [Filter show if Group = 1, 3 or 4]

FQ3: Suppose you deposit 100 euros in a savings account with a fixed interest of 2% per year. If you do not make any other deposits or withdraw any money, how much money will be in the account at the end of the first year, after interest is paid? (In this account there are no commissions or taxes)

a.

b. I do not know [Filter show if Group = 1, 3 or 4]

FQ4 [Big Five.2: Interest Rates and Compounding]: Again, if you do not make any deposits or withdraw any money, how much money will be in the account after 5 years, after the interest payment is paid? (Remember that the savings account has a fixed interest of 2% per year).

a. More than 110 Euros

b. Exactly 110 Euros

c. Less than 110 Euros

d. It is impossible to say with the information given

e. I do not know [Filter show if Group = 1, 3 or 4]

FQ5 [Big Five.3: Risk Diversification]: Generally, it is possible to reduce the risk of investing in the stock market by buying a wide variety of stocks. True or false?

a. True

b. False

c. I do not know [Filter show if Group = 1, 3 or 4]

FQ6 [Big Five.4: Mortgages]: A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest paid over the life of the loan will be less. True or false?

a. True

b. False

c. I do not know [Filter show if Group = 1, 3 or 4]

FQ7 [Big Five.5: Bond Pricing]: What happens to the price of the bonds if the interest rate increases?

a. Falls

b. Goes up

c. Stays the same

d. The price of the bonds is not related to the interest rate

e. I do not know [Filter show if Group = 1, 3 or 4]

FQ8: In a lottery, the probability of winning a prize is 1%. How many people do you think will win a prize if 1,000 people each buy a single different ticket?

a.

b. I do not know [Filter show if Group = 1, 3 or 4]

FQ9a: If 5 machines take 5 minutes to make 5 objects, how long would it take 100 machines to make 100 objects?

FQ9b: If 5 microwaves take 5 minutes to heat 5 plates, how long would it take 100 microwaves to heat 100 plates?

FQ9c: If 5 microwaves take 5 minutes to heat 5 plates, how long would it take 100 microwaves to heat 100

plates?

a. 15 minutes

b. 10 minutes

c. 100 minutes

d. 200 minutes

e. I do not know [Filter show if Group = 1, 3 or 4]

FQ10: Imagine that you are reviewing your household budget. What is 10 – 2?

a. 3

b. 8

c. 10

d. 20

[Final part of the questionnaire: variables measuring behavioral traits.]

POSTQ1: Thinking about this section with 10 questions, how many do you think you have answered correctly?

POSTQ2: If the Spanish population answered these same 10 questions, on average, how many correct answers do you think the following groups would have? The average grade can take values between 0 and 10 where 0 indicates that none would be correct and 10 indicates that all would answer the 10 questions correctly.

a. The entire population:

b. Women:

c. Men:

d. Young people (between 18 and 30 years old):

POSTQ3: Do you have any of the following products? Click on any of the four possible answers: Yes-No-I do not know-I do not know the product

a. Checking account:

b. Savings accounts or deposits:

c. Credit card:

d. Mortgage:

e. Personal loans:

f. Individual or company pension plans:

g. Mutual funds or stocks:

h. Cryptocurrencies:

i. Life or medical insurance:

POSTQ4: In the last 12 months, have you done banking in any of the following ways? (Check all that apply)

a. By personally visiting a bank branch

b. Using an ATM

c. Calling on the phone

d. Using the computer or tablet

e. Using mobile phone apps

f. Otherwise, which one?

g. None of the above

POSTQ5: Of the following options, mark all that you have done at least once:

a. I applied for a promotion at work

b. I asked for an increase in my payroll/salary/pay

c. I was a class representative at school/institute/university

d. I competed in an individual sport (for example: swimming, tennis, judo, fencing, etc.)

e. I competed in a team sport (for example: soccer, gymnastics, basketball, volleyball, etc.)

f. None of the above

POSTQ6: Tell us to what extent you agree with each of the statements. Use a scale of 1 to 5, where 1 indicates strongly disagree and 5 strongly agree.

a. I can solve most problems if I put in the necessary effort

b. I am confident that I can handle unexpected events efficiently

c. I tend to ask questions in class/work meetings

d. Men tend to handle financial problems better than women

POSTQ7: How complicated did you find the survey? Use a scale from 0 to 10, where 0 indicates no complexity and 10 maximum complexity