Active Learning Improves Financial Education:

Experimental Evidence

Tim Kaiser (University of Koblenz-Landau and DIW Berlin)
Lukas Menkhoff (DIW Berlin)

Discussion Paper No. 131

December 20, 2018

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Dr. Klaus M. Schmidt, University of Munich, 80539 Munich, Germany
+49 (89) 2180 3405 | info@rationality-and-competition.de

Active learning improves financial education: Experimental evidence

Tim Kaiser and Lukas Menkhoff

Abstract

We conduct a randomized field experiment to study the effects of two financial education interventions offered to small-scale retailers in Uganda. The treatments contrast "active learning" with "traditional lecturing" within standardized lesson-plans. We find that active learning has a positive and economically meaningful impact on savings and investment outcomes, in contrast to insignificant impacts of lecturing. These results are not conditional on prior education or financial literacy. Tentative evidence suggests that only active learning stimulates several cognitive and non-cognitive mechanisms; moreover, a social mechanism may be at play as treated individuals join social groups discussing financial matters.

JEL-Classification: I21 (analysis of education), A20 (general), D14 (personal finance), O16 (savings)

Keywords: financial behavior, financial literacy, active learning, lecturing, training method, field experiment

December 2018. We thank the vendors in Western Uganda for participating in our study as well as all the project members at *Mountains of the Moon University*, Fort Portal (Uganda), especially Robert Mawenu and Oliver Schmidt. We greatly appreciate the Agricultural and Rural Finance Program (AGRUFIN) of the *Gesellschaft für internationale Zusammenarbeit* (*GIZ)* in Uganda for generously funding the treatments and data collection. Special thanks go to Dirk Steinwand, Julia Kirya, and Esther Nanjovu. Finally, we appreciate comments from seminar participants in Berlin, Dresden, Fort Portal, Freiburg, Hamburg, Kampala, Kassel, Kiel, St. Gallen, Vienna, Washington D.C., and Zürich, in particular Rob Alessie, Florian Artinger, Toman Barsbai, Christian Biener, Christine Binzel, Martin Brown, Shawn Cole, Stefan Eichler, Nathan Fiala, Thomas Gehrig, Xavier Giné, Antonia Grohmann, Ralph Hertwig, Rajshri Jayaraman, Katharina Lehmann-Uschner, Annamaria Lusardi, Pierre-Carl Michaud, Olivia Mitchell, Anita Mukherjee, Bettina Rockenbach, Helke Seitz, William Skimmyhorn, Carly Urban, Lore Vandewalle, William Walstad, Joachim Winter, and Bilal Zia. Financial support by DFG through CRC TRR 190 is gratefully acknowledged. This field-experiment was pre-registered at the AEA RCT Registry on October 10, 2015 (ID: AEARCTR-0000906);
see: https://www.socialscienceregistry.org/trials/906/history/6694.

Tim Kaiser, University of Koblenz-Landau and German Institute for Economic Research (DIW Berlin), 10108 Berlin, Germany; kaiser@uni-landau.de

Lukas Menkhoff, Humboldt-University of Berlin and German Institute for Economic Research (DIW Berlin), 10108 Berlin, Germany; lmenkhoff@diw.de

Active learning improves financial education: Experimental evidence

1 Introduction

Financial education is a high priority for policymakers around the world. Many of the largest economies, including the vast majority of OECD member countries, have implemented policies enhancing financial education in order to promote financial inclusion and to contribute to financial stability (OECD, 2015). Together, these financial education policies seek to reach more than five billion people in sixty countries. As enormous resources are being invested into financial education policies, the effectiveness of financial education programs is an exigent and grave issue.

This issue of effectiveness is debated in the literature (see Willis, 2011; Hastings et al., 2013; Lusardi and Mitchell, 2014) because early experiments find muted effects on financial outcomes (e.g. Cole et al., 2011). While some evidence shows that financial education not only impacts financial knowledge but also the intended financial behavior (van Roij et al., 2011, 2012; Skimmyhorn, 2016; Kaiser and Menkhoff, 2017), the causal effects of these interventions are economically small on average (cf. Fernandes et al., 2014; Miller et al., 2015). As even large-scale financial education programs may have only modest effects on financial behavior (cf. Bruhn et al., 2014, 2016; Brown et al., 2016; Cole et al., 2016; Berry et al., 2018), policymakers and researchers are evaluating alternatives and complements to the typical and pervasive lecture-based classroom-programs (cf. Drexler et al., 2014; Campos et al., 2017; Carpena et al., 2017). Many of these efforts to increase the effectiveness of financial education involve either tailoring the intervention to narrow target groups (Doi et al., 2014; Drexler et al., 2014) or introducing personalized elements such as counseling (Carpena et al.,

2017). Personalization, however, implies a low student to teacher ratio that is costly and difficult to operate at scale in nationwide financial education programs.

In contrast, we study a cost-effective intervention targeting a general audience that keeps the participant to teacher ratio unchanged, only changing *how* the content is taught. Interactive teaching methods that engage and involve participants are expected to yield greater impacts than the exposition-centered education that currently dominates many training programs. While the unambiguous advantage of this "active learning" approach over "traditional lecturing" is documented in large empirical literatures of other domains, such as science, technology, engineering, and mathematics instruction (e.g., Deslauriers et al., 2011; Ruiz-Primo et al., 2011; Freeman et al., 2014; Wieman, 2014), no study experimentally examines the effects of how *financial* education content is taught. Thus, we study the importance of teaching methods in the field of financial education.

A "virtue" of financial education relative to many other educational domains is that respective domain-specific *field behavior* can be observed. Accordingly, we are – to the best of our knowledge – the first study examining the potential advantage of an active learning intervention with respect to impact on field behavior, and not just regarding increases in knowledge.

We conduct a large-scale cluster-randomized field experiment contrasting an "active learning" program that introduces group problem-solving and interactive elements with a "traditional lecturing" program, while keeping content, teachers, and intensity of training constant. The randomized field-experiment is conducted in rural Western Uganda with 1,291 self-employed.

Three results emerge: First, and our main finding, is that the group randomly allocated to the "active learning" program experiences a significant improvement in financial behavior whereas there are no effects on the group allocated to "traditional lecturing." Specifically, the

active learning program causes an increase in three out of five desired outcome-domains, each captured by an index. The strongest impacts occur in the savings and investment domains: total savings increase by 21 percent relative to the control group, with active learning having a direct effect on investments into the own business and business formalization. In addition, there are weaker, but still marginally significant, effects on an index of debt-related behavior. Thus, this approach generates larger treatment-effects than the insignificant, positive treatment effects realized by "traditional lecturing." Moreover, a differential treatment effect can be confirmed for outcomes in the savings domain. Thus, our study is the first in the entire literature on active learning to document causal effects of active learning on field behavior (not just performance on knowledge tests).

Second, regarding potential mechanisms of "active learning" treatments, we provide tentative evidence that the strong impact of this approach may be supported by stimulating an increase in financial literacy, self-control and financial confidence. An additional social mechanism may explain the observed treatment effects to an even greater extent, i.e. active learning leads individuals to join social groups in order to discuss their personal financial matters. An analysis shows that the main improved outcomes of the active learning treatment are mediated via group membership, but that the observed effects are not driven by an increase in group-based forms of saving (e.g. group accounts) but changes of financial behavior in individual accounts.

Third, we show in heterogeneity analyses that the benefits of "active learning" are not conditional on low levels of prior education or financial literacy. This is in contrast to other studies where training is mainly effective for relatively low-skilled individuals (Bjorvatn and Tungodden, 2010; Drexler et al., 2014; Fort et al., 2016).

Our research adds to the growing literature that advances the understanding of the differing impact of delivery channels of financial education and seeks to contribute to the

debate on how to design effective and scalable training programs. Drexler et al. (2014), varying the content of the financial education curriculum, show the differential impact of two financial education curricula in the Dominican Republic. They provide evidence that a heuristics-based approach, relying on the simplification of complex financial concepts ("rule-of-thumb-training"), generates larger behavioral impacts than the teaching of traditional curricula ("full technical training"), mainly for low-skilled individuals and individuals with low baseline financial literacy and motivation. Skimmyhorn et al. (2016) rerun this type of experiment, but do not find evidence to support differential effects regarding the two types of curricula, probably because their sample of respondents has high levels of *ex-ante* human capital, confirming the result by Drexler et al. (2014) that the benefits of the "rule-of-thumb" approach may be driven by strong effects on low-skilled individuals. Lusardi et al. (2017a) present evidence from online-experiments conducted in the U.S. that interactive tools, narratives, and financial education videos may be more effective than written informational content in affecting financial knowledge and self-efficacy.

Moreover, Berg and Zia (2017) show that financial education interventions that primarily target non-cognitive channels through mass media can significantly impact financial behavior. Campos et al. (2017) demonstrate that a psychology-based training program for entrepreneurs, teaching a pro-active mindset, outperforms traditional business training for self-employed individuals in West Africa with strong differential effects on business profits. Finally, Carpena et al. (2017) show that complementing classroom instruction with personalized elements, like individualized counselling and goal-setting, yields higher treatment effects on financial behaviors, such as keeping track of household expenses through a written budget, starting (informal) savings, and avoiding expensive borrowing.

Much of this evidence suggests that interventions relying on traditional lecturing within classroom-settings appear to have small or zero effects. Thus, alternative approaches to

lecture-based education seem to be especially important when targeting low-skilled and low-income individuals. Some argue that "one-size-fits-all" (Drexler et al., 2014, p.25; Carpena et al., 2017, p.2) classroom programs are not suitable for improving financial behavior and programs must be complemented with individualized elements. In contrast, we show that even within a one-size-fits-all classroom program, active learning techniques can provide the missing link between knowledge creation and behavioral change. Encouragingly, results are not conditional on low prior human capital.

This paper is structured into seven further sections: Section 2 describes the financial education programs, the experimental design, and empirical strategy. Section 3 provides an overview over response rates, descriptive statistics, and a discussion of randomization-balance. Section 4 presents main results. Section 5 provides a discussion of potential causal mechanisms leading to behavioral change. Section 6 presents an investigation of distributional treatment effects and treatment effect heterogeneity. Section 7 summarizes robustness exercises and Section 8 discusses the results and concludes.

2 Treatments and experimental design

In this section we describe the two financial education programs (Section 2.1), the experimental design (Section 2.2), and the empirical strategy employed (Section 2.3).

2.1 Financial education interventions

As a step toward promoting financial inclusion, the Bank of Uganda (*BoU*), the country's central bank, established a national strategy for financial literacy in Uganda.[1] This strategy seeks to foster the personal financial knowledge and behavior of different target groups, including school students, young adults in urban settings, and adults in rural areas.

[1] See: https://www.bou.or.ug/opencms/bou/bou-downloads/Financial_Inclusion/Strategy-for-Financial-Literacy-in-Uganda_August-2013.pdf; last accessed February 9, 2018.

The *BoU* partnered with the German Development Cooperation (*GIZ)* to design effective financial education interventions. While the specific curricula and dissemination formats vary depending on the target group, they all target financial knowledge and behavior within the five sub-domains of (i) budgeting and personal financial management, (ii) credit, (iii) savings, (iv) business investing, and (v) payments and financial service providers.

Focusing on one of the financial education strategies' major objective of improving rural outreach, *GIZ* developed educational formats for the target population of the rural self-employed. These programs teach how to create a written household budget and how to keep track of household's financial inflows and outflows; they also encourage household savings, explain the costs and benefits of various forms of credit, explain the trade-off between risk and return regarding productive investments into micro-enterprises, highlight the benefits of diversification among sources of income and investments, and inform about the costs and benefits of using financial services provided by regulated financial institutions. The trainings emphasize the importance of direct and opportunity costs associated with each type of financial product which sometimes also means considering how financial resources may be protected against the demands of peers (see also Schaner, 2017). Thus, this training promotes the use of formal financial services, without generally discouraging semi-formal types of financial products (e.g. saving in village savings and loans associations (VSLAs) or rotating savings and credit associations (ROSCAs). The trainings do, however, caution against both the use of expensive credit provided by informal money-lenders and the take up of costly loans to finance consumption expenditures in general.

Currently, among other interventions, *GIZ* offers two different educational interventions for the same target group of rural self-employed, creating the opportunity to study the differential impact of alternative delivery approaches to financial education. While these two programs are identical with regards to the content areas covered (they both address the

domains (i) to (v)) and their intensity (two hours), they differ in their concrete method of instruction, i.e. *how* this content is taught (cf. Table A1 in Appendix A).

Version A of the financial education training uses *active learning* methods and is highly learner-centered (see Ruiz-Primo et al., 2011; Freeman et al., 2014). Its main feature are five distinct stations, designed to provide problem-based learning opportunities and encourage discussion among the participants. Exercises and materials are designed to engage participants with the subject matter, e.g. through completing budgeting exercises, sorting different savings- and investment options with regard to their safety risk/return-profiles, differentiating between sensible and non-sensible reasons to take up a loan, and completing an exercise on whether financial services are regulated by the central bank or not. Respondents are encouraged to share their experiences and complete the exercises. The trainer has the role of a learning facilitator.

Version B, on the other hand, is organized as a community lecture, i.e. an "exposition-centered" (Freeman et al., 2014, p.8410) *traditional lecturing* approach, relying mostly on lecturing with the aid of a flip-chart and some room for participants to ask questions. Here, the lecturer explains important concepts and demonstrates how to create a written budget, and how different financial products and choices could be categorized. The participants are expected to listen to the input of the trainer and are allowed to ask questions. Sharing of experiences is also encouraged, but the trainer can be characterized as a lecturer rather than a facilitator. Thus, the experiment is designed to contrast constructivist versus exposition-centered (transmissive) course designs as two stylized ideal-types along the continuum of instructional approaches. Despite these important differences in instructional methods, both versions of the financial education training are standardized to cover the same content and to have the same intensity of about two hours. For the purpose of evaluating the impact of the financial education programs, we employ the same group of newly recruited teachers to

deliver both financial education versions A and B to the target groups. Thus, the different versions are not confounded with idiosyncratic characteristics of the teachers and implementation is not heterogeneous across experimental sites. The teachers are all enrolled in a program on banking and microfinance at a local university and have participated in a training of trainers offered by *BoU* and *GIZ*.

2.2 Experimental design

We organize a cluster-randomized experiment to study the differential impact of the two financial education interventions on financial behavior. The main outcomes of interest are changes in *financial behavior* within five sub-domains addressed by the financial education interventions, i.e. (i) budgeting, (ii) usage and handling of credit, (iii) savings, (iv) business investing and formalization, as well as (v) the use of formal payment- and other financial services such as formal insurance. Randomization is done at the market-level, across 83 rural marketplaces in the seven districts of the Rwenzori Region in Western Uganda, collectively forming the sample of clusters considered in this study (see Figure 1).

<Figure 1 about here>

To the best of our knowledge, the dataset covers all relevant permanent and regular marketplaces in the region. Because prior information about the marketplaces (such as number of vendors and primary goods traded) is limited at the time of randomization (prior to the individual-level baseline survey), we perform a non-stratified randomization procedure to allocate the 83 clusters in our sample to either receive financial education treatment A (n=27) or B (n=28), or to be allocated to the control group (C) (n=28). The trainings each aim for reaching a target group of about 15 to 16 vendors, which leads to a sample population of 1,245 to 1,328 individuals.

Vendors were mobilized to participate in our survey and the financial education sessions by the local market-chairpersons. The treatment status was unknown to the market-chairperson, so no differential selection (mobilization) should be in place due to the reliance on local market-chairpersons. However, selection biases could arise if market-chairpersons favor specific groups (such as their peers) over others in general. This would indeed impact the *external* validity of our experiment in the sense that treatment effects may causally interact with unobserved features of our specific context and sample. The *internal* validity, however, is not affected because these selection mechanisms will occur in all three groups. After baseline survey, the treatment groups received either financial education treatment A (active learning) or B (traditional lecturing) on the same day, immediately after the interviews. The control group (C) did not receive any treatment. After baseline-survey and subsequent financial education treatments, we conducted follow-up surveys approximately 6 months later.

We report power-calculations for minimum detectable effect sizes in order to be able to rule out imprecisely estimated small- or zero-effects of financial education treatment. Intra-cluster-correlation (ICC) ranges from 0.030 (financial literacy score and budgeting index) to < 0.001 (savings index) for the outcomes. Thus, our experiment has 80 percent power to precisely detect (at $\alpha = 0.05$) effect sizes as small as 0.15 to 0.18 standard deviation units. To put these minimal detectable effect sizes into perspective with results from the most recent meta-analysis of the literature: The average effect size on financial behaviors in a sample of 90 studies is about 0.9 SD units (Kaiser and Menkhoff, 2017). Note that type II error rates are higher than 20 percent for coefficients estimated to be smaller than 0.15 (0.18); i.e. we cannot reject positive impacts of financial education treatments smaller than 0.15 standard deviation units. However, these may be regarded as economically insignificant in our setting (see Section 4).

2.3 Empirical strategy

To estimate the causal effect of the financial education interventions on financial literacy and behavior, we compare the two treatment groups A (active learning) and B (traditional lecturing) with the control group C (no treatment) at the time of the follow-up survey. As selection into treatment is randomized, with balanced groups at baseline (see Section 3.2 and Appendix A), the unbiased intent-to-treat (ITT) treatment effect (average effect of being assigned to a treatment-cluster) can be estimated within the following ANCOVA framework:

$$y_{ic(t)} = \alpha + \delta_1 y_{ic(t-1)} + \beta_1 A_c^T + \beta_2 B_c^T + \theta_d + \varepsilon_{ict} \quad (1)$$

Here, $y_{ic(t)}$ denotes the outcome variable (measure of financial behavior) for individual i in cluster c at the time of follow-up (t). $y_{ic(t-1)}$ controls for the lagged value of the outcome at baseline and θ_d are district-level fixed-effects. A_{ic}^T and B_{ic}^T are dummy variables equal to one for respondents situated in a cluster assigned to treatment A or B, respectively. Thus, β_1 captures the (ITT) treatment effect of financial education intervention A, and β_2 captures the (ITT) treatment effect of intervention B. ε_{ict} denotes the error-term. Standard errors are clustered at the market-level to account for the level of randomization.

For binary outcomes, linear probability models are used (see Karlan and Valdivia, 2011; Cole et al., 2013, and Drexler et al., 2014, for this approach). However, results are not sensitive to changing the estimations to non-linear (logit or probit) models for binary outcomes.

To avoid problems inherent to testing multiple hypotheses (type-I-error inflation), we aggregate multiple related outcomes into index-measures of outcome families: Following Kling et al. (2007), Karlan and Valdivia (2011), and Drexler et al. (2014), we define $y^*$to be an equally weighted average z-score index of its components y_k^*. Thus, for each component of

a given outcome family, we first rescale each outcome such that positive values indicate desirable treatment effects. Next, we standardize the component to have a mean of zero and standard deviation of one for the control-group: $y_k^* = \frac{y_k - \mu_k}{\sigma_k}$, with μ_k denoting the mean of y_k for the control group (C) and σ_k denoting the standard deviation of y_k for the control group. The aggregate index then takes the following form: $y^* = \frac{\sum_k y_k^*}{k}$. Finally, we standardize the outcome index (y^*) to have a mean of zero and standard deviation of one for the control-group. Thus, coefficients on A_{ic}^T and B_{ic}^T can be interpreted as standardized mean differences (Glass's Δ).

Finally, to probe causal mechanisms we employ two different types of analyses: First, we follow Sayinzoga et al. (2016) and Bulte et al. (2017) and use an IV-setup to study potential causal mechanisms through intermediating variables (see Section 5). Thus, in the first stage of this set-up, an intermediate outcome (such as an increase in financial literacy) is explained by the treatment dummies. In the second stage, the intermediate outcome is instrumented by the treatment dummies. Note that, although useful to explore the potential causal mechanisms at play, the second stage is not necessarily expected to provide an unbiased estimate of the treatment effect as the exclusion restriction is likely to be violated. Thus, we also employ a causal mediation analysis (see Steinert et al., 2018, for such an approach in the financial education literature) which allows to quantify the sensitivity of causal mediation effects to violations of the identifying assumptions (Imai et al., 2011).

3 Data

After mapping of the markets, piloting the survey tools and interventions, and randomization, we conducted a comprehensive baseline survey between November 1st and December 19th of 2015. This dataset covers all vendors invited and interested to participate in

our survey (n=1,291). The questionnaires were translated into three local languages widely spoken in the area and the enumerators conducting the face-to-face interviews in the local languages were trained extensively prior to the field-activities.

3.1 Response rates

Following the baseline-survey and the subsequent financial education treatments, we conducted follow-up surveys between April 6th and July 19th of 2016. After this first round of tracking efforts, we had followed-up with 1,094 vendors (i.e. the attrition rate was at 15.26 percent). To minimize attrition, we undertook extensive tracking efforts to follow up with another 67 respondents (see timeline Figure A1 in Appendix A). Thus, our final response rate is high, given the kind of relatively mobile target group: We follow up with roughly 90 percent of the initial sample at endline survey (see Table 1).

<Table 1 about here>

Unfortunately, attrition rates vary by experimental condition: While the control group (C) and treatment group (A) have attrition rates of 8.55 percent and 7.25 percent, respectively, the attrition rate in treatment group (B) is almost twice as high as in group (A) at 14.25 percent. While this may indicate selective attrition, we show in comprehensive robustness exercises that this does not affect our results: we probe the sensitivity of our results by estimating bounds on the treatment effects with several scenarios imputing missing observations at the endline and applying inverse probability weighting of selection into endline survey to our regressions. Details are provided in Section 7 and Appendix B.

3.2 Baseline descriptive statistics

Table 2 reports summary statistics for the full sample and each experimental condition at baseline (see Table A2 in Appendix A for descriptive statistics for additional baseline-covariates).

<Table 2 about here>

Panel A shows variables that measure characteristics at the household level. The average household size is 6.83 people, with an average of 2.17 adults contributing to the household's income, a mean of 4.17 children being supported and a mean of 0.36 adults who do not generate external income, such as elderly (plus 0.13 for missing values). Several currency denominated outcomes had a long right tail, possibly indicating enumeration errors. Therefore, we winsorize all currency denominated outcomes at the 99th percentile (see e.g. Blattman et al., 2014, for a study in Uganda). The mean (winsorized) monthly household consumption value is about 593,000 UGX ($519 in 2017 PPP).

Panel B reports variables at the respondent-level. The mean (winsorized) monthly individual income is around 220,000 UGX ($192 in 2017 PPP). Household consumption is higher than hypothetical added individual incomes because of subsistence farming, as reported by 83 percent of the sample (see Table A2). Our sample is predominantly comprised of women (80 percent) and the average age is 36.2 years. Only 25 percent report to have participated in education beyond primary school.

Our survey also includes a measure of financial literacy (see Appendix C) and psychological variables, such as self-control and financial confidence, which are standardized to have a mean of zero and a standard-deviation equal to one for the control group (see Appendix A). These are used for the investigation of possible causal mechanisms (see Section 5) and for the purpose of probing randomization balance at baseline.

Panel C shows descriptive statistics for outcome measures of financial behavior indices (standardized to have a mean of zero and a standard deviation equal to one) at baseline (see descriptive statistics for individual index components in Table A3 in Appendix A).

3.3 Randomization balance

Causal inference within the estimation framework introduced in Section 2.4 rests on the random assignment of clusters to the treatment conditions, which achieves balanced observed and unobservable characteristics. Randomization balance is probed by comparing the means between the control group and the treatment groups, as reported in columns (4) and (6) of Table 2. These differences are estimated within a simple regression framework, where standard errors are clustered at the market-level. Due to randomization, only a small difference exists: The treatment groups are estimated to be slightly younger, on average, than the control group. However, this difference is only significant at the 10-percent level (see Appendix A for further baseline covariates which support this picture). This minor imbalance is what can be expected to occur by chance. *Panel C* shows descriptive statistics for these outcome indices of financial behavior at baseline. Again, no statistically significant differences exist between the three experimental groups (see Table A3 in Appendix A for descriptive statistics and randomization balance for single index components). Reassuringly, a joint test of orthogonality (where a categorical variable indicating the experimental group is regressed on all baseline covariates and the index-measures for outcomes at baseline) results in low explanatory power and a p-value of 0.71. Thus, orthogonality and balance are reassured using this test on a reduced sample of 934 respondents due to missing values for some of the covariates.

4 Results

This section reports on the main treatment effects of the two financial education interventions (Section 4.1) and detailed analyses of treatment effects on selected index components (Section 4.2).

4.1 Main treatment effects

Table 3 reports the average (intent-to-treat) treatment effects of the financial education interventions A and B on five domains of financial behavior plus a summary effect aggregating the information across these five domains into a single index. All coefficients are estimated within an ANCOVA framework and include district-fixed-effects to account for district-level unobservable characteristics (see Appendix B for a discussion of district-level events that suggest including these district dummies).

<Table 3 about here>

Indices of financial behavior. The five domains of financial behavior addressed by these trainings are each measured by an index and are presented in Table 3 in the order of the content areas addressed in the financial education sessions. The indices aggregate the single items as they are presented fully in Appendix A and selectively discussed in Section 4.2. Overall, coefficients are – with the exception of one point-estimate – consistently positive, indicating that financial behavior can be impacted in the intended direction.

Among these domains of financial behavior, the overall effects are strongest for the *investment index* with effect sizes being relatively large (more than a quarter of a standard deviation) and statistically significant at the 1-percent level (column 4). Only treatment A results in a statistically significant change in behavior, here investment behavior. Treatment B, i.e. lecturing, is estimated to have a statistically insignificant effect size of about 0.11 standard deviation units.

Similarly, the effect on the *savings behavior index* (column 2) is sizeable for treatment A (about 0.17 standard deviation units) and statistically significant at the 5-percent level. Again, treatment A has a positive effect, while treatment B is statistically and economically insignificant (effect size of 0.01). The resulting large difference between both trainings generates a statistically significant difference of treatment A over treatment B.

Surprisingly, relative to low expectations from the literature on the effectiveness of financial education to change *debt behavior* (Fernandes et al., 2014; Miller et al., 2015; Kaiser and Menkhoff, 2017), we find a small but marginally significant effect for treatment A (column 3). As the coefficient sign of treatment B is even negative, the difference between both treatments is marginally significant, as well. However, we do not emphasize this result since it appears to be driven by a single variable as we show in the appendix. Regarding training effects on the *budgeting index* (column 1), the effect size estimates are positive but statistically insignificant. The coefficient levels are higher for the *financial services index* (column 5); this is the only case where the coefficient of treatment B is higher than the one of treatment A, but both coefficients are estimated with a large standard error and remain statistically insignificant from zero and from each other.

Finally, we aggregate information on all outcome domains into a *summary index* (column 6) which is defined as an equally weighted z-score index of the five financial behavior indices, informing about the overall impact of both treatments. The effect size of lecturing (treatment B) is about 0.03 standard deviation units, and thus of plausible size given that the training lasts only two hours relative to the average treatment effect of 0.08 standard deviation units at an average intensity of 1 to 2 days reported in a meta-analysis of forty RCTs (Kaiser and Menkhoff, 2017). The effect size of active learning (treatment A) is about 0.21. Thus, the overall treatment effect of active learning is significantly different from zero ($p<0.05$) and from lecturing ($p<0.1$).

Overall, we see that financial education tends to have desired effects, but effect sizes remain small and statistically insignificant unless the active learning program (treatment A) is implemented. Active learning results in a significant effect on three out of five financial behaviors addressed by the training. Lecturing (treatment B) on the other hand has no statistically significant effect on any of the addressed financial behaviors. However, a meaningful differential treatment effect between active learning and lecturing can only be confirmed for the savings domain.

4.2 Effects on single index components

This section complements the summary results on indices from above by providing and discussing results about the single index components of interest and considering their economic significance in terms of financial outcomes. In this respect, we focus on those two indices of financial behavior where we found statistically significant results at the 5-percent level, i.e. savings and investments outcomes, while the other program objectives (improving budgeting behavior, debt behavior, and financial services behavior) are discussed in more detail in Appendix A.

Components of the savings index. As shown above, active learning (treatment A) positively affects the savings index, which is both statistically different from zero and from traditional lecturing (treatment B). Now we look at the three detailed results on financial outcomes as they are aggregated in the index, i.e. the variables "any savings," "total savings," and "net savings." These results are shown in the first three columns of Table 4.

<Table 4 about here>

Given that 87.8 percent of the control group report any savings, the increase due to financial education is expected to be modest by definition and is, in effect, just 3.8 percentage points for treatment A, which is statistically significant at the 10-percent level (column 1).

The effect of treatment B is smaller in size at 2.7 percentage points and, thus, statistically insignificant. While this effect size may be economically modest, it is noteworthy that financial education has an effect on savings at the extensive margin that is similar to other studies in various contexts (e.g. Duflo and Saez, 2003; Drexler et al., 2014; Jamison et al., 2014). Thus, one may speculate whether treatment effects on the extensive margin may be larger in magnitude for samples with a lower *ex-ante* share of savers. The main index components driving the overall positive treatment effect of the index, however, are strong increases in financial outcomes in the form of total savings and net savings at the intensive margin.

Respondents in treatment group A report an average increase of 109,186 UGX ($95 in 2017 PPP) in total savings, which amounts to a treatment effect of about 0.18 standard deviation units, or an increase in savings by 21 percent over the mean of the control group (column 2). In contrast, the effect of treatment B is estimated to be economically small (with a negative sign) and is estimated to be statistically insignificant. Testing for differential impacts, active learning is more successful than traditional lecturing (p=0.052). The effect on net savings appears to be even stronger. This variable captures the moderate reduction in debt volume together with the strong positive treatment effect on total savings so that net savings increase by 145,480 UGX ($127 in 2017 PPP), i.e. 38 percent, relative to the control group (column 3). Thus, this effect appears quite strong and statistically significant at the 5-percent level, resulting in the equality of effects for treatments A and B being rejected (p=0.025). Note that these effect sizes on savings outcomes are similar to other recent interventions, but which operate with much higher intensity of instruction (e.g. Carpena et al., 2017; Calderone et al., 2018).

Components of the investment index. Turning to total investments into the own business, effect sizes are estimated to be 90,173 UGX ($79 in 2017 PPP) for treatment A and

41,801 UGX ($37 in 2017 PPP) for treatment B (column 4 in Table 4). Despite relatively large standard errors, the effect of treatment A is economically significant: The increase in investment by 90,173 UGX is equivalent to an increase of 30 percent relative to the control group. This corresponds to an effect size of 0.17 standard deviation units. The effect of treatment B, again, is estimated to be less than half the size and statistically insignificant.

The second component of the investment index looks at business formalization. The survey data capture whether respondents report having formally registered the business with any authorities. Again, treatment A results in a statistically and economically significant effect, whereas treatment B yields effect sizes insignificant from zero: Given that only 23.2 percent of the control group state to have formally registered their business with authorities, an increase of 7.7 percentage points is sizeable (an increase in formalization by 33 percent) and significantly different from zero. In contrast, the effect of treatment B, i.e. lecturing, is estimated to be insignificant from zero.

The third component of the investment index captures the number of productive durable goods owned by the respondents. These items are defined here as those that can be used for productive purposes (e.g. mobile phones, motorcycles, generators) or livestock. The mean number of productive durable assets in the control group is 51.38. Treatment A results in an increase of these assets by 5.41 items ($p<0.001$), i.e. 10.5 percent. In contrast the estimated increase in investment due to treatment B is, again, less than half the size (2.22 items or 4.32 percent) and insignificant. Thus, the two treatments are significantly different at the 10-percent level.

Outcomes on further items within indices. Even though the aggregate impact on the indices regarding the remaining outcome families is only marginally significant and insignificant, respectively (cf. Appendix A for a complete discussion of these results), we note results on single components here, since these qualitatively complement the general

picture, that effect sizes of active learning are consistently estimated to be larger than effects of traditional lecturing.

The *budgeting index* consists of five binary items capturing changes in budgeting and record keeping behavior. None of these items is significantly impacted to a positive extent. Thus, the aggregate impact is also statistically insignificant from zero on average (see Table A4 in Appendix A).

The *borrowing index* consists of six items (see Table A5 in Appendix A). Items aggregated into this index capture the structure and volume of debt along with a proxy for debt-bearing capacity and binary items indicating whether respondents would take up loans without a plan or are able to distinguish between sensible and non-sensible reasons to take up a loan with interest payments. The coefficients on the volume of loans intended for consumption purposes and productive investments are extremely small and, thus, insignificant from zero. There is a significant effect for the active learning treatment, however, on the self-report of respondents to take-up a loan (if offered) when they had no clear plan of how to use the money: approximately 14 percent of the control group report to be willing to take up a loan, even if they had no clear plan on how to use the borrowed amount. The marginal effect of the active learning treatment results in a 5.5 percentage point decrease in the willingness to take up a loan without a clear plan of its utilization. The effect of the traditional lecturing treatment, again, is insignificant from zero.

Finally, the *financial service index* includes two binary items indicating whether respondents were ever covered by a formal insurance product or if they have ever used mobile money (payment) services (cf. Jack and Suri, 2014) (see Table A6 in Appendix A). About 44 percent of the control group report having used mobile money services at least once and active learning has a marginal effect of 6.3 percentage points on mobile money use at the extensive margin. Lecturing, however, has a much smaller effect size (3.5 percentage points)

and is statistically insignificant from zero. Both treatments have zero-effects on formal insurance use.

5 Exploring causal pathways

The remarkable difference we uncover between the two versions of financial education intervention raises the question of potential causal mechanisms. Note, that explorations of potential causal pathways were not pre-specified at the time of registration of the experiment. Thus, any analyses should be treated as exploratory. We motivate our procedure on psychological (Section 5.1) and social mechanisms (Section 5.2) and complement our results with mediation analyses (Section 5.2 and Appendix A).

5.1 Possible cognitive and non-cognitive psychological mechanisms

The early literature on the evaluation of financial education often focuses on the general effect size, because the existence of such a positive effect of financial education was heavily debated in the literature (cf. Fernandes et al., 2014). While many RCTs clearly demonstrate that there is a positive effect on both financial literacy and downstream financial behavior, average effects are generally quite small in size and highly heterogeneous across studies. This raises interest in potential determinants of effective interventions: Meta-analyses find that education intensity matters for its impact and that the timing, participation conditions, and features of the target group can contribute to understanding the reported impact heterogeneity (cf. Fernandes et al., 2014; Miller et al., 2015; Kaiser and Menkhoff, 2017). However, little is known about *how and why* differently designed programs lead to heterogeneous impacts on financial behaviors. An important part of this discussion is to better understand the potential causal mechanisms by which financial education impacts behavior.

Regarding the psychological and economic content of such mechanisms, the first candidate is, of course, an improvement in financial literacy that enables individuals to make better financial decisions, i.e. financial education would impact financial behavior through a cognitive channel. Evidence on this possible causal pathway is documented (cf. Fort et al., 2016; Sayinzoga et al., 2016) and also appears to be supported by a larger sample of experimental work (see Kaiser and Menkhoff 2017, p.617). However, it is a robust insight of (financial) education research that a good transfer of knowledge into behavior is fostered by additional elements (cf. Carpena et al., 2017). Generally, in this respect, evidence shows that better self-control, and in line with it also future-oriented time-preferences, i.e. non-cognitive variables, seem to be associated with more savings (e.g. Ashraf et al., 2006). With regard to the literature on the causal effects of financial education, Alan and Ertac (2018) and Lührmann et al. (2018) demonstrate that financial education can increase the quality of intertemporal decision-making. Similarly, Berg and Zia (2017) show that an intervention implemented into mainstream-media harnessing emotional connections also impacts financial behavior without necessarily impacting cognitive components of general financial literacy. Finally, financial confidence and attitudes may play an important role in financial behavior (e.g. Berg and Zia, 2017; Carpena et al., 2017; Carpena and Zia, 2018). We explore these potential psychological mechanisms, i.e. increased financial literacy, self-control, and financial confidence.

In this context of mechanisms, we would like to note that we do not observe significant increases of income or consumption over the six months observation period (see Table A7 in Appendix A). Thus, increased savings are not an indirect consequence of higher income but due to a change in behavior.

In the following we describe the formation of the three aforementioned variables. "Financial literacy" is assessed through five standard questions that are aggregated into a

psychometrically valid scale of financial literacy (see Appendix C). "Self-control" is assessed by a survey item asking respondents to reply to the question: "If you get money, do you tend to spend it too quickly?" on a 1 (often) to 4 (never) rating scale. Responses are transformed into a z-score, scaled by the mean and standard deviation for the control group (see descriptive statistics in Table 2 and in Table A8 in Appendix A). "Financial confidence" is assessed by multiple items that are aggregated into an unweighted z-score-index of its components as detailed in Section 2.3. The index covers responses to binary questions and statements that are answered on a rating-scale. Questions were asked on whether or not respondents felt that a complaint to a financial services provider would not change anything, whether respondents feel confident to inquire about the details of a financial product and to choose the financial product that best meets their needs, and whether respondents consider various products and options before making a financial decision (see definitions and descriptive statistics in Table A8 in Appendix A).

Results are presented in Table 5. Active learning (A) appears to have a significant impact on all three considered mediators, while lecturing (B) only impacts the measure of financial confidence but neither the measure of financial knowledge nor the survey-measure of self-control. Thus, given that only treatment A has an effect on financial behavior and financial outcomes, we can conclude that changes in financial confidence alone (as is the case for treatment B) may not be sufficient to realize a change in behavior.

<Table 5 about here>

Next, we examine which content areas have been better understood by those individuals trained with the active learning program. We find that the (marginally) significant impact on the financial literacy score is primarily caused by a better understanding of two items dealing with simple and compound interest (see Table A9 and Table C1 in the Appendix). Thus, one may speculate whether a better understanding of interest compounding has contributed to an

increased willingness to save and invest available funds. Second, we observe that treatment A also causes an increase on a (self-reported) measure of self-control which is in line with recent experiments on financial education and intertemporal decision making (e.g. Alan and Ertac, 2018; Lührmann et al., 2018). Thus, the increase in savings may also be caused by this non-cognitive mechanism.

To probe this intuition, we complement our analysis with IV regressions (see Sayinzoga et al., 2016; Bulte et al., 2017) and a mediation analysis (Imai et al., 2011). Looking at the second-stage regressions of the IV-approach (Table A10 in Appendix A), we see that improved savings behavior may indeed be impacted via better self-control, while the investment index may be impacted via improved financial literacy and financial confidence. As noted earlier, however, this estimation may suffer from the violation of the exclusion restriction. Thus, we turn to results from causal mediation analysis to probe the sensitivity of these results (see Table A11 in Appendix A). We find that the mediation effect of our considered mediators is estimated to be less certain (and possibly insignificant from zero) and may account for less than 5 percent of the total causal effect of treatment A. Overall, active learning appears to have a direct effect which is not well explained by our considered mediators. Thus, we also turn to other possible mechanisms that may explain the impact of active learning.

5.2 Possible social mechanisms

Active learning methods hinge critically on involving participants as a group and fostering interaction between them. Individuals jointly engage in interactive problem solving and information sharing. The lecture, in contrast, does not depend on group interaction, since communication is mainly between the lecturer and the individual participants. Thus, we explore whether the active learning treatment may lead people to join a community of peers in

order to discuss their financial affairs. The intuition is that this group experience may motivate individuals who have not been involved in social groups at baseline, to join groups following the treatment. Column (1) of Table 6 shows results on any group membership, i.e. not only including organized groups of people saving together. We find that active learning increases the fraction of individuals involved in any social groups at endline by 4.7 percentage points ($p<0.1$), relative to the control group mean of 80.5 percent of individuals being involved in any social groups. The lecture treatment, on the other hand, has no effect on joining a social group (-1.3 percentage points). Thus, this difference between the two treatments is statistically significant at the 5-percent level.

<Table 6 about here>

Next, we analyze whether group membership mediates the effect of the active learning treatment. Indeed, we find that group membership is a significant mediator of active learning for the outcomes saving (column 2), investing (column 3) and the aggregate summary index of financial behaviors (column 4). Specifically, we find that the increase in group membership caused by active learning may account for up to 54.62 percent of the increase in the savings index (upper bound of the estimated 90-percent CI) up to 25.14 percent (upper bound of the estimated 90% CI) of the overall impact. The impact on increased investments, however, appears to be explained to a lesser extent by this mediator (upper bound of 9.43 percent). Thus, in addition to this mediation effect there is still an unexplained direct effect of the active learning treatment for which we do not observe the potential mediator. It is noteworthy, however, that none of the estimated average causal mediation effects (ACME) of group membership includes zero in their 90-percent CI, making it a robust finding with respect to explaining the causal effects realized by active learning.

To further investigate the qualitative meaning of this mediation effect we look at whether the active learning treatment causes individuals to save *within* these groups (e.g.

creating savings groups such as ROSCAs or opening group accounts at formal financial institutions) or whether the active learning treatment encourages people to discuss financial affairs and, thus, indirectly benefiting *individual* savings. We therefore disaggregate total savings by distinguishing between individual accounts and locations and group accounts (see Appendix Table A12). We find that the average treatment effect increase in savings is primarily driven by an increase in individual savings (average effect of treatment A is +101,028 UGX, i.e. an increase of over 50 percent, $p<0.1$), and not by savings in group accounts (+ 55,354 UGX, i.e. an increase of 17.61 percent, $p>0.1$). Thus, active learning has not resulted in participants saving together with peers. Rather, active learning appears to have fostered a community of peers discussing their individual financial matters and subsequently adjusting their individual behavior.

6 Treatment effect heterogeneity

This section investigates potentially heterogeneous effects from treatment by an examination of how treatment effects may be conditional on levels of outcomes (Section 6.1) and how treatment effects may interact with observable traits of the target group (Section 6.2).

6.1 Distributional treatment effects

Indeed, simultaneous quantile regressions show that the advantageous effects of the financial education trainings are heterogeneous. Note that the following interpretations make the implicit assumption of rank invariance. Starting with the financial literacy index, the effect of treatment A (active learning) is independent from the outcome level of financial literacy (see Table 7, column 1). Graphically speaking, the entire distribution of knowledge levels appears to be shifted to the right in response to the active learning treatment. By contrast, the effect of treatment B (traditional lecturing) is quite strong only below the median

(although not much stronger than that of treatment A) while coefficients are about zero above the median. These differential effects become especially apparent at the top of the financial literacy distribution ($p<0.1$). We conclude that treatment B may be an alternative approach in this respect if only individuals with very low levels of financial literacy belong to the target group. Active learning, however, appears to be beneficial across the entire range of the financial literacy distribution.

<Table 7 about here>

Regarding the outcome level of the savings index, treatment A has clearly larger and significant effects above the median, while effects of treatment B are very small and insignificant across the whole distribution (see column 2). Thus, differential treatment effects of active learning within the savings domain are driven by large effects in the 60^{th} and 80^{th} percentile, while effect sizes are small below the median. This indicates that financial education may increase savings-inequality with respect to our target group which is consistent with a recent theoretical model of endogenous financial knowledge and wealth inequality (Lusardi et al., 2017). A similar pattern arises with respect to the investment index, although at a higher level of effect sizes (for details see column 3).

Finally, we show distributional treatment effects for the summary index of all financial behavior indices (column 4). Again, effect sizes are consistently larger for treatment A than treatment B across the entire distribution. Differential treatment effects are confirmed at the 20^{th} and 60^{th} percentile. The strongest impacts, again, are estimated to be present at the top of the distribution (80^{th} percentile) with a treatment effect of a quarter of a standard deviation of active learning on the aggregate financial behavior index.

Overall, we draw two conclusions: first, active learning is effective in increasing financial literacy scores across the entire distribution; and second, regarding the savings, the investment and the summary financial behavior indices, effects are largest for those

individuals with higher levels of outcomes and possibly zero for people at the bottom of the outcome distributions (for savings and investing). Thus, financial education seems to be less effective for the poorest individuals in our sample (cf. Lusardi et al., 2017).

6.2 Subgroup analyses

We complement these results by investigating treatment effects by subgroups along observable characteristics of the respondents. We examine the possibility of heterogeneous treatment effects for three subgroups that are known to have different levels of *ex-ante* financial literacy and may respond differently to financial education programs (cf. Lusardi and Mitchell, 2014). First, gender differences are treated as a stylized fact in the literature, with men scoring higher on financial knowledge tests than women in most surveys (see Hsu, 2016). Second, financial literacy is correlated with general educational attainment. In our sample, about 32 percent of individuals cannot read or write in any language and only about 25 percent have more than primary education. Thus, differential impacts conditional on general educational attainment may occur. Third, we examine differential impacts conditional on baseline financial literacy levels. Generally, we do not find evidence for differential treatment effects, accordingly results of these complementary analyses are shown in Appendix Table A13.

One exception may be that those who have above average financial literacy scores at baseline appear to respond more positively to the treatments with respect to budgeting behavior, i.e. respondents may only be able to translate the benefits of the training into action (e.g. creating a written household budget) if they are relatively knowledgeable at baseline.

Interestingly, impacts on downstream financial behaviors, such as savings and investments, are not conditional on prior skill-levels (education or financial literacy). This is in contrast to other experiments: Bjorvatn and Tungodden (2010) report that business training

in Tanzania is most effective for low-skilled individuals. Drexler et al. (2014) show that benefits of "rules of thumb" training are especially driven by large impacts on low-skilled respondents (cf. Drexler et al., 2014, p.25). Finally, Fort et al. (2016) document that exogenous variation in bank information policies impacts financial literacy and financial behavior, with effects highest for low-educated elderly households. Benefits of active learning, however, appear to be universal for our sample and generally not contingent on low ability.

7 Robustness

We demonstrate in various exercises that results of the main paper are robust. These checks are briefly reported here, while details are provided in Appendix B, and cover four main areas: (i) As the sample is characterized by mild attrition, we carefully address this issue showing that there is no reason for concern, among others, by inverse probability weighting for selection into endline survey and bounds analysis testing results under several scenarios; (ii) we show OLS-results using only endline data; (iii) while the measure of financial literacy used here has better psychometric properties than simple unweighted sum-scores of standard items, we show that also a more common measure of sumscores leads to a qualitatively similar result; and finally, (iv) we show results for an alternative method of creating the indices for outcome variables.

8 Conclusion

Our research contributes to revealing the determinants of successful financial education. As a potentially crucial determinant of effective financial education, we compare an active learning approach with traditional lecturing. Active learning is shown to be clearly superior in

the field of science education, thus providing strong motivation to test the approach in financial education.

The main result is clear cut: We study five outcome groups of financial behavior, such as savings and investments, finding that active learning outperforms traditional lecturing. In three out of five cases, active learning has the intended effect on financial behavior to a significant degree, while traditional lecturing never has a significant effect. Coefficients of active learning tend to be clearly larger, although the differential treatment effects compared to lecturing are only significant for the savings domain and for a single item in the borrowing index. Still, given the relatively limited power of our cluster-RCT, with 1,162 vendors in the endline survey and a short educational intervention of only two hours, the advantage of active learning seems quite strong.

In a second step, we aim to explore possible mechanisms intermediating financial education into changes of financial behavior. For this purpose, we examine the role of cognitive and non-cognitive variables, i.e. financial literacy, self-control, and financial confidence. We find that active learning has a positive impact on all three considered intermediating variables, while lecturing only impacts financial confidence. However, evidence on the importance of these intermediate outcomes as causal pathways to behavior change remains uncertain. We also investigate potential social mechanisms that may explain the effectiveness of active learning in financial education: Active learning which relies heavily on social interactions with peers has an effect on increased social group membership which may partly mediate the effects on saving and investing. Last, we document that benefits from active learning are not conditional on the human capital of participants although effects may be especially evident for individuals at the top of the wealth distribution.

Overall, active learning methods may be an effective way to improve existing classroom programs without increasing the costs that result from other approaches, such as

lowering the student to teacher ratio or moving to strictly personalized interventions such as counseling. An important area for future research would be to study the effects of active learning over a longer time horizon and to apply the approach to a broader set of target groups and circumstances.

References

Ashraf, N., Karlan, D., and Yin, W. (2006). Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. *Quarterly Journal of Economics*, 121(2): 635–672.

Berg, G. and Zia, B. (2017). Harnessing emotional connections to improve financial decisions. Evaluating the impact of financial education in mainstream media. *Journal of the European Economic Association*, 15(5): 1025–1055.

Berry, J., Karlan, D., and Pradhan, M. (2018). The impact of financial education for youth in Ghana. *World Development*, 102: 71–89.

Bjorvatn, K, and Tungodden, B. (2010). Teaching business in Tanzania: Evaluating participation and performance. *Journal of the European Economic Association*, 8(2-3): 561–570.

Blattman, C., Fiala, N., and Martinez, S. (2014). Generating skilled self-employment in developing countries: Experimental evidence from Uganda. *Quarterly Journal of Economics*, 129(2): 697–752.

Bruhn, M., Ibarra, G.L., and McKenzie, D. (2014). The minimal impact of a large-scale financial education program in Mexico City. *Journal of Development Economics*, 108: 184–189.

Bruhn, M., de Souza Leao, L., Legovini, A., Marchetti, R., and Zia, B. (2016). The impact of high school financial education: Evidence from a large-scale evaluation in Brazil. *American Economic Journal: Applied Economics,* 8(4): 256–295.

Brown, M., Grigsby, J., van der Klaauw, W., Wen, J., and Zafar, B. (2016). Financial education and the debt behavior of the young. *Review of Financial Studies*, 29(9): 2490–2522.

Bulte, E., Lensink, R., and Vu, N. (2017). Do gender and business trainings affect business outcomes? Experimental evidence from Vietnam. *Management Science*, 63(9): 2885–2902.

Calderone, M., Fiala, N., Mulaj, F. Sadhu, S., and Sarr, L. (2018). Financial education and savings behavior: Evidence from a randomized experiment among low income clients of branchless banking in India. *Economic Development and Cultural Change*, forthcoming.

Campos, F., Frese, M., Goldstein, M., Iacovone, L., Johnson, H. C., McKenzie, D., and Mensmann, M. (2017). Teaching personal initiative beats traditional training in boosting small business in West Africa. *Science*, 357(6357): 1287–1290.

Carpena, F., Cole, S., Shapiro, J., and Zia, B. (2017). The ABCs of financial education. Experimental evidence on attitudes, behavior, and cognitive biases. *Management Science,* https://doi.org/10.1287/mnsc.2017.2819.

Cole, S., Sampson, T., and Zia, B. (2011). Prices or knowledge? What drives demand for financial services in emerging markets? *Journal of Finance*, 66(6): 1933–1967.

Cole, S., Gine, X., Tobacman, J., Topalova, P., Townsend, R., and Vickery, J. (2013). Barriers to household risk management: Evidence from India. *American Economic Journal: Applied Economics*, 5(1): 104–135.

Cole, S., Paulson, A., and Shastry, G.K. (2016). High school curriculum and financial outcomes: The impact of mandated personal finance and mathematics courses. *Journal of Human Resources*, 51(3): 656–698.

Deslauriers, L., Schelew, E., and Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science*, 332(6031): 862–864.

Doi, Y., McKenzie, D., and Zia, B. (2014). Who you train matters: Identifying combined effects of financial education on migrant households. *Journal of Development Economics*, 109: 39–55.

Drexler, A., Fischer, G., and Schoar, A. (2014). Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics*, 6(2): 1–31.

Duflo, E. and Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics*, 118(3): 815–842.

Fernandes, D., Lynch Jr., J.G., and Netemeyer, R.G. (2014). Financial literacy, financial education, and downstream financial behaviors. *Management Science*, 60(8): 1861–1883.

Fort, M., Manaresi, F., and Trucchi, S. (2016). Adult financial literacy and households' financial assets: The role of bank information policies. *Economic Policy*, 31(88): 743–782.

Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H., and Wenderoth, M.P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23): 8410–8415.

Hastings, J.S., Madrian, B.C., and Skimmyhorn, W.L. (2013). Financial literacy, financial education, and economic outcomes. *Annual Review of Economics*, 5: 347–373.

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

Imai, K., L. Keele, D. Tingley, and Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review,* 105: 765–789.

Jack, W., and Suri, T. (2014). Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution. *American Economic Review*, 104(1): 183–223.

Jamison, J.C., Karlan, D, and Zinman, J. (2014). Financial education and access to savings accounts: Complements or substitutes? Evidence from Ugandan youth clubs. *NBER Working Paper 20135.*

Kaiser, T. and Menkhoff, L. (2017). Does financial education impact financial behavior, and if so, when? *World Bank Economic Review*, 31(3): 611–630.

Karlan, D. and Valdivia, M. (2011). Teaching entrepreneurship: Impact of business training on microfinance clients and institutions. *Review of Economics and Statistics*, 93(2): 510–527.

Kling, J.R., Liebman, J.B., and Katz, L.F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1): 83–119.

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

Lusardi, A., Michaud, P.-C., and Mitchell, O.S. (2017). Optimal financial knowledge and wealth inequality. *Journal of Political Economy*, 125(2): 431–477.

Lusardi, A., Samek, A.S., Kapteyn, A., Glinert, L., Hung, A., and Heinberg, A. (2017a). Visual tools and narratives: New ways to improve financial literacy. *Journal of Pension Economics and Finance,* 16(3): 297–323.

Lührmann, M., Serra-Garcia, M., and Winter, J. (2018). The impact of financial education on adolescents' intertemporal choices. *American Economic Journal: Economic Policy,* forthcoming.

Miller, M., Reichelstein, J., Salas, C., and Zia, B. (2015). Can you help someone become financially capable? A meta-analysis of the literature. *World Bank Research Observer*, 30(2): 220–246.

OECD (2015). National strategies for financial education. OECD/INFE policy handbook, https://www.oecd.org/finance/National-Strategies-Financial-Education-Policy-Handbook.pdf.

Ruiz-Primo, M.A., Briggs, D., Iverson, H., Talbot, R., and Shepard, L.A. (2011). Impact of undergraduate science course innovations on learning. *Science*, 331(6022): 1269–1270.

Sayinzoga, A., Bulte, E.H., and Lensink, R. (2016). Financial literacy and financial behaviour: Experimental evidence from rural Rwanda. *Economic Journal*, 126(594): 1571–1599.

Schaner, S. (2017). The cost of convenience? Transaction costs, bargaining power, and savings account use in Kenya. *Journal of Human Resources,* 52(4), 919–945.

Skimmyhorn, W. (2016). Assessing financial education: Evidence from boot camp. *American Economic Journal: Economic Policy*, 8(2): 322–343.

Skimmyhorn, W.L., Davies, E.R., Mun, D., and Mitchell, B. (2016). Assessing financial education methods: Principles vs. rules-of-thumb approaches. *Journal of Economic Education*, 47(3): 193–210.

Steinert, J. I., Cluver, L. D., Meinck, F., Doubt, J., and Vollmer, S. (2018). Household economic strengthening through financial and psychosocial programming: Evidence from a field experiment in South Africa. *Journal of Development Economics*, 134: 443–466.

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

van Rooij, M., Lusardi, A., and Alessie, R. (2012). Financial literacy, retirement planning and household wealth. *Economic Journal* 122(560): 449–478.

Wieman, C. E. (2014). Large-scale comparison of science teaching methods sends clear message. *Proceedings of the National Academy of Sciences*, 111(23): 8319–8320.

Willis, L.E. (2011). The financial education fallacy. *American Economic Review: Papers and Proceedings,* 101(3): 429–434.

Figure 1: Location and treatment status of 83 clusters

Table 1: Sample overview and response-rates

	Wave	Control group	Treatment groups		Total
			Treatment A	Treatment B	
Clusters (marketplaces) (%)	Baseline	28 (33.73%)	27 (32.54%)	28 (33.73%)	83 (100%)
	Endline	28 (33.73%)	27 (32.54%)	28 (33.73%)	83 (100%)
Individuals (%)	Baseline	456 (35.32 %)	414 (32.07%)	421 (32.61%)	1,291 (100%)
	Endline	417 (35.89%)	384 (33.05%)	361 (31.06%)	1,162 (100%)
	Attrition (individuals)	39 (8.55%)	30 (7.25%)	60 (14.25%)	129 (9.99%)

Notes: Randomization was done in Stata and is fully reproducible. The cluster-level dataset contained one duplicate cluster (market) that was known by two different names in the local languages. Thus, randomization was done with 84 markets. This was discovered only after randomization and initial field activities. The duplicate (which was allocated to group A) was removed *ex post*. The baseline survey was conducted between November 2nd and December 19th, 2015. The follow-up survey was conducted between April 6th and July 19th of 2016 with additional tracking efforts and surveys in October 2016 and February 2017 (see Timeline Figure A1 in Appendix A).

Table 2: Summary statistics and randomization-balance at baseline

		Full sample	Control (C)	Treatment (A)		Treatment (B)	
	Obs.	Mean (SD)	Mean (SD)	Mean (SD)	Diff. from C [SE]	Mean (SD)	Diff. from C [SE]
Panel A: Household-level characteristics at baseline							
Household size	1,259	6.83	6.86	7.00	0.14	6.64	-0.22
		(3.77)	(3.52)	(4.25)	[0.33]	(3.52)	[0.33]
No. of contributors	1,277	2.12	2.12	2.06	0.06	2.19	0.06
		(1.98)	(2.01)	(1.95)	[0.14]	(1.98)	[0.12]
No. of children	1,273	4.17	4.29	4.29	0.00	3.94	-0.35
		(2.95)	(2.88)	(3.22)	[0.27]	(2.75)	[0.24]
Monthly consumption^{+}	1,286	592,775	592,219	616,570	24,350	569,925	-22,294
		(408,015)	(402,390)	(448,328)	[49,884]	(370,276)	[39,842]
Panel B: Respondent-level characteristics at baseline							
Monthly income^{+}	1,250	219,867	222,400	203,232	-19,168	233,565	11,165
		(327,192)	(337,538)	(283,402)	[32,471]	(355,164)	[33,255]
Female	1,265	0.80	0.80	0.79	-0.01	0.80	0.00
		(0.40)	(0.40)	(0.41)	[0.05]	(0.40)	[0.06]
Age	1,277	36.23	37.72	35.38	-2.34*	35.46	-2.26*
		(11.89)	(12.36)	(11.53)	[1.18]	(11.59)	[1.19]
Education (>primary)	1,291	0.25	0.28	0.22	-0.06	0.26	-0.02
		(0.43)	(0.45)	(0.41)	[0.04]	(0.44)	[0.04]
Self-control (z)	1,273	0.01	0.00	0.08	0.08	-0.04	-0.04
		(1.01)	(1.00)	(1.02)	[0.08]	(0.99)	[0.08]
Fin. knowledge (z)	1,291	0.03	0.00	0.10	0.10	0.00	0.00
		(0.98)	(1.00)	(0.96)	[0.11]	(0.99)	[0.11]
Fin. confidence (z)	1,259	-0.05	0.00	-0.04	-0.04	-0.11	-0.11
		(0.99)	(1.00)	(0.98)	[0.11]	(1.01)	[0.10]
Any group membership	1,232	0.73	0.74	0.71	-0.03	0.74	-0.00
		(0.44)	(0.44)	(0.45)	[0.04]	(0.44)	[0.04]
Class size	1,291	16.75	17.16	16.70	-0.47	16.36	-0.81
		(4.21)	(3.90)	(4.36)	[1.18]	(4.35)	[1.19]
Panel C: Outcome measures at baseline							
(1) Budgeting index	1,248	-0.02	0.00	-0.10	-0.10	0.04	0.04
		(0.96)	(1.00)	(0.91)	[0.08]	(0.97)	[0.08]
(2) Savings index	1,284	0.06	0.00	0.09	0.09	0.09	0.09
		(1.05)	(1.00)	(1.12)	[0.08]	(1.01)	[0.10]
(3) Borrowing index	1,248	0.06	0.00	0.02	0.02	0.15	0.15
		(1.25)	(1.00)	(1.27)	[0.07]	(1.46)	[0.10]
(4) Investments index	1,191	-0.01	0.00	-0.06	-0.06	0.03	0.03
		(1.03)	(1.00)	(1.03)	[0.13]	(1.05)	[0.12]
(5) Fin. services index	1,241	0.02	0.00	-0.04	-0.04	0.11	0.11
		(1.07)	(1.00)	(0.99)	[0.10]	(1.07)	[0.11]
F-test of joint orthogonality (p-value)					0.71		
Observations					934		
Clusters					83		

Notes: + indicates that the currency denominated outcome (in Ugandan Shilling (UGX)) is winsorized at the 99th percentile. Differences between treatment and control groups are estimates from OLS-regressions. Standard errors (clustered at the market-level) are reported in square brackets. Tests are unadjusted for multiple hypothesis testing. The F-test of joint orthogonality regresses a categorical variable indicating the three groups on the full set of variables. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Main experimental results (ANCOVA)

	(1) Budgeting index (z)	(2) Savings index (z)	(3) Borrowing index (z)	(4) Investment index (z)	(5) Fin. services index (z)	(6) Summary index (z)
Treatment A	0.045	0.165**	0.108*	0.258***	0.109	0.207**
	(0.087)	(0.071)	(0.057)	(0.091)	(0.077)	(0.083)
Treatment B	0.009	0.013	-0.037	0.113	0.149	0.032
	(0.088)	(0.077)	(0.075)	(0.095)	(0.095)	(0.090)
$A - B = 0$ (p-value)	0.690	0.077*	0.047**	0.150	0.678	0.073*
R^2	0.108	0.152	0.009	0.326	0.130	0.237
Mean (SD) of y_t in	0.000	0.000	0.000	0.000	0.000	0.000
control group	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Observations	1,114	1,161	1,110	1,007	1,111	895
Clusters	83	83	83	83	83	83
District FEs	yes	yes	yes	yes	yes	yes
$y_{(t-1)}$ covariate	yes	yes	yes	yes	yes	yes

Notes: Table shows OLS regression results of ANCOVA models. The dependent variables (y_t) are equally weighted z-score indices of financial behavior and are standardized to have a zero mean and a standard deviation of one for the control group. Thus, coefficients can be interpreted as effect sizes (Glass's Δ). All currency denominated outcomes (in Ugandan Shilling (UGX)) within the indices in columns (2), (3) and (4) are winsorized at the 99th percentile. Column (6) shows results for an equally weighted z-score index of all five financial behavior indices. All models include the lagged outcome at baseline and district-level fixed effects. Standard errors (clustered at the market-level) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Analysis of individual index components saving and investing

	(a) Savings index			(b) Investment index		
	(1) Any savings	(2) Total savings+	(3) Net savings+	(4) Total investments+	(5) Business formally registered	(6) Total number of durable items owned
Treatment A	0.038*	109,186*	145,481**	90,173*	0.077**	5.408***
	(0.020)	(65,132)	(64,785)	(47,461)	(0.034)	(2.027)
Treatment B	0.027	-32,519	-14,226	41,801	0.060	2.215
	(0.023)	(68,588)	(66,477)	(63,124)	(0.037)	(1.625)
$A - B = 0$ (p-value)	0.601	0.052*	0.025**	0.487	0.640	0.088*
R^2	0.024	0.244	0.131	0.184	0.035	0.520
Mean (SD) of y_t	0.878	513,629	380,568	301,067	0.232	51.384
in control group	(0.328)	(937,119)	(973,769)	(526,957)	(0.423)	(34.344)
Observations	1,161	1,162	1,162	1,053	1,110	1,162
Clusters	83	83	83	83	83	83
District FEs	yes	yes	yes	yes	yes	yes
$y_{(t-1)}$ covariate	yes	yes	yes	yes	yes	yes

Notes: Table shows OLS regression results of ANCOVA models. Columns (1) and (5) are linear probability models. All models include the lagged outcome at baseline and district-level fixed effects. + indicates that the outcome (in Ugandan Shilling (UGX)) is winsorized at the 99th percentile. Standard errors (clustered at the market-level) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Results on potential psychological mediators

	(1) Fin. literacy (z)	(2) Self-control (z)	(3) Fin. confidence (z)
Treatment A	0.151*	0.147*	0.159*
	(0.086)	(0.079)	(0.082)
Treatment B	0.076	-0.063	0.192**
	(0.076)	(0.092)	(0.083)
$A - B = 0$ (p-value)	0.375	0.019**	0.697
R^2	0.026	0.022	0.009
Mean (SD) of y_t	0.000	0.000	0.000
in control group	(1.000)	(1.000)	(1.000)
Observations	1,162	1,157	1,027
Clusters	83	83	83
District FEs	yes	yes	yes

Notes: Results from OLS regressions. Standard errors (clustered at the market-level) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Results on social mechanisms for the active learning treatment

	(1) Any group membership	(2) Savings index (z)	(3) Investment index (z)	(4) Summary Index (z)
Treatment A	0.047*			
	(0.026)			
Treatment B	-0.013			
	(0.029)			
$A - B = 0$ (p-value)	0.041**			
R^2	0.111			
ACME (Any group membership)		0.022	0.010	0.016
[90% CI]		[0.003, 0.043]	[0.001, 0.021]	[0.001, 0.034]
Direct Effect (Treatment A)		0.126	0.244	0.188
[90% CI]		[0.014, 0.235]	[0.095, 0.389]	[0.050, 0.323]
Total Effect (Treatment A)		0.148	0.254	0.204
[90% CI]		[0.029, 0.259]	[0.101, 0.400]	[0.059, 0.339]
% of Total effect mediated		14.88	3.84	7.96
[90% CI]		[8.00, 54.62]	[2.41, 9.43]	[4.71, 25.14]
Mean (SD) of y_t	0.805	0.000	0.000	0.000
in control group	(0.397)	(1.000)	(1.000)	(1.000)
Observations	1,155	1,154	1,002	891
Clusters	83	83	83	83
District FEs	yes	yes	yes	yes
$y_{(t-1)}$ covariate	yes	yes	yes	yes

Notes: Results from causal mediation analysis (Imai et al. 2011). All modes include the lagged dependent variable $y_{(t-1)}$ and district fixed effects. Standard errors (clustered at the market-level) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Simultaneous-quantile regressions for indices of financial knowledge, savings, and investment

		(1) Fin. literacy (z)	(2) Savings index (z)	(3) Investment index (z)	(4) Summary index (z)
20th percentile	Treatment A	0.161	0.038	0.051	0.123*
		(0.119)	(0.032)	(0.040)	(0.073)
	Treatment B	0.281**	0.017	0.038	-0.043
		(0.120)	(0.034)	(0.050)	(0.082)
	$A - B = 0$ (p-value)	0.372	0.412	0.801	0.032**
40th percentile	Treatment A	0.180*	0.027	0.102*	0.130
		(0.109)	(0.020)	(0.058)	(0.079)
	Treatment B	0.204*	-0.001	0.089	0.065
		(0.108)	(0.020)	(0.057)	(0.094)
	$A - B = 0$ (p-value)	0.804	0.210	0.839	0.482
Median	Treatment A	0.108	0.037	0.180**	0.112
		(0.111)	(0.027)	(0.075)	(0.085)
	Treatment B	0.101	-0.003	0.098	0.055
		(0.096)	(0.027)	(0.060)	(0.083)
	$A - B = 0$ (p-value)	0.941	0.177	0.292	0.509
60th percentile	Treatment A	0.154	0.068**	0.255***	0.163*
		(0.135)	(0.029)	(0.091)	(0.096)
	Treatment B	-0.022	0.007	0.101	0.007
		(0.115)	(0.029)	(0.070)	(0.091)
	$A - B = 0$ (p-value)	0.154	0.064*	0.098*	0.090*
80th percentile	Treatment A	0.136	0.137*	0.471***	0.258***
		(0.095)	(0.072)	(0.127)	(0.098)
	Treatment B	-0.037	-0.037	0.227	0.096
		(0.096)	(0.057)	(0.149)	(0.114)
	$A - B = 0$ (p-value)	0.091*	0.021**	0.090*	0.165
Observations		1,162	1,162	1,007	895
Clusters		83	83	83	83
District FEs		yes	yes	yes	yes
$y_{(t-1)}$ covariate		yes	yes	yes	yes
Mean (SD) of y_t in control group		0.000 (1.000)	0.000 (1.000)	0.000 (1.000)	0.000 (1.000)

Notes: Results present simultaneous-quantile regressions for the impact of the financial education treatments A and B on indices of financial knowledge (1), savings (2), and investments (3). The dependent variables are equally weighted z-score indices of their respective components as discussed in Section 2.4 and are standardized to have a zero mean and a standard deviation of one for the control group. Thus, coefficients can be interpreted as effect sizes (Glass's Δ). The output presents treatment effects for each quartile and the median. All currency denominated outcomes (in Ugandan Shilling (UGX)) within the indices in columns (2) and (3) are winsorized at the 99th percentile. All models include the lagged outcome at baseline and district-level fixed effects. Standard errors in parentheses are bootstrapped with 1,000 replications. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix

(online appendix not intended for print publication)

to accompany

"**Active learning improves financial education: Experimental evidence**"

Appendix A: Supplementary tables and figures

Appendix B: Robustness checks

Appendix C: Measuring financial literacy

Appendix A: Supplementary tables and figures

Figure A1: Timeline

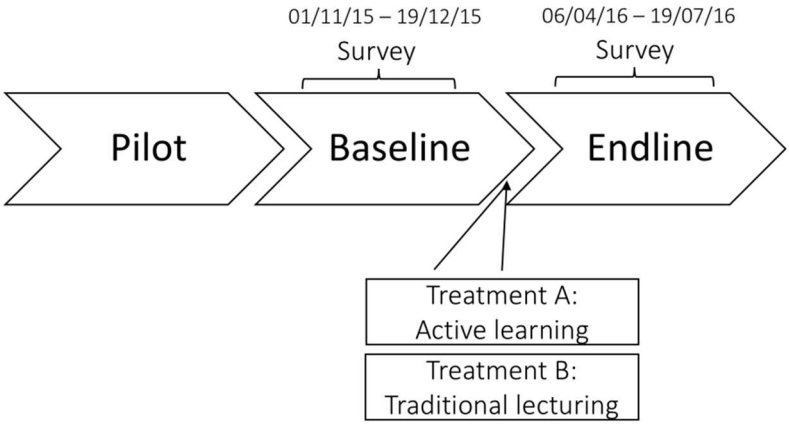

Table A1: Overview of financial education treatments

	Topic	Treatment A: Active learning		Treatment B: Traditional lecturing	
		Time (mins)	Activity	Time (mins)	Activity
1	Personal financial management	3	The trainer introduces the topic by asking the participants what they think is involved in personal finance.	20	The lecturer writes the headings of each topic on a flip chart and tells the participants about the learning objectives of the lecture. Learning objectives of the first topic are: - The participants are able to differentiate between inflows and outflows - The participants are able to differentiate between wants and needs - The participants know how to create a simple budget
		10	The trainer reads out a case study of a family with five children, three sources of income and several expenditures. The trainer explains the concept of the "money tree" (metaphor for household budget) with a display and invites participants to identify the family's inflows (roots) and outflows (leaves) and to place them on the "money tree". The trainer asks the participants to calculate the total amounts of income and expenditures. Total expenditures are larger than total incomes.		
		5	The trainer asks the participants what they notice about the family's money tree (defining the problem of overspending) and what could be done about it. The trainer then introduces the concept of needs and wants and asks participants to differentiate between what they consider are the family's "needs" and "wants" and to remove the wants from the money tree and cluster them elsewhere.		The lecturer starts with the topic of "personal financial management" and discusses the following keywords: - Financial priorities (separating needs from wants) - What to use money for: spend, save, invest - Inflows/outflows - Budget - Setting goals and targets - Keeping track of spending
		4	The trainer introduces the new family budget (after removing the "wants") and links it to the money tree. A picture of a written household budget is shown to the participants and the trainer asks the participants to give their views on how the family could spend the resulting surplus.	5	The lecturer asks the participants whether they have questions or comments.
		3	The session is being wrapped up by asking the participants how they would rate the importance of financial management and whether the introduced concepts are relevant for their own lives.		
2	Savings	3	The trainer, again, refers to the case of the hypothetical family and pins up a definition of "saving" followed by personal questions to the participants: Who is saving? What are you saving for? How often are you saving? Why are you saving? Trainer pins mentioned reasons on a poster and adds those mentioned and notion model cards.	20	The lecturer writes the topic of this section on a flip chart and defines the learning objectives: - The participants know the meaning of savings - The participants know at least 3 ways to save - The participants know at least 2 benefits of savings - The participants know the trade-off between ease of access (liquidity) and return
		8	The trainer introduces different ways to save (account, cash, durables,…) and explains that each way of saving comes with a specific "ease of access and return" profile and that these generally resemble conflicting goals. The trainer now asks the participants to indicate their preferences with regard to this trade-off by means of placing themselves on a line on the floor where the far left of the continuum indicates "easy to access" and the far right indicates "high return". Trainer asks the participants to share the rationale behind their decisions. The trainer now chooses one example of saving forms at a time and asks participants where they would place the different examples of ways of saving. The trainer encourages discussion about these decisions.		The lecturer writes the following keywords on a flipchart and discusses the topics with the participants. - Saving in kind vs. in cash - Conflicting savings-goals: - safety - return - ease of access The lecturer draws a line on the flipchart, illustrating the ease of access and return trade-off. The lecturer illustrates this using the examples of 'money under the mattress' and 'money in a fixed deposit account' on the line.
		9	The trainer reads statements on "saving" and	5	The lecturer asks the participants whether they

			asks the participant to cross out wrong statements. This results in a collection of ten statements on the benefits of savings which are subsequently repeated.		have questions or comments.
3	Borrowing	5	The trainer asks the participants to share their experiences with loans and subsequently displays a card with a formal definition of a loan.	20	The lecturer writes the topic of this section on a flip chart and defines the learning objectives: - The participants are able to differentiate between a sensible and non-sensible reasons to take up a loan - The participants know typical direct and indirect costs of borrowing
		5	The trainer refers to the hypothetical family and adds more detail to their financial situation. The participants are asked to identify three things that the hypothetical family plans to borrow money for (consumption vs. productive investments).		The lecturer writes the following keywords on a flipchart and discusses the topics with the participants: - Reasons for borrowing: - to finance productive investments - to finance consumption goods - to cater for emergencies
		7	The trainer asks the participant what they consider sensible and non-sensible reasons to take up a loan. The participants use cards with example pictures and cluster these on a poster into the two categories. Loans are evaluated on the basis on whether they may put to productive use. Participants agree on the advice they would give this family on which loan to take and which loan not to take.		The cost of borrowing - Direct costs vs. indirect costs of borrowing Questions to ask a lender before taking up a loan: - Interest rate - Collateral - Installments (how much, how often) - Penalties for delinquency
		4	The trainer informs the group that a loan comes with (direct and indirect) costs and is usually tied to specific conditions. The participants are asked to place cards with the various costs of borrowing on the poster.	5	The lecturer asks the participants whether they have questions or comments.
		2	The trainer concludes with the station and asks whether the participants have any further questions regarding borrowing. The trainer cautions the participants against the use of expensive credit to finance consumption expenditures.		
4	Investment	4	The trainer pins up a card with the word "investment" and asks the participants for a definition. The trainer complements this discussion with a formal definition.	20	The lecturer writes the topic of this section on a flip chart and defines the learning objectives: - The participants know what an investment is - The participants know different forms of investment - The participants know the trade-off between safety (minimal risk) and return
		13	The trainer requests the participants to reflect on the discussion they have just had about and asks the participants to share their experiences with regard to investments they have made themselves. The trainer then introduces illustrative cards that display either consumption or investment activates and asks participant to assess whether the cards indicate productive investments. The trainer informs participants that each form of investment comes with a unique "safety-return" profile. "Safety" and "return" represent conflicting goals and that an investment always comes with certain risks. The trainer now asks the participants to indicate their preferences with regard to this trade-off by means of placing themselves on a line where the far left of the continuum indicates "safety" and the far right indicates "high return". Trainer asks the participants to share the rationale behind their decisions. The trainer now chooses one example of investment forms at a time and asks participants where they would place the different examples of ways of investing. The trainer encourages discussion about these decisions.		The lecturer writes the following keywords on a flipchart and discusses the topics with the participants: Forms of investments: - Animals - Land - Business (own and other's) - Buildings - Why invest? - Create wealth and security - Increase the ability to earn more income - Planning for old age - Create employment opportunities for oneself and others

				- Short term investment examples - Medium term investment examples - Long term investment examples
	8	The trainer refers to the case of the hypothetical family and asks the participants to summarize the investment decisions the family has taken. <u>The trainer asks the participants to summarize the associated risks and benefits of these investment decisions.</u> The trainer asks the participants whether they are aware of strategies to manage these risks. Afterwards he introduces the notion of formal and informal insurance, as well as insurance through diversification.		<u>Investment Risks:</u> - Loss of value (depreciation) - Theft - Mismanagement - Assets destroyed or damaged The lecturer draws a line on the flipchart, illustrating the ease of "safety" and "return" trade-off. The lecturer illustrates this using the examples of "land" and " livestock" on the line. <u>Risk management strategies:</u> - Diversification ("do not put all eggs in one basket") - Formal and informal insurance: - Insurance premium - Insurance coverage - Insurance contract
			5	The lecturer asks the participants whether they have questions or comments.
Financial service providers *-continued-*	6	The trainer refers to the hypothetical case of the family and <u>asks the participants to name institution</u>s where the family could save money at. The trainer classifies these answers into regulated and non-regulated institutions (by the central bank).	20	The lecturer writes the topic of this section on a flip chart and defines the learning objectives: - The participants know the difference between regulated and unregulated financial service providers - The participants know rights and responsibilities of financial service users - The participants know different options to make money-transfers and payments
	6	The trainer discusses advantages and disadvantages of financial institutions regulated or not-regulated by the central bank and <u>asks the participants to give the hypothetical family advice</u> on where to save the money.		The lecturer writes the following keywords on a flipchart and discusses the topics with the participants: Classification of financial services sector in Uganda: - Tier i - Commercial Banks - Tier ii – Credit Institutions - Tier iii – Micro Deposit Taking Institutions - Tier iv – Other Financial Institutions (e.g. VSLAs, ROSCA Unregulated vs. regultated by the Bank of Uganda (Tier i to iii)
	7	Trainer introduces the aspect of rights and responsibilities of consumers of financial services. The trainer informs participants that they have rights and responsibilities as Financial service consumers/users. The trainer asks the <u>participants to complete a true/false exercise</u> on statements related to consumer protection rights.		- Rights and responsibilities of consumers - <u>Payments:</u> - Understanding the costs involved - Keeping ones personal information secure - Make safer payments (track who got it) - Mobile money & transaction costs - Automated Teller Machines (ATMs)
	5	The trainer moves to a discussion of payment services and <u>asks the participants to name different ways of transferring money</u> (i.e. for remittances) and asks the participants to discuss the costs attached to these services. The trainer closes the station by encouraging the participants to compare prices and to analyze all options available to them to make sound financial decisions.	5	The lecturer asks the participants whether they have questions or comments.

Further checks of randomization balance.

Table A2 reports additional summary statistics for a rich set of covariates in the full sample and each experimental condition at baseline.

< Table A2 about here>

Panel A shows variables that measure characteristics at the household level. The average household size is 6.83 people, with an average of 2.17 adults contributing to the household's income, a mean of 4.17 children being supported and a mean of 0.36 adults who do not generate external income, such as elderly (plus 0.13 for missing values). The mean (winsorized) monthly household consumption value is about 593,000 UGX.

Panel B reports variables at the respondent-level. The mean (winsorized) monthly individual income is around 220,000 UGX (about 60 USD). Household consumption is higher than hypothetical added individual incomes because of subsistence farming, being reported by 83 percent of the sample. Our sample is predominantly comprised of women (80%) and the average age is 36.2 years. On average, participants have been vending goods on markets for 7.4 years. Only 14 percent report to be selling non-food items (mainly second-hand clothing). The other 86 percent of the sample sell either fresh agricultural products or prepared food. Over two thirds (68 percent) are able to read and write in at least one language and about 25 percent participated in education beyond primary school. About two thirds of the sample report to be married. 70 percent of the respondents state that they are the main contributors of income to the household, while 55 percent report to be the "head of the household". Only 16 percent report to be economically dependent on others. Responding to another survey question, 22 percent of the sample receive aid or assistance from either NGOs or government programs.

We elicit general and domain-specific risk attitudes using common non-incentivized survey items popularized by Dohmen et al. (2011). These survey-items ask for willingness to

take risk on a 0 to 10 scale with highest risk tolerance at 10. On average, respondents are relatively risk-averse. This applies both to the general risk attitudes (mean of 3.61) and to risk attitudes regarding the financial domain (mean of 3.78). The mode and median are at 3 for both the general and the domain-specific case, and thus lower than for a representative sample of the German population with a modal response of 5 (Dohmen et al., 2011). Our survey also includes a measure of numeracy and several psychological variables which are standardized into z-scores to have a mean of zero and a standard-deviation equal to one in the pooled sample. These are mainly used as controls, for the purpose of probing randomization balance in multiple dimensions, and for the investigation of possible causal mechanisms (see Section 5).

Panel C shows descriptive statistics for outcome measures of financial literacy and financial behavior indices (standardized to have a mean of zero and a standard deviation equal to one) at baseline. Again, the outcomes are clearly balanced as no statistically significant differences exist. This also extends to an analysis of the descriptive statistics for individual index components in Table A3.

< Table A3 about here>

Apart from one outcome within the borrowing-index, all components are fully balanced at baseline and a joint test of orthogonality, again, confirms that the three groups are balanced with respect to their levels of the main outcome variables at baseline (p=0.687).

Table A2: Additional descriptive statistics and randomization balance for a rich set of baseline covariates

		Full sample	Control (C)	Treatment (A)		Treatment (B)	
	Obs.	Mean (SD)	Mean (SD)	Mean (SD)	Diff. from C [SE]	Mean (SD)	Diff. from C [SE]
Panel A: Household characteristics at baseline							
Household size	1,259	6.83	6.86	7.00	0.14	6.64	-0.22
		(3.77)	(3.52)	(4.25)	[0.33]	(3.52)	[0.33]
No. of contributors	1,277	2.12	2.12	2.06	0.06	2.19	0.06
		(1.98)	(2.01)	(1.95)	[0.14]	(1.98)	[0.12]
No. of children	1,273	4.17	4.29	4.29	0.00	3.94	-0.35
		(2.95)	(2.88)	(3.22)	[0.27]	(2.75)	[0.24]
No. of rooms	1,284	3.23	3.37	3.14	-0.23	3.16	-0.21
		(1.84)	(1.94)	(1.78)	[0.17]	(1.79)	[0.17]
Owns dwelling	1,291	0.74	0.78	0.70	-0.08*	0.72	-0.56
		(0.44)	(0.41)	(0.46)	[0.05]	(0.45)	[0.05]
Assets (z-score)	1,162	0.00	0.07	-0.06	-0.13	-0.01	-0.08
		(1.00)	(1.07)	(0.99)	[0.13]	(0.93)	[0.11]
Tap water	1,291	0.46	0.40	0.50	0.10	0.47	0.07
		(0.50)	(0.49)	(0.50)	[0.08]	(0.50)	[0.09]
Monthly consumption[+]	1,286	592,775	592,219	616,570	24,350	569,925	-22,294
		(408,015)	(402,390)	(448,328)	[49,884]	(370,276)	[39,842]
Panel B: Respondent characteristics at baseline							
Monthly income[+]	1,250	219,867	222,400	203,232	-19,168	233,565	11,165
		(327,192)	(337,538)	(283,402)	[32,471]	(355,164)	[33,255]
Female	1,265	0.80	0.80	0.79	-0.01	0.80	0.00
		(0.40)	(0.40)	(0.41)	[0.05]	(0.40)	[0.06]
Age	1,277	36.23	37.72	35.38	-2.34*	35.46	-2.26*
		(11.89)	(12.36)	(11.53)	[1.18]	(11.59)	[1.19]
Education	1,282	6.83	7.11	6.61	-0.49	6.74	-0.36
		(3.69)	(3.66)	(3.71)	[0.38]	(3.70)	[0.35]
Literate	1,238	0.68	0.70	0.64	-0.06	0.68	-0.02
		(0.47)	(0.46)	(0.48)	[0.05]	(0.47)	[0.04]
Econ. dependent	1,285	0.16	0.15	0.16	0.01	0.16	0.01
		(0.36)	(0.36)	(0.37)	[0.03]	(0.37)	[0.03]
Receives aid	1,277	0.22	0.24	0.21	-0.03	0.21	-0.03
		(0.41)	(0.43)	(0.41)	[0.04]	(0.40)	[0.03]
Married	1,291	0.62	0.59	0.60	0.01	0.66	0.07
		(0.49)	(0.49)	(0.49)	[0.04]	(0.48)	[0.04]
Main contrib.	1,291	0.70	0.70	0.74	0.04	0.67	-0.03
		(0.46)	(0.46)	(0.44)	[0.04]	(.47)	[0.04]
HH head	1,291	0.55	0.55	0.58	0.03	0.51	-0.04
		(0.50)	(0.50)	(0.49)	[0.05]	(0.50)	[0.05]
Years as vendor	1,263	7.42	8.25	7.77	-0.48	6.18	-2.07**
		(7.46)	(8.08)	(7.41)	[0.90]	(6.63)	[0.86]
Sells nonfood items	1,291	0.14	0.14	0.14	0.00	0.14	0.00
		(0.35)	(0.35)	(0.35)	[0.04]	(0.35)	[0.04]
Numeracy (z-score)	1,291	0.05	0.01	0.07	0.06	0.07	0.06
		(0.97)	(0.97)	(0.94)	[0.08]	(0.98)	[0.09]
Self-control (z-score)	1,273	0.01	0.00	0.08	0.08	-0.04	-0.04
		(1.01)	(1.00)	(1.02)	[0.08]	(0.99)	[0.08]
Fin. literacy score	1,291	0.03	0.00	0.10	0.10	0.00	0.00
		(0.98)	(1.00)	(0.96)	[0.11]	(0.99)	[0.11]
Trust (z-score)	1,291	0.00	0.00	-0.04	-0.04	0.04	0.05
		(1.00)	(0.98)	(1.02)	[0.08]	(1.00)	[0.08]
Altruism (z-score)	1,267	0.00	-0.04	0.05	0.09	0.00	0.04
		(1.00)	(0.99)	(1.02)	[0.08]	(0.98)	[0.06]
Fatalist worldview (z-score)	1,253	0.00	0.03	-0.02	-0.05	-0.01	-0.04
		(1.00)	(0.99)	(1.03)	[0.08]	(0.98)	[0.10]
General risk attitude (0-10)	1,262	3.61	3.53	3.67	0.14	3.66	0.13
		(2.42)	(2.39)	(2.40)	[0.17]	(2.48)	[0.19]
Specific risk attitude (0-10)	1,272	3.78	3.72	3.78	0.06	3.85	0.13
		(2.52)	(2.53)	(2.46)	[0.16]	(2.57)	[0.19]
F-test of joint orthogonality (p-value)					0.79		
Observations					886		
Clusters					83		

Table A3: Descriptive statistics and randomization-balance for components of outcome indices at baseline

		Full Sample	Control (C)	Treatment (A)		Treatment (B)	
	Obs.	Mean (SD)	Mean (SD)	Mean (SD)	Diff. C [SE]	Mean (SD)	Diff. C [SE]
(1) Budgeting index	1,248	-0.02	0.00	-0.10	-0.10	0.04	0.04
		(0.96)	(1.00)	(0.91)	[0.08]	(0.97)	[0.08]
Knows how to write a budget	1,291	0.11	0.12	0.10	-0.02	0.10	-0.02
		(0.31)	(0.33)	(0.31)	[0.02]	(0.30)	[0.02]
Usually keeps track of spending	1,283	0.61	0.59	0.59	0.00	0.64	0.05
		(0.49)	(0.49)	(0.49)	[0.05]	(0.48)	[0.04]
Separates business and private budget	1,261	0.32	0.32	0.28	-0.04	0.36	0.04
		(0.47)	(0.47)	(0.45)	[0.04]	(0.48)	[0.04]
Keeps a written budget	1,275	0.23	0.24	0.20	-0.04	0.25	0.01
		(0.42)	(0.43)	(0.40)	[0.03]	(0.43)	[0.04]
Has kept track of spending in last 6 month	1,280	0.23	0.24	0.20	-0.04	0.24	-0.00
		(0.42)	(0.43)	(0.40)	[0.03]	(0.43)	[0.04]
(2) Savings index	1,284	0.06	0.00	0.09	0.09	0.09	0.09
		(1.05)	(1.00)	(1.12)	[0.08]	(1.01)	[0.10]
Any savings	1,284	0.79	0.78	0.78	0.00	0.81	0.03
		(0.41)	(0.42)	(0.41)	[0.03]	(0.40)	[0.03]
Total savings+	1,291	379,602	336,102	409,960	73,859	396,865	60,763
		(761,556)	(667,250)	(845,120)	[73,423]	(770,338)	[78,023]
Net savings+	1,291	259,957	224,729	294,348	69,619	264,296	39,567
		(768,922)	(760,300)	(821,473)	[55,647]	(723,618)	[66,127]
(3) Borrowing index	1,248	0.06	0.00	0.02	0.02	0.15	0.15
		(1.25)	(1.00)	(1.27)	[0.07]	(1.46)	[0.10]
Loans for consumption+	1,291	4,780	3,592	5,219	1,627	5,634	2,042
		(29,378)	(24,862)	(30,854)	[2,267]	(32,290)	[2,048]
Loans for productive investment+	1,291	169,408	151,022	164,710	13,688	193,943	42,921
		(434,795)	(387,398)	(419,685)	[38,571]	(493,829)	[39,317]
Debt to asset ratio (z)	1,288	0.02	0.00	0.00	0.00	0.05	0.05
		(0.94)	(1.00)	(0.84)	[0.07]	(0.96)	[0.07]
Loan if no clear plan	1,291	0.03	0.01	0.03	0.02**	0.05	0.04***
		(0.18)	(0.10)	(0.18)	[0.01]	(0.22)	[0.02]
Can identify “bad” loan	1,250	0.87	0.89	0.84	-0.05*	0.86	-0.03
		(0.34)	(0.31)	(0.36)	[0.03]	(0.34)	[0.03]
(4) Investments index	1,191	-0.01	0.00	-0.06	-0.06	0.03	0.03
		(1.03)	(1.00)	(1.03)	[0.13]	(1.05)	[0.12]
Total investments+	1,229	419,777	395,130	418,060	22,930	447,655	52,525
		(805,892)	(753,875)	(842,834)	[87,898]	(823,972)	[91,532]
Business formally regist.	1,252	0.24	0.23	0.24	0.01	0.27	0.04
		(0.43)	(0.42)	(0.43)	[0.03]	(0.44)	[0.03]
Total number of items owned	1,291	49.10	51.32	46.95	-4.37	48.80	-2.52
		(31.85)	(35.58)	(30.50)	[4.25]	[28.61]	[3.42]
(5) Fin. services index	1,241	0.02	0.00	-0.04	-0.04	0.11	0.11
		(1.07)	(1.00)	(0.99)	[0.10]	(1.07)	[0.11]
Ever purchased a formal Insurance product	1,261	0.07	0.07	0.06	-0.01	0.08	0.01
		(0.25)	(0.25)	(0.24)	[0.02]	(0.27)	[0.02]
Ever used mobile money	1,262	0.46	0.45	0.44	-0.01	0.50	0.05
		(0.50)	(0.50)	(0.50)	[0.05]	(0.50)	[0.06]
F-test of joint orthogonality (p-value)						0.543	
Observations						1,098	
Clusters						83	

Notes: + indicates that the currency denominated outcome (in Ugandan Shilling (UGX)) is winsorized at the 99th percentile. Standard errors (clustered at the market-level) are reported in square brackets. Tests are unadjusted for multiple hypothesis testing. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Treatment effects on further components within indices.

The *budgeting index* consists of five binary items capturing changes in budgeting and record keeping behavior (see Table A4). These binary items measure whether respondents know how to create a written budget and ask for budgeting behavior, i.e. separating personal and business records and regularly keeping track of expenditures. Of the ten reported marginal effects (considering the two financial education treatments), five have a size of about 0.04 to 0.05 percentage point improvement (on averages of 9 to 66 percent in the control group), out of which four have a positive sign; however, the single negatively signed coefficient is the only significant coefficient, indicating adverse effects of treatment B on the ability to correctly draft a budget. Thus, overall average effects in this domain are insignificant from zero.

< Table A4 about here>

The *borrowing index* consists of six items (see Table A5). Items aggregated into this index capture the structure and volume of debt along with a proxy for debt-bearing capacity and binary items indicating whether respondents would take up loans without a plan or are able to distinguish between sensible and non-sensible reasons to take up a loan with interest. As predicted by the existing literature, treatment effects in this domain are very modest, especially with regard to financial outcomes (cf. Miller et al., 2015; Kaiser and Menkhoff, 2017). The coefficients on the volume of loans intended for consumption purposes and productive investments are extremely small and, thus, insignificant from zero. The same is true for the effect on the asset to debt ratio proxying the debt-bearing capacity and the ability of respondents to distinguish between sensible and non-sensible reasons to take up a loan with interest in hypothetical settings. There is a significant effect for the active learning treatment, however, on the self-report of respondents to take-up a loan (if offered) when they had no clear plan of how to use the money: approximately 14 percent of the control group report to

be willing to take up a loan, even if they had no clear plan on how to use the borrowed amount. The marginal effect of the active learning treatment results in a 5.5 percentage point decrease in the willingness to take up a loan without a clear plan of its utilization. Thus, this particular effect is relatively large (39.2 percent relative to the control group) resulting in the reported overall small and marginally-significant positive impact on the aggregated index discussed above. Treatment B, on the other hand, has a much smaller effect (reduction of 3.8 percentage points or 27 percent) and is statistically insignificant from zero.

< Table A5 about here>

Finally, the *financial service index* (see Table A6) includes two binary items indicating whether respondents were ever covered by a formal insurance product or if they have ever used mobile money (payment) services. Both treatments seem to have the expected positive sign, and the coefficient of treatment A on mobile money use is marginally significant. 44 percent of the control group report to have ever used mobile money services and active learning has a marginal effect of 6.3 percentage points on mobile money use. Lecturing, however, has a much smaller effect size (3.5 percentage points) and is statistically insignificant from zero. Both treatments have zero-effects on formal insurance use (an increase of 0.6 and 5 percentage points) relative to a control-group mean of 9.9 percent).

< Table A6 about here>

Overall, even if the three aggregated indices on budgeting, and financial services do not show the desired significant coefficients (see Table 3), a closer look at the disaggregated items suggests that there are selected subdomains where treatments which were more focused or more intensive may be able to impact financial behavior.

Reference in Appendix A

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3): 522–550.

Table A4: Budgeting index - individual components

	(1) Knows how to write a budget	(2) Usually keeps track of spending	(3) Separates business and private budget	(4) Keeps a written budget	(5) Has kept track of spending in last 6 month
Treatment A	0.004	0.046	0.046	-0.008	-0.007
	(0.025)	(0.041)	(0.038)	(0.037)	(0.035)
Treatment B	-0.047**	0.039	0.064	0.003	0.007
	(0.021)	(0.039)	(0.043)	(0.034)	(0.033)
$A - B = 0$ (p-value)	0.015**	0.842	0.661	0.751	0.681
R^2	0.011	0.046	0.032	0.149	0.149
Mean of y_t in C	0.098	0.665	0.450	0.284	0.274
Observations	1,162	1,160	1,131	1,150	1,153
Clusters	83	83	83	83	83
District FEs	yes	yes	yes	yes	yes
$y_{(t-1)}$ covariate	yes	yes	yes	yes	yes

Notes: Coefficients show results from linear probability models. All models include the lagged outcome at baseline and district-level fixed effects. Standard errors (clustered at the market-level) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A5: Borrowing index – individual components

	(1) Loans for consumption*	(2) Loans for productive investment	(3) Debt to asset ratio (z)*	(4) Would take loan if no clear plan*	(5) Can correctly identify a "bad" loan
Treatment A	2,142	-15,615	-0.086	-0.055**	-0.001
	(5,840)	(37,458)	(0.073)	(0.023)	(0.010)
Treatment B	12,855	-43,846	0.007	-0.038	0.004
	(9,345)	(36,785)	(0.090)	(0.023)	(0.010)
$A - B = 0$ (p-value)	0.231	0.4043	0.272	0.495	0.900
R^2	0.010	0.108	0.119	0.014	0.006
Mean (SD) of y_t in control group	20,763	151,021	0.000	0.139	0.029
	(98,688)	(387,398)	(1.000)	(0.346)	(0.167)
Observations	1,158	1,158	1,155	1,116	1,162
Clusters	83	83	83	83	83
District FEs	yes	yes	yes	yes	yes
$y_{(t-1)}$ covariate	yes	yes	yes	yes	yes

Notes: Columns (1) to (3) show OLS-regressions. Columns (4) and (5) are linear probability models. Dependent variable in Column (1) is the amount of credit intended for consumption purposes. Dependent variable I Column (2) is the amount of debt intended for productive investments. Dependent variable in Column (3) reports the ratio between debt and household assets and is transformed to a z-score. Dependent variables in Columns (4) and (5) are binary items reporting whether a respondent would be willing to take up a loan if he or she had no plans on how to use the money or whether a respondent can separate between good and bad reasons to take up a loan as stated in a hypothetical example. Items marked with an asterisk $(*)$ are later rescaled for the composition of the index such that positive values indicate better outcomes. All models include the lagged outcome at baseline and district-level fixed effects. + indicates that the currency denominated outcome (in Ugandan Shilling (UGX)) is winsorized at the 99th percentile. Standard errors (clustered at the market-level) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Financial services index – individual components

	(1) Ever purchased a formal insurance product	(2) Has ever used mobile money
Treatment A	0.006	0.063*
	(0.027)	(0.037)
Treatment B	0.050	0.035
	(0.035)	(0.045)
$A - B = 0$ (p-value)	0.159	0.514
R^2	0.036	0.194
Mean (SD) of y_t	0.099	0.442
in control group	(0.299)	(0.497)
Observations	1,133	1,136
Clusters	83	83
District FEs	yes	yes
$y_{(t-1)}$ covariate	yes	yes

Notes: Coefficients show results from linear probability models. All models include the lagged outcome at baseline and district-level fixed effects. Standard errors (clustered at the market-level) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A7: Impact of financial education treatments on individual-level income and household-level consumption

	(1) Income^{+}	(2) Consumption^{+}
Treatment A	-0,004	-20,299
	(37,321)	(35,919)
Treatment B	-8,664	-18,658
	(46,435)	(31,662)
$A - B = 0$ (p-value)	0.850	0.967
R^2	0.118	0.246
Mean (SD) of y_t in control group	389,893	720,541
	(599,475)	(489,268)
Observations	1,104	1,161
Clusters	83	83
District FEs	yes	Yes
$y_{(t-1)}$ covariate	yes	yes

Notes: Table shows OLS regression results of ANCOVA models. + indicates that the outcome (in Ugandan Shilling (UGX)) is winsorized at the 99th percentile. All models include the lagged outcome at baseline and district-level fixed effects. Standard errors (clustered at the market-level) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Descriptive statistics for intermediating index variables at baseline

		Full Sample	Control (C)	Treatment (A)		Treatment (B)	
	Obs.	Mean (SD)	Mean (SD)	Mean (SD)	Diff. C [SE]	Mean (SD)	Diff. C [SE]
(1) Fin. literacy score	1,291	0.03	0.00	0.10	0.10	0.00	0.00
		(0.98)	(1.00)	(0.96)	[0.11]	(0.99)	[0.11]
FL Item 1	1,291	0.62	0.63	0.61	-0.02	0.62	-0.01
		(0.48)	(0.48)	(0.49)	[0.48]	(0.49)	[0.05]
FL Item 2	1,291	0.60	0.59	0.60	0.01	0.62	0.03
		(0.49)	(0.49)	(0.49)	[0.05]	(0.49)	[0.05]
FL Item 3	1,291	0.47	0.43	0.51	0.08	0.47	0.04
		(0.50)	(0.50)	(0.50)	[0.03]	(0.50)	[0.04]
FL Item 4	1,291	0.44	0.45	0.50	0.05	0.39	-0.06
		(0.50)	(0.50)	(0.50)	[0.05]	(0.49)	[0.05]
FL Item 5	1,291	0.39	0.38	0.43	0.04	0.37	-0.01
		(0.49)	(0.49)	(0.50)	[0.04]	(0.48)	[0.05]
(2) Fin. confidence	1,259	-0.05	0.00	-0.04	-0.04	-0.11	-0.11
(z-score)		(0.99)	(1.00)	(0.98)	[0.11]	(1.01)	[0.10]
(3) Self-control	1,273	0.01	0.00	0.08	0.08	-0.04	-0.04
(z-score)		(1.01)	(1.00)	(1.02)	[0.08]	(0.99)	[0.08]
F-test of joint orthogonality (p-value)						0.32	
Observations						1,248	
Clusters						83	

Notes: For further information on the individual components of financial literacy score see Appendix C. Fin. confidence is an aggregate measure (equally weighted z-score index of its components and standardized to have a zero mean and a SD of one for the control group) of the following four items: (i) "In case you are dissatisfied with a financial service provider and you complain, do you think that the financial service provider is more powerful than you, and that the complaint will therefore not lead to anything?" (y/n). (ii) "I am confident enough to approach a bank and ask questions to learn more about their products." (disagree strongly – agree strongly). (iii) "I am confident that among a range of loans offered by different banks, I can choose the loan that best suits my specific needs" (disagree strongly – agree strongly). (iv) Which of the following statements best describes how you last chose a financial product? (a) I considered several products from different companies before making my decision. (b) I considered the various products from one company. (c) I didn't consider any other products at all (d) I looked around but there were no other products to consider. Self-control is assessed through the following survey question: "If you get money, do you tend to spend it too quickly? (a) often (b) sometimes (c) rarely (d) never. Standard errors (clustered at the market-level) are reported in square brackets. Tests are unadjusted for multiple hypothesis testing. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Analysis of effects on individual items in the FL-score (ANCOVA)

	(1) Item 1	(2) Item 2	(3) Item 3	(4) Item 4	(5) Item 5
Treatment A	-0.013	0.022	0.084**	-0.016	0.073*
	(0.029)	(0.032)	(0.036)	(0.039)	(0.040)
Treatment B	0.002	0.007	0.039	0.027	0.028
	(0.031)	(0.029)	(0.034)	(0.047)	(0.042)
$A - B = 0$ (p-value)	0.636	0.629	0.275	0.304	0.224
R^2	0.054	0.009	0.028	0.019	0.016
Mean (SD) of y_t in control group	0.653	0.672	0.444	0.436	0.512
	(0.477)	(0.470)	(0.497)	(0.496)	(0.500)
Observations	1,150	1,157	1,158	1,158	1,158
Clusters	83	83	83	83	83
District FEs	yes	Yes	yes	yes	yes
$y_{(t-1)}$ covariate	yes	yes	yes	yes	yes

Notes: Coefficients show results from linear probability models. For the results on the composite index reported in the main text, predicted scores from the 2PLM-IRT-model are used (cf. Appendix C). All models include the lagged outcome at baseline and district-level fixed effects. Standard errors (clustered at the market-level) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: 2SLS-regressions to probe the causal mechanisms for psychological variables

	1st stage			2nd stage		
	Panel A: Savings outcomes					
	(1) Fin. literacy	(2) Self-control	(3) Fin. confidence	(4) Savings index(z)	(5) Savings index(z)	(6) Savings index(z)
Treatment A	0.152*	0.152*	0.160**			
	(0.086)	(0.079)	(0.081)			
Treatment B	0.076	-0.058	0.193**			
	(0.076)	(0.092)	(0.083)			
Fin. literacy*				1.320		
				(0.880)		
Self-control*					0.864*	
					(0.524)	
Fin. confidence *						0.536
						(0.480)
Observations	1,162	1,156	1,026	1,162	1,157	1,027
	Panel B: Investment outcomes					
	Fin. literacy	Self-control	Fin. confidence	Investment index (z)	Investment index (z)	Investment index (z)
Treatment A	0.181*	0.134	0.142*			
	(0.094)	(0.086)	(0.080)			
Treatment B	0.097	-0.073	0.219***			
	(0.083)	(0.093)	(0.075)			
Fin. literacy*				1.479		
				(0.955)		
Self-control*					0.615	
					(0.672)	
Fin. confidence*						1.112
						(0.688)
Observations	1,037	1,033	920	1,037	1,033	920
	Panel B: Summary index					
	Fin. literacy	Self-control	Fin. confidence	Summary index (z)	Summary index (z)	Summary index (z)
Treatment A	0.185*	0.116	0.148*			
	(0.094)	(0.090)	(0.081)			
Treatment B	0.101	-0.056	0.218***			
	(0.086)	(0.100)	(0.076)			
Fin. literacy*				0.980		
				(0.664)		
Self-control*					0.790	
					(0.711)	
Fin. confidence*						0.633
						(0.560)
Observations	995	991	887	995	991	887
Clusters	83	83	83	83	83	83
District FEs	yes	yes	yes	yes	yes	yes
$y_{(t-1)}$ covariate	no	no	no	no	no	no

Notes: Results show two stage regressions. All models district-level fixed effects. Standard errors (clustered at the market-level) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Results from mediation analysis for psychological variables

	(1) Fin. literacy	(2) Self-control	(3) Fin. confidence	(4) Savings index	(5) Invest. index	(6) Summary index
Treatment A	0.151*	0.147*	0.159*			
	(0.086)	(0.079)	(0.082)			
Treatment B	0.076	-0.063	0.192**			
	(0.076)	(0.092)	(0.083)			
$A - B = 0$ (p-value)	0.375	0.019**	0.697			
Observations						
Panel A: Mediation effect of financial literacy						
ACME (Fin. literacy)				-0.001	0.004	0.010
[90% CI]				[-0.009, 0.005]	[-0.002, 0.014]	[-0.001, 0.024]
Direct Effect (Treatment A)				0.163	0.250	0.193
[90% CI]				[0.048, 0.276]	[0.103, 0.394]	[0.062, 0.322]
Total Effect (Treatment A)				0.162	0.254	0.203
[90% CI]				[0.047, 0.276]	[0.105, 0.399]	[0.066, 0.335]
| % of Total effect mediated |				0.8	1.71	4.96
[90% CI]				[0.05, 2.6]	[1.10, 4.09]	[2.96, 14.52]
Observations				1,161	1,007	895
Panel B: Mediation effect of self-control						
ACME (Self-control)				0.003	0.002	0.006
[90% CI]				[-0.004, 0.010]	[-0.004, 0.010]	[-0.003, 0.018]
Direct Effect (Treatment A)				0.159	0.251	0.199
[90% CI]				[0.041, 0.275]	[0.100, 0.399]	[0.067, 0.330]
Total Effect (Treatment A)				0.162	0.253	0.205
[90% CI]				[0.042, 0.279]	[0.101, 0.401]	[0.068, 0.337]
| % of Total effect mediated |				1.57	0.75	2.93
[90% CI]				[0.88, 5.26]	[0.47, 1.84]	[1.75, 8.57]
Observations				1,145	992	883
Panel C: Mediation effect of financial confidence						
ACME (Fin. confidence)				0.004	0.004	0.015
[90% CI]				[-0.003, 0.014]	[-0.004, 0.015]	[0.0009, 0.335]
Direct Effect (Treatment A)				0.141	0.225	0.184
[90% CI]				[0.015, 0.265]	[0.062, 0.384]	[0.032, 0.334]
Total Effect (Treatment A)				0.145	0.229	0.199
[90% CI]				[0.018, 0.269]	[0.065, 0.389]	[0.041, 0.351]
| % of Total effect mediated |				2.78	1.83	7.64
[90% CI]				[1.34, 9.77]	[1.06, 6.09]	[4.10, 26.93]
Observations				1,026	894	800
Mean (SD) of y_t	0.000	0.000	0.000	0.000	0.000	0.000
in control group	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Clusters	83	83	83	83	83	83
District FEs	yes	yes	yes	yes	yes	yes
$y_{(t-1)}$ covariate	yes	yes	yes	yes	yes	yes

Notes: Results from causal mediation analysis (Imai et al. 2011). All modes include the lagged dependent variable $y_{(t-1)}$ and district fixed effects. Standard errors (clustered at the market-level) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Impact on individual savings vs. group savings

	(1) Individual savings^{+}	(2) Group savings^{+}
Treatment A	101,028*	55,355
	(59,981)	(51,200)
Treatment B	77,154	-69,918
	(61,997)	(53,611)
$A - B = 0$ (p-value)	0.749	0.017**
R^2	0.036	0.034
Mean (SD) of y_t in control group	199,355	314,274
	(783,963)	(904,339)
Observations	1,162	1,162
Clusters	83	83
District FEs	yes	Yes
$y_{(t-1)}$ covariate	yes	yes

Notes: Table shows OLS regression results of ANCOVA models. + indicates that the outcome (in Ugandan Shilling (UGX)) is winsorized at the 99th percentile. All models include the lagged outcome at baseline and district-level fixed effects. Standard errors (clustered at the market-level) are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Heterogeneous treatment effects for subgroups

	(1) FL score (z)	(2) Budgeting index (z)	(3) Savings index (z)	(4) Borrowing index (z)	(5) Investment index (z)	(6) Fin. services index (z)	(7) Summary index (z)
Panel A: Male respondents							
Treatment A	0.0463	0.0285	0.186**	0.143**	0.200**	0.0613	0.167**
	(0.0995)	(0.0842)	(0.0753)	(0.0648)	(0.0830)	(0.0767)	(0.0830)
Treatment B	0.0331	0.00140	0.0416	0.0469	0.106	0.167*	0.0813
	(0.0862)	(0.0887)	(0.0717)	(0.0870)	(0.0964)	(0.0901)	(0.0969)
Male	-0.0973	0.306**	0.307**	0.404***	0.283*	0.360***	0.589***
	(0.101)	(0.147)	(0.125)	(0.122)	(0.146)	(0.109)	(0.137)
Treatment A × Male	0.387**	0.0796	-0.0463	-0.195	0.237	0.156	0.115
	(0.167)	(0.214)	(0.204)	(0.188)	(0.341)	(0.194)	(0.262)
Treatment B × Male	0.164	0.0582	-0.104	-0.403**	-0.0381	-0.0116	-0.140
	(0.155)	(0.182)	(0.243)	(0.194)	(0.212)	(0.189)	(0.193)
Obs.	1,138	1,091	1,137	1,087	987	1,088	878
R2	0.053	0.123	0.159	0.019	0.345	0.156	0.282
Panel B: Respondents with beyond primary education							
Treatment A	0.113	0.0656	0.186**	0.139**	0.237**	0.109	0.216**
	(0.0947)	(0.0821)	(0.0769)	(0.0654)	(0.0998)	(0.0819)	(0.0876)
Treatment B	0.146	0.0449	0.0638	-0.0424	0.128	0.104	0.0537
	(0.0901)	(0.0809)	(0.0693)	(0.0993)	(0.112)	(0.105)	(0.0898)
Educated	0.165	0.304***	0.335***	0.184*	0.0553	0.226**	0.313***
	(0.104)	(0.106)	(0.0978)	(0.109)	(0.113)	(0.0994)	(0.0911)
Treatment A × Educated	0.132	-0.0521	-0.0321	-0.105	0.229	0.0314	0.0175
	(0.147)	(0.147)	(0.164)	(0.139)	(0.215)	(0.167)	(0.191)
Treatment B × Educated	-0.276*	-0.106	-0.174	0.0358	0.169	0.198	-0.0256
	(0.163)	(0.173)	(0.156)	(0.187)	(0.267)	(0.138)	(0.158)
Obs.	1,162	1,114	1,161	1,110	1,007	1,111	895
R^2	0.058	0.120	0.166	0.014	0.150	0.146	0.254
Panel C: Respondents with above average financial literacy at baseline							
Treatment A	0.0893	-0.139	0.265**	0.162*	0.198*	0.0832	0.101
	(0.106)	(0.0938)	(0.105)	(0.0866)	(0.110)	(0.117)	(0.104)
Treatment B	0.0889	-0.121	0.161	-0.0672	0.0705	0.0991	-0.0252
	(0.0991)	(0.104)	(0.107)	(0.0981)	(0.109)	(0.137)	(0.102)
Fin. literate	0.0811	-0.122	0.261**	0.0450	0.0546	0.117	0.0457
	(0.131)	(0.0789)	(0.111)	(0.0954)	(0.0907)	(0.0924)	(0.0916)
Treatment A × Fin. literate	0.0837	0.335**	-0.198	-0.0995	0.145	0.0357	0.174
	(0.142)	(0.132)	(0.143)	(0.131)	(0.143)	(0.135)	(0.124)
Treatment B × Fin. literate	-0.0301	0.250*	-0.287**	0.0602	0.188	0.101	0.110
	(0.152)	(0.145)	(0.142)	(0.124)	(0.166)	(0.159)	(0.139)
Obs.	1,162	1,114	1,161	1,110	1,007	1,111	895
R^2	0.052	0.115	0.158	0.010	0.150	0.136	0.244
Clusters	83	83	83	83	83	83	83
District Fes	yes	yes	yes	yes	yes	yes	yes
$y_{(t-1)}$ controls	yes	yes	yes	yes	yes	yes	yes

Notes: Panel A shows results by gender, with "male" being an indicator variable for male respondents. Panel B shows interactions between treatments and education with "Educated" being an indicator identifying respondents with above primary education at baseline. Panel C reports interactions between the treatments and baseline financial literacy with "Fin. literate" being an indicator for a respondent scoring higher on the baseline financial literacy assessment than the average respondent in the full baseline sample. Standard errors, clustered at the market-level, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix B: Robustness checks

This appendix mainly deals with the issue of selective attrition and probes the sensitivity of our results to changes in the empirical strategy. In the following, we analyze determinants of attrition and examine the implications of attrition on the estimated treatment effects. Correcting for differential attrition according to treatment groups and districts through inverse-probability weighting of selection into endline yields similar results to an unweighted analysis that ignores attrition. More importantly, we estimate bounds for the treatment effects (A) and (B) under several plausible or extreme scenarios for outcomes of those who are not met at the follow-up survey. These analyses strongly support the findings of the analyses presented in the main text. Next, we show that simple OLS results using only the endline data are similar to the ANCOVA results. Finally, we show treatment effects on the measure of financial literacy using the more common raw sumscore rather than the predicted latent trait from the IRT model.

Selective attrition. As mentioned in Section 3.1, attrition varies between the treatment and control groups. At first glance, it appears that there is a significant effect of being assigned to treatment B on attrition. To investigate this issue, we first look at attrition per district: Just before our follow-up survey in April 2016, the districts of Bundibugyo and Kasese experienced violent outbreaks related to disputed local elections and tribal differences. These violent episodes lead to several deaths, military involvement and, according to local media, displacement of several thousand individuals from the district of Bundibugyo. 22 out of 83 markets are located in this district (and another 12 markets in the district of Kasese). These clusters make up roughly 39% of individuals in our entire sample at baseline. Because randomization at the cluster-level was not stratified by district, the distribution of treatment groups along these clusters is heterogeneous: In the case of Bundibugyo, only 23% of clusters belong to the control group. Thus, 77% of clusters have been assigned to financial education

treatment in the district. Out of these assigned to treatment, 52% have been assigned to treatment B which experiences the highest attrition. Thus, attrition appears may be driven by this exogenous shock.

To investigate determinants of attrition in a formal setting, we estimate a logit model of being absent at follow-up on the two treatment dummies and, alternatively include market or district fixed effects in this model.

<Table B1 about here>

Column 1 of Table B1 shows marginal effects after logit without accounting for district-level effects. This indicated that those assigned to treatment are 5 percentage points (marginally significant) more likely drop out of the survey than the control group. Turning to Column 2 which included market dummies, however, shows that this result disappears when accounting for these market-level events. Similarly, the coefficient decreases in size when accounting for district fixed-effects: Column 3 shows results including district dummies with the above-mentioned district of "Bundibugyo" being the omitted category. All of the other six districts are estimated to have lower attrition than Bundibugyo, further indicating that attrition is not endogenous but may be driven by this exogenous district and market-level shocks. To probe the sensitivity of our results to attrition, we next use the predicted probability to drop out from the study to adjust the regressions.

Inverse probability weighting. We weight all regressions with inverse probability of selection into endline tracking as estimated by a logit-model based in the specification in Column 3 of Table B1.

<Table B2 about here>

The results are near identical to the unweighted specification discussed in the main text.

Bounds estimates for the financial education treatments. To further address selective attrition, we estimate bounds for the worst-case scenarios for the treatment effects on the

financial literacy and savings. Similar to Karlan and Valdivia (2011) and Drexler et al (2014), we follow Horrowitz and Manski (2000) and Lee (2002), and impute plausible values for missing observations to estimate bounds for the treatment effect. For the extreme lower bound, we impute the largest value of the 1st percentile of each outcome variable in the observed distribution of y^* to the attriters in the treatment groups, and the smallest value of 99th percentile of the observed control-group distribution to the attriters in the control group. For the extreme upper bound, we impute the smallest value of the 99th percentile of each outcome variable in the observed treatment distributions to the attriters in the treatment groups, and the largest value of the 1st percentile of the observed control distribution to the attriters in the control group (cf. Karlan and Valdivia, 2011, p. 522). More plausible lower bounds are created by imputing missing values for the treatment groups A and B as their respective means for minus 0.1, 0.25, and 0.5 standard deviations of the observed distribution for the group. Missing values for the control group are imputed as the mean of the control group plus 0.1, 0.25, and 0.5 standard deviations, respectively. Plausible upper bounds for 0.1, 0.25, and 0.5 standard deviations are calculated analogously.

<Table B3 about here>

Panel A of Table B3 shows the sensitivity of the unadjusted treatment effect to the uncertainty around attrition on the savings index. Encouragingly, the main result is unaffected by considering even relatively extreme scenarios. Treatment A is always estimated to have larger (or less adverse in the extreme lower bound scenario) effects on savings except at the extreme upper bound where both treatments show large impacts. Additionally, differential treatment effects (test of equality of Treatment A and B) are rejected across the entire distribution of considered scenarios except the extreme upper bound. Statistical and economic significance for treatment A is maintained up to a lower bound of -0.25 standard deviations. In contrast, treatment B only reaches statistical significance at the extreme upper bound.

Panel B of Table B3 shows results on the investment index. Again, treatment A is estimated to have larger coefficients than treatment B in several of the considered scenarios. The effect of treatment A remains to be relatively large and statistically significant up to a lower bound scenario of -0.4 standard deviations. Differential effects between treatment A and B are present at all lower bound estimates. Thus, the results presented here give strong support to conclude that the main results are not sensitive to selective attrition.

Panel C of Table B3 shows results on the summary index. Again, differential effects are confirmed for all scenarios but the extreme upper bound. Effect sizes of treatment A are positive and significantly large from scenarios considering less than -0.25 SD and upwards.

OLS using only endline data. Table B4 shows results using OLS and only the endline data. This approach yields less power and precision, and thus, is not preferred (McKenzie, 2012). Nevertheless, it may be reassuring that results are similar to the ANCOVA estimation. Treatment effects of treatment A and B are estimated to be even larger, however, with less precision.

< Table B4 about here >

Standard FL-score. Table B5 shows regression-results using a standard sum score of correct answers as the measure of financial literacy. As the IRT-model clearly has better psychometric properties (cf. Appendix C), we do not prefer this procedure. However, it may be reassuring that effect sizes are similar and the difference in magnitude of the coefficients (unstandardized regression coefficient from OLS in Column1 and logged odds from ordered-logit and ordered-probit models) is similar to the result presented in the text, albeit with larger standard errors due to the relative imprecise measure of financial literacy.

< Table B5 about here >

Indices from PCA. Finally, we probe the sensitivity of results to relying on alternative methods of summary index creation.

< Table B6 about here >

Table B6 shows results for indices that were created with (polychoric) principal components analyses instead of equally weighted z-scores. Results are very similar to the method used in the main paper.

References in Appendix B

Drexler, A., Fischer, G., and Schoar, A. (2014). Keeping it simple: Financial literacy and rules of thumb. *American Economic Journal: Applied Economics*, 6(2): 1–31.

Horowitz, J. L. and Manski, C. F. (2000). Nonparametric analysis of randomized experiments with missing covariate and outcome data. *Journal of the American Statistical Association*, 95(449): 77–84.

Karlan, D. and Valdivia, M. (2011). Teaching entrepreneurship: Impact of business training on microfinance clients and institutions. *Review of Economics and Statistics*, 93(2): 510–527.

Lee, D.S. (2002). Trimming for bounds on treatment effects with missing outcomes. *NBER Technical Working Paper 277.*

McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, 99(2): 210–221.

Table B1: Determinants of attrition

	Binary for attrition		
	(1)	(2)	(3)
Treatment A	-0.016	0.073	-0.021
	(0.028)	(0.107)	(0.026)
Treatment B	0.050*	0.099	0.044*
	(0.028)	(0.102)	(0.025)
District: Kyegegwa			-0.068*
			(0.039)
District: Kyenjojo			-0.102***
			(0.036)
District: Kamwenge			-0.068*
			(0.039)
District: Kasese			-0.028
			(0.046)
District: Kabarole			-0.019
			(0.035)
District: Ntoroko			-0.083*
			(0.047)
Market dummies	No	Yes	No
District dummies	No	No	Yes
Pseudo R^2	0.015	0.106	0.04
n (Individuals)	1,291	1,291	1,291
n (Clusters)	83	83	83

Notes: Marginal effects from logit models. Robust Standard errors, (clustered at the market-level in (2) and (3)), in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B2: Main experimental results with inverse probability weighting to correct for selection into endline survey

	(1) Budget index (z)	(2) Savings index (z)	(3) Borrowing index (z)	(4) Investment index (z)	(5) Fin. services index (z)	(6) Summary index
Treatment A	0.042	0.163**	0.108*	0.254***	0.108	0.205**
	(0.086)	(0.071)	(0.057)	(0.090)	(0.077)	(0.082)
Treatment B	0.005	0.019	-0.044	0.112	0.157	0.035
	(0.086)	(0.076)	(0.075)	(0.093)	(0.095)	(0.088)
$A - B = 0$ (p-value)	0.679	0.092*	0.039**	0.322	0.615	0.078*
R^2	0.109	0.152	0.009	0.147	0.130	0.239
Mean (SD) of y_t in	0.000	0.000	0.000	0.000	0.000	0.000
control group	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Observations	1,114	1,161	1,110	1,007	1,111	895
Clusters	83	83	83	83	83	83
District FEs	yes	yes	yes	yes	yes	yes
$y_{(t-1)}$ covariate	yes	yes	yes	yes	yes	yes
IPW	yes	yes	yes	yes	yes	yes

Notes: WLS regressions with inverse probability weights based on the estimation in Table B1, Column 3. Standard errors, clustered at the market-level, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B3: Bounds estimates for treatments A and B

	(1) Extreme lower bound	(2) -0.50 SD	(3) -0.25 SD	(4) -0.10 SD	(5) Unadjusted treatment effect	(6) + 0.10 SD	(7) +0.25 SD	(8) +0.50 SD	(9) Extreme upper bound
Panel A: Impact on the savings index (z)									
Treatment A	-0.297***	0.0848	0.122*	0.145**	0.164**	0.175***	0.197***	0.235***	0.597***
	(0.0959)	(0.0655)	(0.0653)	(0.0655)	(0.071)	(0.0661)	(0.0668)	(0.0685)	(0.124)
Treatment B	-0.576***	-0.102	-0.0509	-0.0200	0.013	0.0211	0.0519	0.103	0.690***
	(0.107)	(0.0699)	(0.0686)	(0.0682)	(0.077)	(0.0681)	(0.0685)	(0.0697)	(0.138)
$A - B = 0$ (p)	0.006***	0.019**	0.027**	0.034**	0.077*	0.048**	0.062*	0.096*	0.57
Obs.	1,291	1,291	1,291	1,291	1,161	1,291	1,291	1,291	1,291
R^2	0.111	0.136	0.140	0.141	0.152	0.144	0.146	0.148	0.130
Panel B: Impact on the investment index (z)									
Treatment A	-0.429***	0.129	0.208***	0.255***	0.258***	0.319***	0.367***	0.447***	1.115***
	(0.118)	(0.0800)	(0.0778)	(0.0769)	(0.091)	(0.0762)	(0.0761)	(0.0768)	(0.126)
Treatment B	-0.646***	-0.0304	0.0604	0.115	0.113	0.188**	0.243***	0.334***	0.918***
	(0.123)	(0.0793)	(0.0772)	(0.0767)	(0.095)	(0.0768)	(0.0774)	(0.0797)	(0.137)
$A - B = 0$ (p)	0.021**	0.061*	0.080*	0.095*	0.150	0.122	0.148	0.201	0.254
Obs.	1,291	1,291	1,291	1,291	1,007	1,291	1,291	1,291	1,291
R^2	0.284	0.261	0.261	0.261	0.326	0.262	0.263	0.265	0.255
Panel B: Impact on the summary index (z)									
Treatment A	-0.699***	0.0210	0.108	0.160**	0.207**	0.229***	0.281***	0.369***	1.254***
	(0.108)	(0.0672)	(0.0654)	(0.0649)	(0.083)	(0.0650)	(0.0656)	(0.0674)	(0.113)
Treatment B	-1.059***	-0.204***	-0.0970	-0.0327	0.032	0.0536	0.119*	0.230***	1.165***
	(0.127)	(0.0690)	(0.0655)	(0.0644)	(0.090)	(0.0642)	(0.0651)	(0.0682)	(0.134)
$A - B = 0$ (p)	0.001***	0.002***	0.004***	0.006***	0.073*	0.011**	0.021**	0.056*	0.537
Obs.	1,291	1,291	1,291	1,291	895	1,291	1,291	1,291	1,291
R^2	0.194	0.187	0.187	0.187	0.237	0.187	0.186	0.184	0.176
Clusters	83	83	83	83	83	83	83	83	83
District FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$y_{(t-1)}$ covar.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors, clustered at the market-level, in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B4: Main experimental results (OLS using only endline data)

	(1) Budget index (z)	(2) Savings index (z)	(3) Borrowing index (z)	(4) Invest-ments (z)	(5) Fin. services index (z)	(6) Summary index
Treatment A	0.0311	0.200**	0.100*	0.265**	0.0939	0.182*
	(0.0942)	(0.0781)	(0.0548)	(0.124)	(0.0932)	(0.103)
Treatment B	0.0124	0.0482	-0.0211	0.194	0.186	0.0816
	(0.0950)	(0.0871)	(0.0685)	(0.129)	(0.116)	(0.115)
$A - B = 0$ (p-value)	0.843	0.127	0.073*	0.611	0.424	0.394
R^2	0.047	0.025	0.007	0.077	0.032	0.038
Mean (SD) of y_t in control group	0.000	0.000	0.000	0.000	0.000	0.000
	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)	(1.000)
Observations	1,145	1,162	1,139	1,037	1,153	995
Clusters	83	83	83	83	83	83
District FEs	yes	yes	yes	yes	yes	yes
$y_{(t-1)}$ covariate	no	no	no	no	no	no

Notes: Standard errors, clustered at the market-level, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B5: Standard financial knowledge score

	(1) OLS	(2) Ordered logit	(3) Ordered probit
Treatment A	0.155	0.099	0.165
	(0.110)	(0.083)	(0.141)
Treatment B	0.089	0.062	0.109
	(0.101)	(0.078)	(0.126)
$A - B = 0$ (p-value)	0.55	0.67	0.70
R^2	0.028		
Pseudo R^2		0.016	0.015
Mean (SD) of y_t in control group	2.491	2.491	2.491
	(1.531)	(1.531)	(1.531)
Observations	1,162	1,162	1,162
Clusters	83	83	83
District FEs	yes	yes	yes
$y_{(t-1)}$ covariate	yes	yes	yes

Notes: Dependent variable is the financial knowledge sumscore (cf. Appendix C) at the time of the endline survey. (1) shows OLS results. (2) shows results from an ordered-logit model. (3) shows results from an ordered-probit model. Standard errors, clustered at the market-level, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B6: Alternative methods of index creation

	(1) Budget index (PCA)	(2) Savings index (PCA)	(3) Borrowing index (PCA)	(4) Invest-ments (PCA)	(5) Fin. services index (PCA)	(6) Summary index of indices (PCA)
Treatment A	0.036	0.152**	0.044	0.238***	0.111	0.141*
	(0.084)	(0.071)	(0.077)	(0.083)	(0.076)	(0.077)
Treatment B	0.046	-0.010	0.005	0.090	0.147	0.008
	(0.084)	(0.075)	(0.086)	(0.087)	(0.094)	(0.082)
$A - B = 0$ (p-value)	0.933	0.052*	0.637	0.118	0.703	0.148
R^2	0.129	0.176	0.132	0.424	0.133	0.428
Mean (SD) of y_t in control group	0.000 (1.000)	0.000 (1.000)	0.000 (1.000)	0.000 (1.000)	0.000 (1.000)	0.000 (1.000)
Observations	1,114	1,161	1,110	1,007	1,136	995
Clusters	83	83	83	83	83	83
District FEs	yes	yes	yes	yes	yes	yes
$y_{(t-1)}$ covariate	yes	yes	yes	yes	yes	yes

Notes: Indices created via (polychoric) principal components analyses instead of equally weighted z-scores. Standard errors, clustered at the market-level, in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix C: Measuring financial literacy

Financial *literacy* is typically measured through a set of survey-items that measure financial *knowledge* (cf. Huston 2010, p. 303). In a second step, the scores on these dichotomous items (true/false) are summed up to generate a scale of financial knowledge to serve as a proxy for the latent trait "financial literacy". There are standard questions used for the assessment of financial literacy in developed economies (cf. Knoll and Houts 2012; Lusardi and Mitchell 2014, p. 10) and Cole et al. (2011) were the first to translate and adapt these questions to a developing economy context for their study of financial education and the demand for financial products in Indonesia. While these and similar survey items have been widely used in the literature on financial education in developing economies (e.g. Carpena et al. 2011; Sayinzoga et al. 2016), their psychometric properties have not been adequately assessed. This may be surprising since knowledge development is seen to be one of the primary goals of these types of financial education interventions (cf. Skimmyhorn 2016) and the need for a valid measure of financial literacy appears evident.

Thus, we study the psychometric properties of these items and propose an alternative approach to generate valid financial literacy scores for individual respondents: Following Knoll and Houts' (2012) discussion of items used in the assessment of financial knowledge in large-scale household surveys, we use "item response theory" (IRT) to create a valid and reliable scale of financial literacy. IRT is a family of models widely used in educational and psychological measurement (see Rasch 1960 and Lord 1980 for key contributions to this literature). A popular model used to design psychological measurement-scales is the two-parameter logistic model (2PLM) (cf. Birnbaum 1968). Here, the probability of an individual j to solve the item i is defined as

$$P(x_{ij} = 1 \mid \theta_j) = \frac{exp\{a_i(\theta_j - b_i)\}}{1+exp\{a_i(\theta_j - b_i)\}} \quad (1)$$

with

$$\theta_j \sim N(0,1) \quad (2)$$

where a_i and, b_i are the discrimination and difficulty parameters of item i respectively, and θ_j representing the latent trait (e.g. financial literacy) of individual j. Thus, the discrimination parameter a_i describes how well item i discriminates people of lower and higher ability (θ_j), and b_i corresponds to the point on the latent scale (θ) where $P(x_i = 1 | \theta) = 0.5$ (i.e. the point on the latent scale where an individual has a greater possibility to score the item than indicated by chance). Since we assume θ to have a mean of zero by definition, an item i is relatively easy to solve if $b_i < 0$, and an item i is relatively hard to solve if $b_i > 0$. This model requires the assumption of local independence among items (solving an item must not be conditional upon solving another item) and θ to be unidimensional. While local independence is given by the design of the items and implementation into the survey instruments, we tested the assumption of unidimensionality through a polychloric factor analysis. Indeed, only one factor is estimated with an eigenvalue > 1 and, thus, the assumption of a unidimensional θ appears to be met by the items included in the scale (cf. Figure C1 in Appendix C). To arrive at parameter estimates for a_i and, b_i, as well as to predict θ_j for all respondents in the dataset, we estimate equation (1) with five binary items that form the final financial knowledge scale. Standard errors are clustered at the level of randomization (markets).

Table C1 shows the exact wording, discrimination and difficulty for the final set of five items. The items are ordered by their ability to discriminate in ascending order. Thus, item 1 is the least discriminating ($a_1 = 0.981$) and item 5 is the most discriminating ($a_5 = 1.629$). The difficulty ranges from -0.569 (item 2) to 0.463 (item 3). A graphic representation of these item characteristics is depicted in Figure C2 which shows the trace line for each item included in the scale. Item 3 is most difficult (furthest to the right) while item 2 appears to be easiest. Regarding the discrimination, it is obvious that the trace line for item 5 has the

steepest slope while the slope of item 1 is most gradual. Another way to represent the features of each item is to plot the item information functions. Figure C3 shows the item information functions for each item. Figure C4 relates the latent trait back to the sumscores of items solved: Using the critical values of the z-distribution (-1.96 and 1.96) it appears that 95% of randomly selected individuals would solve between 0.451 and 4.51 items with a respondent of average ability (θ=0) scoring 2.56 (two or three) out of five items. Turning to the overall reliability of the scale, Figure C5 shows that the scale is most precise at the mean of θ with smallest standard errors close between -0.1 and 0. Finally, we standardize the scale to have a mean of zero and a standard deviation of one for the control group: Figure C6 shows the full distribution of the estimated ability (θ) for all individuals in our dataset at baseline.

References in Appendix C

Birnbaum, A. 1968. Some latent trait models and their use in inferring an examinee's ability. *Statistical Theories of Mental Test Scores*, F. M. Lord and M. R. Novick (ed.), Reading, MA: Addison–Wesley, 395–479.

Carpena, F., Cole, S., Shapiro, J., and Zia, B. (2011). Unpacking the causal chain of financial literacy. *World Bank Policy Research Working Paper 5798.*

Cole, S., Sampson, T., and Zia, B. (2011). Prices or knowledge? What drives demand for financial services in emerging markets? *Journal of Finance*, 66(6): 1933–1967.

Knoll, M. A. Z. and Houts, C. R. (2012). The financial knowledge scale: An application of item response theory to the assessment of financial literacy. *Journal of Consumer Affairs,* 46(3):381–410.

Lord, F. M. (1980). *Applications of Item Response Theory to Practical Testing Problems.* Mahwah, NJ: Lawrence Erlbaum.

Lusardi, A. and Mitchell, O. S. (2014). The Economic Importance of Financial Literacy: Theory and Evidence. *Journal of Economic Literature* 52 (1), pp. 5–44.

Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests.* Copenhagen: Danish Institute of Educational Research.

Sayinzoga, A., Bulte, E. H., and Lensink, R. (2016). Financial literacy and financial behaviour: Experimental evidence from rural Rwanda. *Economic Journal*, 126(594): 1571–1599.

Skimmyhorn, W. L., Davies, E. R., Mun, D., and Mitchell, B. (2016). Assessing financial education methods: Principles vs. rules-of-thumb approaches. *Journal of Economic Education*, 47(3): 193–2.

Table C1: Items and their psychometric properties of the final FL-scale

Item	Topic	Question and response options	a_i (SE)	b_i (SE)	% correct baseline	% correct endline
1	Diversification	Is it riskier to plant…? A) multiple crops or **B) one crop** Y) Don`t know Z) Refuse to Answer	1.092 (0.166)	-0.569 (0.104)	62.28	64.52
2	Inflation	If you have UGX. 100,000 in a savings account earning 1% interest per annum, and prices for goods and services rise 2% over a 1-year period, can you buy A) more than, **B) less than,** C) or the same amount of goods in 1 year as you could today, with the money in the account?" Y) Don`t know Z) Refuse to Answer	1.692 (0.258)	-0.375 (0.085)	60.50	67.16
3	Interest rate (loan)	If you were offered a loan with 5% monthly interest rate and a loan with 20% annual interest rate, which loan would offer better value? A)5% monthly interest rate **B) 20% annual interest rate** Y) Don`t know Z) Refuse to Answer	1.346 (0.149)	0.130 (0.061)	46.79	48.88
4	Interest rate (loan)	Suppose you need to borrow 500,000 UGX. Two people offer you a loan. Which loan represents a better deal for you? A) One loan requires you to pay back 600,000 UGX in 1 month. **B) The second loan requires you to pay back in 1 month 500,000 UGX plus 15% interest.** Y) Don`t know Z) Refuse to Answer	0.981 (0.152)	0.274 (0.107)	44.46	43.78
5	Compound interest	Suppose you borrow 100,000 UGX at an interest rate of 2% per month, with no repayment for 3 months. After 3 months, do you owe A) less than. 102,000 UGX, B) exactly. 102,000 UGX, **C) or more than 102,000 UGX?** Y) Don`t know Z) Refuse to Answer	1.218 (0.146)	0.463 (0.087)	39.27	52.94

Notes: N=1,291 at baseline. Results from fitting a 2PLM to the 5 items. Standard errors are clustered at the market-level. Items are coded to be binary. The correct response is coded to be equal to one. Wrong answers and response options Y) and Z) are coded to be equal to zero.

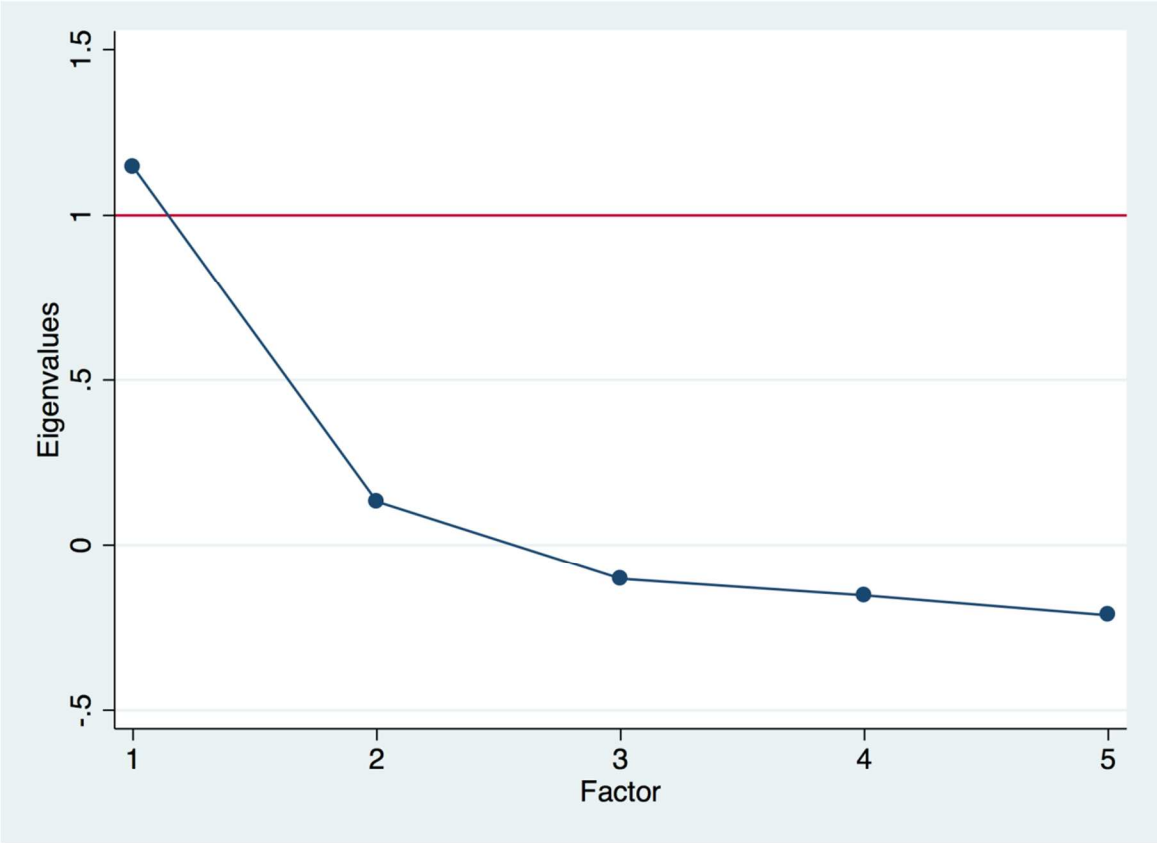

Figure C1: Screeplot of eigenvalues by factor after factor analysis

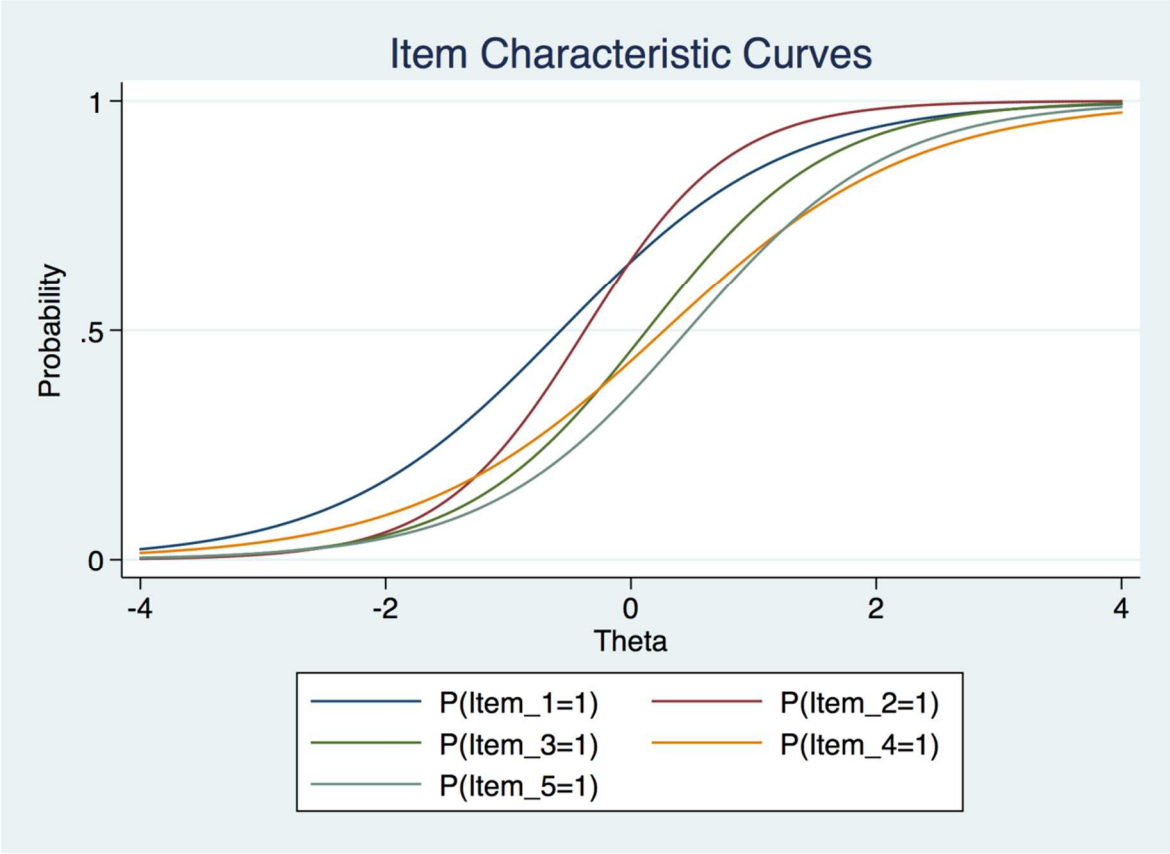

Figure C2: Item characteristic curves for the 2PLM financial literacy scale

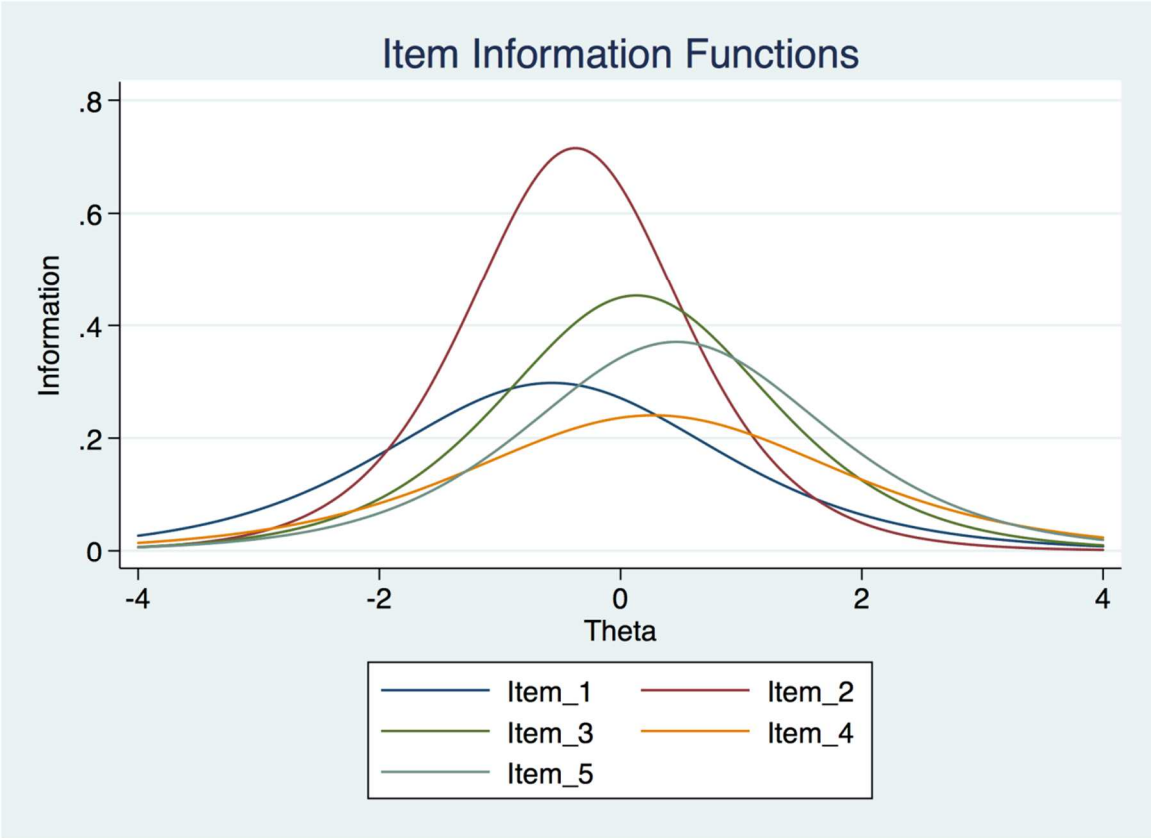

Figure C3: Item information functions for the 2PLM financial literacy scale

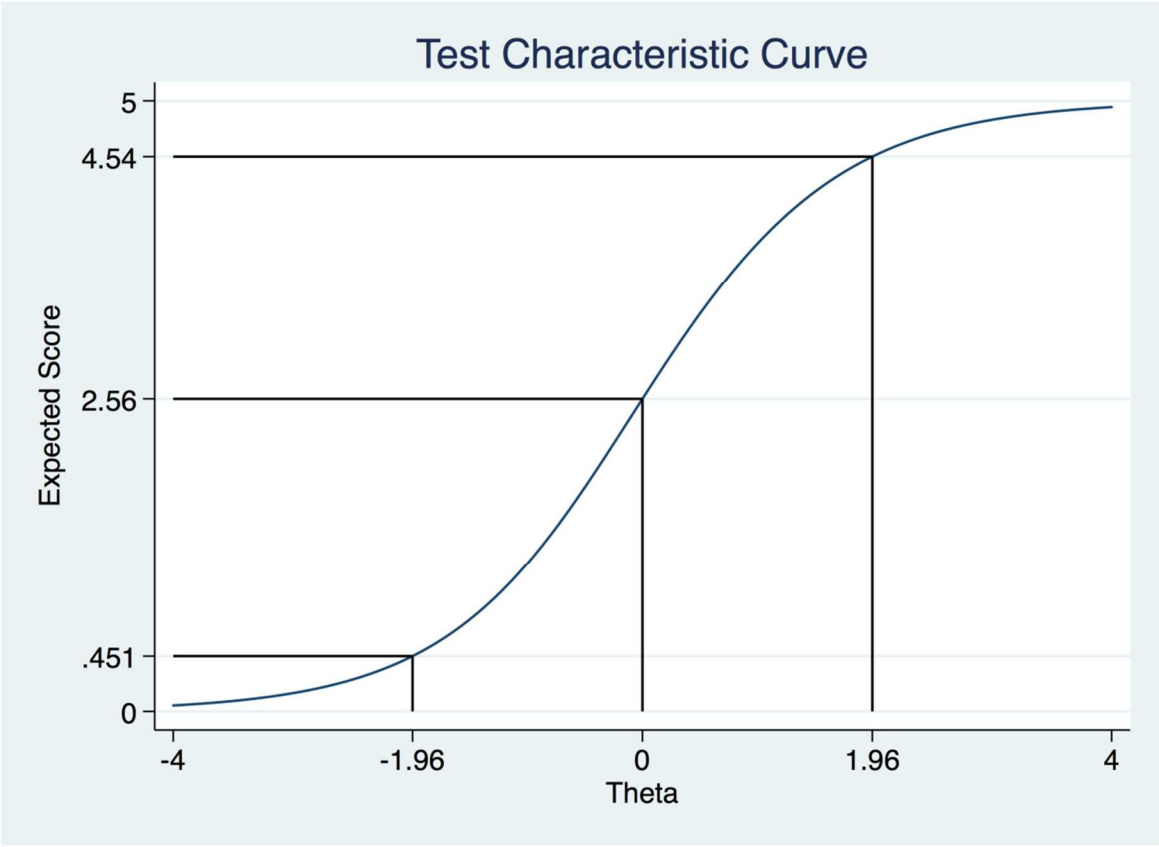

Figure C4: Test characteristic curve for the 2PLM financial literacy scale

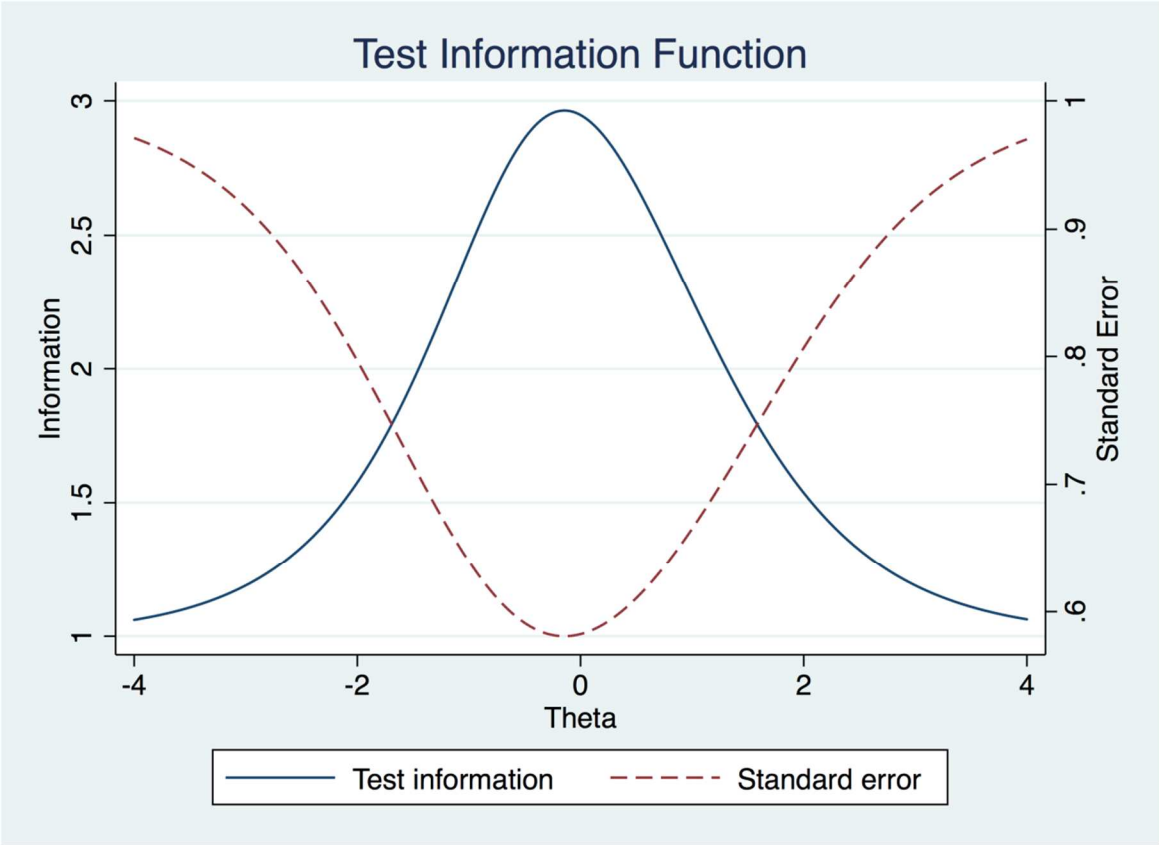

Figure C5: Test information function for the financial literacy scale

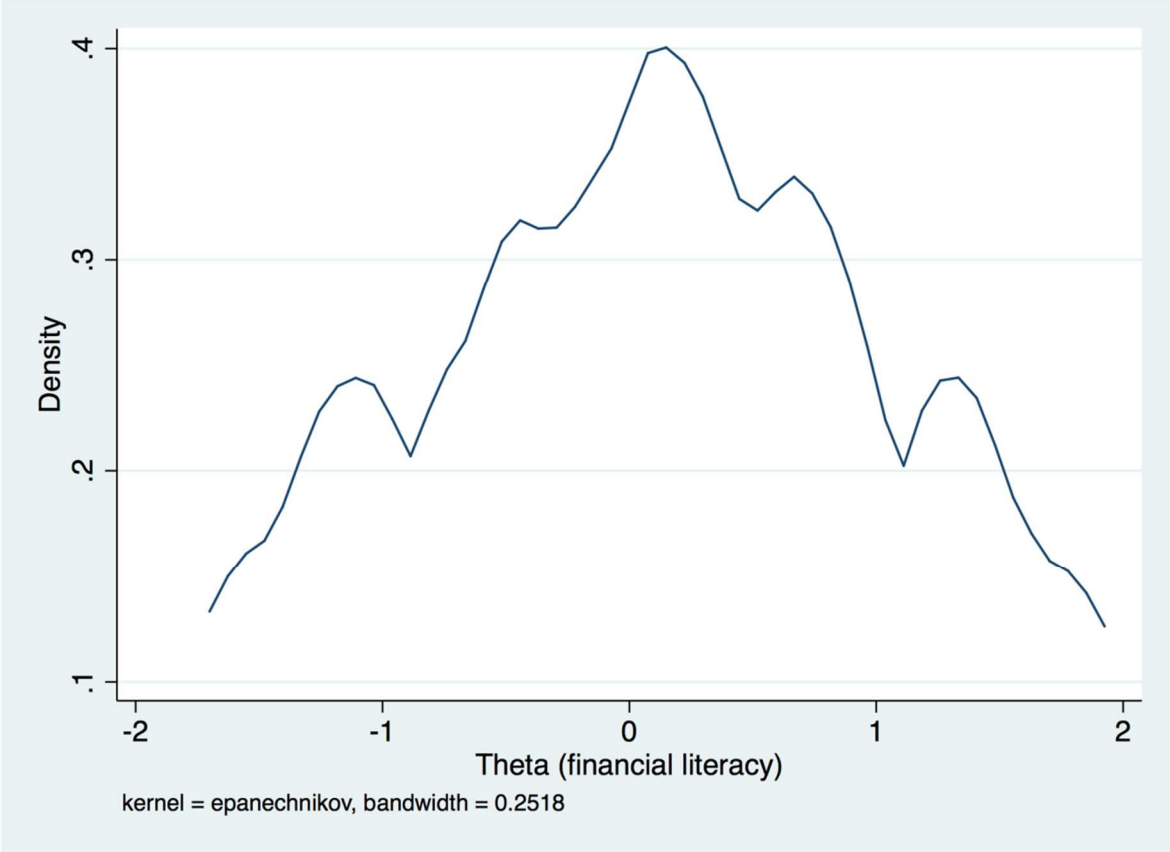

Figure C6: Distribution of standardized financial literacy IRT-scores at baseline