Active learning fosters financial behavior: Evidence from rural Uganda

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 Western Uganda. The treatments contrast "active learning" with "traditional lecturing" within standardized lesson-plans. We find that active learning has a clear positive impact on savings and investment outcomes, but weaker effects on financial literacy and debt-related outcomes, in contrast to precisely estimated insignificant impacts of lecturing. The active learning intervention is superior as it works via three mechanisms, i.e. increased financial literacy, self-control, and financial confidence, while lecturing only affects financial confidence.

JEL-Classification: O16 (savings), D14 (personal finance), I21 (analysis of education)
Keywords: financial literacy, financial behavior, RCT, active learning, lecturing, training method

April 09, 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 highly 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 colleagues and seminar participants in Berlin, Fort Portal, Hamburg, Kampala, Kiel, and Vienna, in particular Florian Artinger, Toman Barsbai, Nathan Fiala, Antonia Grohmann, Ralph Hertwig, Katharina Lehmann-Uschner, Helke Seitz, and Joachim Winter. Financial support by DFG through CRC TRR 190 is gratefully acknowledged. This field-experiment was pre-registered at the AEA RCT Registry on October 10th, 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, German Institute for Economic Research (DIW Berlin) and Humboldt-University of Berlin, 10108 Berlin, Germany; lmenkhoff@diw.de

Active learning fosters financial behavior: Evidence from rural Uganda

1 Introduction

Financial education programs are expected to foster the financial knowledge and behavior of individuals in order to achieve financial inclusion and to promote financial stability. However, the potential impact of these programs is debated in the literature (see Hastings et al., 2013; Lusardi and Mitchell, 2014; van Roij et al., 2014; Brown et al., 2016) because early experiments show relatively muted effects on financial outcomes (e.g. Cole et al., 2011). While more recent evidence clearly tilts toward the intended effectiveness (Kaiser and Menkhoff, 2017), the causal effects of these interventions are economically small on average (cf. Fernandes et al., 2014). Moreover, effects seem to be highly heterogeneous across multiple dimensions (Miller et al., 2015; Kaiser and Menkhoff, 2017). As even large-scale financial education programs appear to have only modest effects on financial behavior (cf. Bruhn et al., 2014, 2016), policymakers and researchers are evaluating alternatives and complements to the typically pervasive lecture-based classroom-programs (cf. Drexler et al., 2014; Carpena et al., 2011, 2017; Campos et al., 2017).

Many of these efforts to increase the effectiveness of financial education involve either altering the content that is taught (Drexler et al., 2014) or introducing personalized elements (Carpena et al. 2017) with a lower student to teacher ratio which is much more costly and difficult to operate at scale. In contrast, we propose using the existing curricula and keeping the participant to teacher ratio unchanged, only changing *how* the content is taught. To the best of our knowledge, despite the fact that the advantages of "active learning" over "traditional lecturing" is demonstrated in other domains, such as science instruction (e.g., Freeman et al., 2014), no study experimentally examines the effects of how financial education content is taught. Interactive teaching methods that engage and involve participants

are expected to yield greater impacts than the exposition centered financial education that currently dominates many programs. Generally, the importance of how information is conveyed is demonstrated to affect behavioral responses in the context of retirement-savings (cf. Saez, 2009).

Thus, we conduct a large-scale cluster-randomized field experiment to contribute to this literature by examining the treatment effects of alternative approaches to financial education. The financial education treatments contrast an *active learning* approach with a *traditional lecturing* approach, while teachers, content, and intensity of training are kept constant. The randomized field-experiment is conducted in rural Western Uganda with 1,291 market vendors.

Two results emerge: First, and our main finding, is that the group allocated to the *active learning* program experiences a significant increase in four out of six 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 financial literacy test-scores and an index of debt-related behavior. Thus, this approach generates larger treatment-effects than the insignificant, positive treatment effects realized by *traditional lecturing*. Differential treatment effects, however, can only be confirmed for outcomes in the savings domain.

Second, we provide tentative evidence that the relatively strong impact of the active learning approach on savings outcomes is not primarily caused by an increase in financial literacy but is impacted through the channel of increased self-control. Investment outcomes, in contrast, appear to be impacted through increased financial literacy and changes in financial confidence.

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 financial education interventions. Drexler et al. (2014) vary the content of the financial education curriculum and show the differential impacts of two different 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"), especially for low-skilled individuals and individuals with low baseline financial literacy and motivation. Skimmyhorn et al. (2016) replicate this type of experiment, but do not find evidence to support differential effects regarding the two different types of curricula, probably because their sample of respondents is relatively well educated. Lusardi et al. (2017) 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.

Meanwhile, Berg and Zia (2017) show that financial education interventions that primarily target non-cognitive channels can significantly impact financial behavior. Targeting entrepreneurs, Campos et al. (2017) show that a psychology-based training program that teaches 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 in a multiple-treatment experiment that complementing traditional instruction with personalized elements, like 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 to cope with unexpected shocks (cf. Carpena et al., 2017, p.2).

Much of this evidence suggests that interventions relying on traditional lecturing appear to have small or zero effects. Thus, alternative approaches to lecture-based education appear 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 fostering financial behavior and that these programs must be complemented with individualized elements. 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. Exposition centered teaching methods, however, yield smaller treatment effects and do not impact financial behavior or financial outcomes.

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 which we evaluate (Section 2.1), the experimental design (Section 2.2) and also the empirical strategy we follow here (Section 2.3).

2.1 Financial education interventions

As one step toward promoting financial inclusion, the central bank of Uganda (*BoU*) has 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. *BoU* has 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* has developed educational formats for the target population of the rural self-employed. These programs teach how to create a written household budget and to keep track of household's financial inflows and outflows, 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 benefits of using financial services provided by regulated financial institutions. Thus, these trainings promote 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 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

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

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)), 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 trainers 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 trainers and implementation is not heterogeneous across experimental sites.

2.2 Experimental design

We organize a cluster-randomized experiment to study the differential impact of the two financial education interventions on financial literacy and financial behavior. The main outcomes of interest are changes in *financial literacy* (see Appendix C for the development of a psychometrically sound measure of financial literacy) and changes in *financial behaviors* within five sub-domains addressed by the training, 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 cluster-level, and 83 rural marketplaces in seven districts of the Rwenzori Region in Western Uganda form 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 when 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 (cf. Muller, 2015). 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. 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 (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) or (B = traditional lecturing) to the control group (C= no treatment) at the time of the follow-up survey. Because selection into treatment is done through randomization and the later Section 3.2 indicates balanced groups at baseline, 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 (McKenzie, 2012):

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

Here, $y_{ic(t)}$ denotes the outcome variable (measure of financial literacy or 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 being 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 (see Appendix B).

Finally, to avoid problems inherent to testing multiple hypotheses (type-I-error inflation), we aggregate multiple related outcomes into index-measures of outcomes 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).

To probe the potential causal mechanisms leading to changes in financial behavior we estimate the following two stage regressions:

$$Z_{ic(t)} = \alpha + \beta_1 A_c^T + \beta_2 B_c^T + \theta_d + \varepsilon_{ict} \quad (2)$$

$$y_{ic(t)} = \alpha + \beta_1 Z_{ic(t)}^* + \theta_d + \varepsilon_{ict} \quad (3)$$

Equation (2) shows the first stage of the regressions. Here, $Z_{ic(t)}$ is a measure of intermediate outcomes in a causal chain that may have an impact on downstream behavior. We test three potential mechanisms: $Z_{ic(t)} = financial\ literacy$, $Z_{ic(t)} = self\ control$, $Z_{ic(t)} = financial\ confidence$ (see Section 6). The first stage estimates the causal effect of financial education treatments A and B on these intermediate outcomes. Thus, in the second stage, $Z_{ic(t)}$is instrumented by the two treatment dummies, and we examine whether predicted levels of the intermediate result ($Z_{ic(t)}^*$) explain variation in financial behavior ($y_{ic(t)}$).

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 enumerators conducting the face to face interviews in the local languages were trained extensively prior to the field-activities.

3.1 Response rates

After baseline-survey and 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) with 14.25 percent. While this may indicate selective attrition, we show in Appendix B that this does not change 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.

< 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 Blattmann et al., 2015). The mean (winsorized) monthly individual income is around 220,000 UGX (about 60 USD). The mean (winsorized) monthly household consumption value is about 593,000 UGX. Household consumption is higher than hypothetical added individual incomes because of subsistence farming, being reported by 83 percent of the sample.

Panel B reports variables at the respondent-level. 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. (2010, 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 (see descriptive statistics for individual index components in Table A7 in Appendix A).

3.3 Randomization balance

Causal inference within the estimation framework introduced in Section 2.4 rests on the random assignment of cluster 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 few small differences exist: In group A, a smaller share of the households seems to own the dwelling they live in. However, this difference is only marginally significant.

Second, the treatment groups are estimated to be slightly younger, on average, than the control group. However, again, this difference is only significant at the 10-percent-level. Thus, the only difference that is estimated to be statistically significant at the 5-percent-level

is the average number of years the respondents work as a market-vendor. On average, respondents in the control group have been vending for about 2 years longer than their counterparts in group B. These minor imbalances are what can be expected to occur by chance. *Panel C* shows descriptive statistics for these outcome indices of financial literacy and financial behavior at baseline. Again, no statistically significant differences exist between the three experimental groups (see Table A7 in Appendix A for descriptive statistics and randomization balance for single index components). Reassuringly, a joint test of orthogonality (where an indicator of being assigned to any treatment is regressed on all observable baseline covariates and the index-measures for outcomes at baseline) results in low explanatory power and a p-value of 0.3. Thus, orthogonality and balance seem to be met in this sample of 800 respondents (491 respondents have missing values on at least one of the covariates included in this fully specified regression covering all variables in Panels A, B, and C.)

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 trainings (A) and (B) on the financial literacy scores and on five domains of financial behavior. 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 >

Financial literacy score. Starting with financial literacy, both treatments show a positive coefficient, indicating that training tends to improve the degree of financial literacy. However, the coefficient sizes are moderate and thus only treatment A, i.e. active learning, has a statistically marginally significant effect relative to the control group. Due to small effect sizes and large standard errors the difference between both treatments is statistically insignificant.

While this result is not very strong in absolute terms, it looks more promising if we consider the circumstances. First, these trainings aim for changing behavior and not primarily for learning about numeracy-related financial knowledge, and second, the training is quite brief, lasting for two hours and covering five domains of behavior, so that the intensity for each subject matter is rather limited. Additionally, the ICC for outcomes can be relatively high with up to 0.003. Thus, the experiment is powered to detect a minimum effect size of 0.18 standard deviation units with small standard errors. Given these conditions, a marginal significant effect size of 0.134 for treatment A may be seen as promising, since we cannot rule out false negatives due to relatively low power for this outcome.

Indices of financial behavior. The five domains of financial behavior being 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 the Appendix A and selectively discussed in the next Section 4.2. Overall, coefficients are consistently positive, indicating that also financial behavior can be improved into the desired direction.

Among these domains of financial behavior, the overall effects are strongest for the *investment index* with effect sizes being large (more than a quarter of a standard deviation) and statistically significant at the 1-percent-level (column 5). 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 0.16 standard deviation units. Because the experiment is powered to detect effect sizes as low as 0.15 standard deviation units for outcomes with low ICC we conclude that this is evidence of precisely estimated zero-effects of treatment B.

Similarly, the effect on the *savings behavior index* (column 3) is sizeable for treatment A (about 0.16 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.

Somewhat surprisingly, relative to low expectations from the literature on the effectiveness of financial education to change *debt behavior* (Miller et al., 2015; Kaiser and Menkhoff, 2017), we find a small but marginally significant effect for treatment A (column 4). As the coefficient sign of treatment B is even negative, the difference between both treatments is marginally significant, as well. Regarding training effects on the *budgeting index* (column 2), the effect sizes are positive but statistically insignificant from zero. The coefficient levels are higher for the *financial services index* (column 6); 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.

Overall, we see that financial education tends to have desired effects, but effect sizes remain small and statistically insignificant unless the active learning program is implemented. Active learning results in a significant effect on a measure of financial literacy and three out of five financial behaviors addressed by the training. Lecturing, on the other hand has no effect on financial literacy or any of the addressed financial behaviors. However, meaningful differential treatment effects can only be confirmed for outcomes in 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 (financial literacy, budgeting behavior, debt behavior, and financial services behavior) are discussed in more detail in the Appendix A.

Components of the savings index. As shown above, active learning has a positive effect on 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. “any savings”, “total savings” and “net savings”. These results are shown in the first three columns of Table 5.

<Table 5 about here >

Given that 87.9 percent of the control group report any savings, the increase due to financial education has to be modest by definition and is in effect just 3.6 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 2.5 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 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 lower ex-ante numbers 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 109,000 UGX in total savings which amounts to a treatment effect of approximately 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 (even with a negative sign) and is precisely 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 even 145,000 UGX (38 percent) relative to the control group (column 3). Thus, this effect appears quite strong and statistically significant at the 5-percent-level, and equality of effects for treatments A and B can clearly be rejected (p=0.025).

Components of the investment index. Turning to total investments into the own business, both treatments yield positive effects. The effect sizes – are estimated to be 94,000 UGX for treatment A and 41,000 UGX for treatment B (column 4 in Table 5). Despite relatively large standard errors, the effect of treatment A is statistically and economically significant: The increase in investment by 94,000 UGX is equivalent to an increase by 29.8 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 captures whether respondents state to have 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 of 33 percent) and significantly different from zero. In contrast, the effect of treatment B, i.e. lecturing, is estimated to be insignificant from zero.

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 confirm the general picture, that active learning consistently outperforms lecturing.

The *financial literacy score* is a combination of five items (see Table A2). Even though the overall effect size is moderate and only marginally statistically significant from zero, the results on single components reveal that active learning treatment had a significant impact on two out of five items (which were of medium and high difficulty; see Table C1 in Appendix C), whereas lecturing had no effect (see Table A2).

The *budgeting index* consists of five binary items capturing changes in budgeting and record keeping behavior. Here, none of the items is impacted to a positive extent. Thus, the aggregate impact is insignificant from zero on average.

The *borrowing index* consists of six items (see Table A4 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 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; Suri and Jack, 2016). 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.

5 Exploring causal pathways

The remarkable difference we have uncovered between the two versions of financial education intervention raises the question of potential causal mechanisms. Given that the financial education treatments are estimated to have zero impacts on incomes and hours worked (cf. Appendix A), the causal pathways from financial education to savings and investment outcomes warrant an investigation. We motivate our procedure here (Section 5.1), then we introduce and discuss three kinds of potentially intermediating variables (Section 5.2), and finally we show exploratory results on causal pathways to behavior change (Section 5.3).

5.1 Potential mechanisms of financial education impact

The earlier literature on the evaluation of financial education often focused 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 recent RCTs have clearly demonstrated that there is a positive effect on both financial literacy and downstream financial behaviors, 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 understand 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 outcomes. An important part of this discussion is to better understand the causal mechanisms by which financial education impacts behavior.

Regarding the economic content of such mechanisms, the first candidate is, of course, an improvement in financial literacy which enables individuals to make better financial decisions. Evidence on this possible causal pathway has been documented in recent studies (cf. Fort et al., 2016; Sayinzoga et al., 2016) and also seems 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., 2011, 2017). Generally, evidence shows in this respect that better self-control, and in line with it also future-oriented time-preferences, seem to be associated with more savings (e.g. Ashraf et al., 2006). Finally, financial confidence and attitudes play an important role in financial behavior which explains why parental education or socialization has clear effects on financial behavior (e.g. Grohmann et al., 2015). Overall, there may be three main mechanisms at work which we analyze in the following, i.e. financial literacy, self-control and financial confidence.

Regarding the empirical test of a causal pathway through these variables we follow recent work by Sayinzoga et al. (2016) by applying a two-stage estimation approach (see Section 2.3).

5.2 Intermediating variables

In the following we describe the formation of the three above mentioned variables. While the concept and measurement of "financial literacy" was introduced above already (see Section 4 and Appendix C), "self-control" and "financial confidence" have not been defined yet. We measure the effect on these intermediate outcomes through items in our survey.

"Self-control" is assessed by a survey item asking to respond to the following question on a 1 (often) to 4 (never) rating scale: "If you get money, do you tend to spend it too quickly?" Responses are transformed into a z-score, scaled by the mean and standard deviation for the control group (see descriptive statistics in Table A6 in Appendix A).

"Financial confidence" is assessed by multiple items which 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 descriptive statistics and definitions in Table A6 in Appendix A).

5.3 Results on causal pathways

The results from applying the two-step estimation approach are presented in Table 5. In the two Panels A and B we each analyze one of the two indices of financial behavior were we

found significant average treatment effects, as they have been introduced above (see Section 4). Focusing first on the results of the first-stage regressions (which differ across the two panels because of different number of observations on outcomes), we find that the active learning treatment has significant effects on all three possible intermediating variables we consider: financial literacy, self-control, and financial confidence. Treatment B, however, only has an effect on financial confidence but neither on our measure of financial literacy nor on our 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 behavior change.

< Table 5 about here >

Looking at the second-stage regressions, we see that improved savings behavior is mainly impacted via better self-control, while the investment index is impacted via improved financial literacy and financial confidence. Thus, all three intermediating variables may play a role, however, in different ways: the nexus between self-control and savings confirms other studies, and investment outcomes are mainly affected by literacy and confidence. We note that treatment A, i.e. the active learning approach, seems to activate all three mechanisms and all of them are helpful to change financial 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

Simultaneous quantile regressions show indeed that the advantageous effects of the financial education training are heterogeneous. Starting with the financial literacy index, the effect of treatment A (active learning) is largely independent from the outcome level of financial literacy (see Table 6, column 1). Graphically speaking, the entire distribution of financial literacy 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, i.e. for individuals who may never have been confronted with any of the contents of the training. Active learning, however, appears to be beneficial at every stage of the financial literacy distribution.

< Table 6 about here >

Regarding the outcome level of the savings index, treatment A has clearly larger and even significant effects above the median, while effects of treatment B are very small and insignificant across the whole distribution. Thus, differential treatment effects of active learning within the savings domain are driven by relatively large effects in the 60^{th} and 80^{th} percentile, while effect sizes are small below the median. This would indicate that financial education may increase savings-inequality with respect to our target group.

A similar pattern arises with respect to the investment index, although at a higher level of effect sizes. Treatment A has significant effects from the 40^{th} percentile upwards. Again, effect sizes are largest at the top of the outcome distribution. Also treatment B appears to generate the intended effects at some points of the outcome distribution, being significant at

the median and at the 60th percentile, but always with smaller estimated effect sizes and overall insignificant average impact.

Overall, we draw two conclusions: first, active learning seems to be effective in increasing financial literacy across the entire distribution; second, regarding the savings and the investment index, effects are largest for those individuals with higher levels of outcomes and zero for people at the bottom of the outcome distributions for savings and investing. Thus, financial education may not be beneficial for the most constrained individuals in our sample, confirming the intuition that investments in financial literacy may not be rational and beneficial for all individuals (cf. Lusardi et al., 2017).

6.2 Subgroup analyses

Turning to an investigation of treatment effects by subgroups along observable characteristics of the respondents, we examine the possibility of heterogeneous treatment effects for three subgroups which generally 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 literacy tests than women in most surveys. Second, financial literacy is correlated with general educational attainment. In our sample, nearly 32 percent cannot read or write in any language and only and 25.35 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.

The three panels of Table 7 show an investigation of heterogeneous treatment effects for the subgroups discussed above. In each panel, binary indicators for each group are interacted with the treatment dummies to estimate the heterogeneous effects. The results are mixed:

Starting with the impact on financial literacy scores (column 1), it can be observed that treatment effects of treatment A appear to be heterogeneous only with respect to gender. First, the positive treatment effect on financial literacy may be driven predominately by a very strong treatment effect on males' financial literacy scores whereas the treatment effect for females appears to be small and insignificant (*Panel A*). Considering the other indicators, no significant interactions appear to exist (*Panels B anc C*).

Turning to outcomes with regard to financial behaviors, however, indicates that treatment effects appear to be less heterogeneous and the average positive treatment effects discussed above are still present when investigating effects for subgroups.

Apart from strong negative interaction effects of the male indicator with the dummy for treatment B (suggesting males respond negatively to this treatment leading to worse outcomes on the debt index but) none of the interactions appear to be statistically significant and meaningful. 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 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

being 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 robustness checks 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 some attrition we carefully address this issue showing that there is no reason for concern, among others by a bounds analysis testing results under extreme assumptions and inverse probability weighting for selection into endline survey . (ii) We rerun the analyses on binary outcomes with logit and probit-models. (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 the more common measures lead to qualitatively identical results.

8 Conclusion

This 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 the "traditional lecturing" that appears to dominate current financial education programs. Active learning is shown to be clearly superior in the field of science education, which provides a strong motivation to test the approach in financial education. In doing so, we employ a design similar to Drexler et al. (2014), comparing two financial education treatment types. However, instead of comparing two different curricula, thus varying the content, we test the impact of *how* financial education is conducted using the same content, teaching the same target group and relying on the same teachers.

The main result is clear cut: We study five outcome groups of financial behavior, such as savings behavior and financial literacy, finding that active learning outperforms traditional lecturing. In four out of six cases, active learning has the intended effect 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 debt index. Still, given the relatively limited power of our cluster-RCT, with 1,162 vendors in the endline survey and a short training of only 120 minutes, the advantage of active learning seems quite strong. Given that existing studies rely on video screenings or information interventions to test the effects of financial education programs in developing countries (e.g. De Mel et al., 2011; Carpena et al., 2017), we argue that one explanation for the muted effects of classroom programs observed in the present literature may lie in the form of exposition centered instructional design that falls short of translating financial literacy into financial behavior.

Thus, our experiment adds to the growing body of high-quality evidence that financial education can positively impact financial behavior, especially savings behavior (cf. Skimmyhorn 2016; Calderone et al., 2017).

In a second step, we aim to better understand the mechanisms intermediating financial education into changes in of financial behavior. For this purpose, we examine the role of financial literacy, self-control, and financial confidence separately by applying a two-step estimation in line with Sayinzoga et al. (2016), where the intermediating variables are instrumented by the financial education treatments. We find, first, that active learning has a positive impact on all three intermediating variables, while lecturing only impacts financial confidence. Second, we reveal that financial outcomes seem to be impacted by changes in specific intermediating variables, suggesting that the transfer from education to a change in behavior differs depending upon the kind of behavior. This would call for financial education

that is not only specific to target groups (cf. Gibson et al., 2014; Doi et al., 2016) but also to the outcome variables. Obviously, more research in this direction is necessary in order to fully understand the mechanisms at play.

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.

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): 256–95.

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.

Calderone, M., Fiala, N., Mulaj, F. Sadhu, S., and Sarr, L. (2017). 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. (2011). Unpacking the causal chain of financial literacy. *World Bank Policy Research Working Paper,* No. 5798.

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.

De Mel, S., McKenzie, D., and Woodruff, C. (2011). Getting credit to high return microentrepreneurs: The results of an information intervention. *World Bank Economic Review*, 25(3): 456–485.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100(3): 1238–1260.

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.

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.

Gibson, J., McKenzie, D., and Zia, B. (2014). The impact of financial literacy training for migrants. *World Bank Economic Review*, 28(1): 130–161.

Grohmann, A., Kouwenberg, R., and Menkhoff, L. (2015). Childhood roots of financial literacy. *Journal of Economic Psychology,* 51: 114–133.

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.

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., & 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. (2017). Visual tools and narratives: New ways to improve financial literacy. *Journal of Pension Economics and Finance,* 16(3): 297–323.

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

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.

Muller, S. M. (2015). Causal interaction and external validity: Obstacles to the policy relevance of randomized evaluations. *World Bank Economic Review*, 29: S217–S225.

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.

Saez, E. (2009). Details matter: The impact of presentation and information on the take-up of financial incentives for retirement saving. *American Economic Journal: Economic Policy*, 1(1): 204–228.

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. (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.

Suri, T. and Jack, W. (2016). The long-run poverty and gender impacts of mobile money. *Science*, 354(6317): 1288–1292.

van Rooij, M., A. Lusardi, and Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics* 101(2): 449–72.

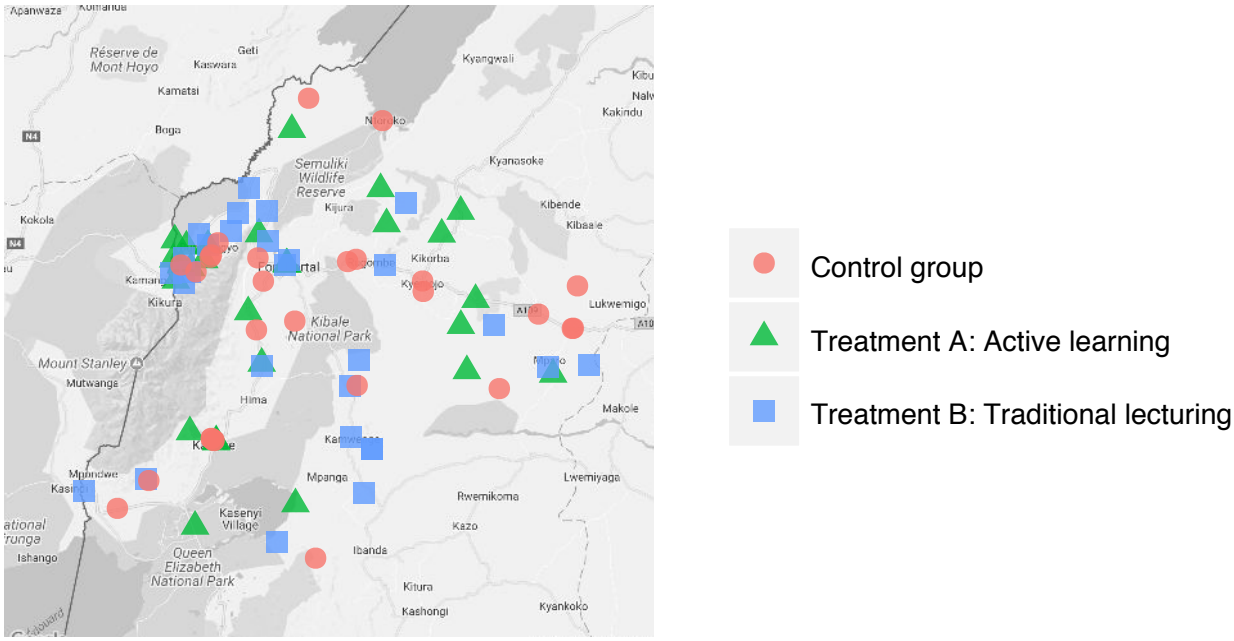

Figure 1: Location and treatment status of 83 clusters

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) which 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 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 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]
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							
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]
Patience (z-score)	1,280	0.00	0.02	-0.07	-0.09	0.05	0.04
		(1.00)	(1.01)	(1.04)	[0.09]	(0.95)	[0.07]
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]

-continued -

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	1,253	0.00	0.03	-0.02	-0.05	-0.01	-0.04
(z-score)		(1.00)	(0.99)	(1.03)	[0.08]	(0.98)	[0.10]
General risk attitude	1,262	3.61	3.53	3.67	0.14	3.66	0.13
(0-10)		(2.42)	(2.39)	(2.40)	[0.17]	(2.48)	[0.19]
Specific risk attitude	1,272	3.78	3.72	3.78	0.06	3.85	0.13
(0-10)		(2.52)	(2.53)	(2.46)	[0.16]	(2.57)	[0.19]
Panel C: Outcome measures at baseline							
(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]
(2) 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]
(3) Savings index	1,161	0.04	0.00	0.07	0.07	0.05	0.05
		(1.04)	(1.00)	(1.14)	[0.09]	(1.00)	[0.10]
(4) Borrowing index	1,126	0.06	0.00	0.01	0.01	0.17	0.17
		(1.25)	(1.00)	(1.32)	[0.08]	(1.43)	[0.11]
(5) Investments index	1,142	0.04	0.00	0.00	0.00	0.10	0.10
		(1.07)	(1.00)	(1.08)	[0.11]	(1.14)	[0.11]
(6) 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: any treatment (p-value)					0.308		
Observations					862		
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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Main experimental results (ANCOVA)

	(1) Fin. literacy (z)	(3) Budget index (z)	(4) Savings index (z)	(5) Debt index (z)	(6) Investment index (z)	(7) Fin. services index (z)
Treatment A	0.134*	0.045	0.162**	0.104*	0.284***	0.109
	(0.079)	(0.087)	(0.071)	(0.057)	(0.097)	(0.077)
Treatment B	0.079	0.009	0.011	-0.038	0.168	0.149
	(0.071)	(0.088)	(0.077)	(0.075)	(0.119)	(0.095)
$A - B = 0$ (p-value)	0.494	0.690	0.079*	0.052*	0.353	0.678
R^2	0.051	0.108	0.151	0.009	0.144	0.130
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	1,160	1,114	1,160	1,108	1,007	1,136
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 indices of financial literacy and 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 (4), (5), and (6) are 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 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
Treatment A	0.036*	109,186*	145,480**	90,173*	0.077**
	(0.020)	(65,132)	(64,784)	(47,461)	(0.034)
Treatment B	0.025	-32,519	-14,226	41,801	0.060
	(0.023)	(68,588)	(66,476)	(63,124)	(0.037)
$A - B = 0$ (p-value)	0.612	0.052*	0.025**	0.487	0.640
R^2	0.023	0.244	0.131	0.184	0.035
Mean (SD) of y_t in control group	0.879 (0.326)	513,629 (937,119)	380,568 (973,769)	301,067 (526,957)	0.232 (0.423)
Observations	1,160	1,162	1,162	1,053	1,110
Clusters	83	83	83	83	83
District FEs	yes	yes	yes	yes	yes
$y_{(t-1)}$ covariate	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 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 5: 2SLS-Regressions to probe the causal mechanisms from financial education to savings outcomes

	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	0.156*	0.152*	0.160**			
	(0.087)	(0.079)	(0.081)			
Treatment B	0.083	-0.058	0.193**			
	(0.076)	(0.092)	(0.083)			
Fin. literacy*				1.248		
				(0.824)		
Self-control*					0.860*	
					(0.518)	
Fin. confidence *						0.515
						(0.476)
Observations	1,161	1,156	1,026	1,161	1,156	1,026
	Panel B: Investment outcomes					
	Fin. literacy	Self-control	Fin. confidence	Investment index (z)	Investment index (z)	Investment index (z)
Treatment A	0.188 **	0.134	0.141*			
	(0.094)	(0.086)	(0.079)			
Treatment B	0.098	-0.073	0.217***			
	(0.083)	(0.093)	(0.074)			
Fin. literacy*				1.683*		
				(0.955)		
Self-control*					0.734	
					(0.660)	
Fin. confidence*						1.321*
						(0.742)
Observations	1,037	1,033	920	1,037	1,033	920
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 6: Simultaneous-quantile regressions for indices of financial knowledge, savings, and investment

		(1) FL score (z)	(2) Savings index (z)	(3) Investment index (z)
20th percentile	Treatment A	0.161	0.035	0.027
		(0.119)	(0.030)	(0.017)
	Treatment B	0.281**	0.015	0.015
		(0.120)	(0.031)	(0.020)
	$A - B = 0$ (p-value)	0.372	0.382	0.560
40th percentile	Treatment A	0.180*	0.026	0.095**
		(0.109)	(0.020)	(0.042)
	Treatment B	0.204*	-0.002	0.044
		(0.108)	(0.020)	(0.035)
	$A - B = 0$ (p-value)	0.804	0.217	0.315
Median	Treatment A	0.108	0.037	0.155**
		(0.111)	(0.027)	(0.0649)
	Treatment B	0.101	-0.003	0.108**
		(0.096)	(0.027)	(0.055)
	$A - B = 0$ (p-value)	0.941	0.180	0.527
60th percentile	Treatment A	0.154	0.068**	0.301**
		(0.135)	(0.029)	(0.143)
	Treatment B	-0.022	0.007	0.137*
		(0.115)	(0.029)	(0.076)
	$A - B = 0$ (p-value)	0.154	0.052*	0.263
80th percentile	Treatment A	0.136	0.134*	0.466**
		(0.095)	(0.073)	(0.208)
	Treatment B	-0.037	-0.037	0.187
		(0.096)	(0.057)	(0.206)
	$A - B = 0$ (p-value)	0.091*	0.022**	0.162
Observations		1,162	1,160	1,007
Clusters		83	83	83
District FEs		yes	yes	yes
$y_{(t-1)}$ covariate		yes	yes	yes
Mean (SD) of y_t in control group		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. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 7: Heterogeneous treatment effects for subgroups

	(1) FK score (z)	(2) Budget index (z)	(3) Savings index (z)	(4) Debt index (z)	(5) Investment index (z)	(6) Fin. services index (z)
Panel A: Male respondents						
Treatment A	0.053	0.029	0.184**	0.141**	0.224**	0.062
	(0.100)	(0.085)	(0.076)	(0.065)	(0.0863)	(0.077)
Treatment B	0.044	0.001	0.0391	0.0470	0.168	0.168*
	(0.086)	(0.089)	(0.072)	(0.088)	(0.106)	(0.091)
Male	-0.101	0.307**	0.306**	0.407***	0.406**	0.362***
	(0.101)	(0.147)	(0.125)	(0.122)	(0.172)	(0.109)
Treatment A × Male	0.392**	0.079	-0.0442	-0.205	0.274	0.157
	(0.170)	(0.214)	(0.204)	(0.185)	(0.348)	(0.195)
Treatment B × Male	0.155	0.059	-0.101	-0.409**	0.0849	-0.0121
	(0.157)	(0.182)	(0.244)	(0.194)	(0.285)	(0.190)
Obs.	1,138	1,091	1,136	1,085	987	1,088
R2	0.053	0.123	0.158	0.018	0.177	0.156
Panel B: Respondents with beyond primary education						
Treatment A	0.121	0.066	0.183**	0.134**	0.237**	0.110
	(0.095)	(0.082)	(0.077)	(0.065)	(0.100)	(0.082)
Treatment B	0.149	0.044	0.061	-0.044	0.128	0.104
	(0.091)	(0.081)	(0.067)	(0.100)	(0.112)	(0.106)
Educated	0.165	0.303***	0.334***	0.184*	0.055	0.228**
	(0.105)	(0.106)	(0.098)	(0.108)	(0.113)	(0.100)
Treatment A × Educated	0.126	-0.053	-0.032	-0.102	0.229	0.032
	(0.147)	(0.147)	(0.164)	(0.139)	(0.215)	(0.168)
Treatment B × Educated	-0.256	-0.107	-0.172	0.037	0.169	0.199
	(0.164)	(0.174)	(0.156)	(0.187)	(0.267)	(0.138)
Obs.	1,162	1,114	1,160	1,108	1,007	1,111
R^2	0.0577	0.120	0.164	0.014	0.150	0.146
Panel C: Respondents with above average financial literacy at baseline						
Treatment A	0.099	-0.140	0.260**	0.162*	0.198*	0.084
	(0.105)	(0.0942)	(0.105)	(0.087)	(0.110)	(0.118)
Treatment B	0.092	-0.121	0.158	-0.068	0.071	0.100
	(0.099)	(0.104)	(0.108)	(0.099)	(0.109)	(0.137)
Fin. literate	0.090	-0.122	0.258**	0.046	0.055	0.118
	(0.132)	(0.079)	(0.111)	(0.095)	(0.091)	(0.093)
Treatment A × Fin. literate	0.079	0.336**	-0.194	-0.106	0.145	0.036
	(0.143)	(0.132)	(0.144)	(0.131)	(0.143)	(0.136)
Treatment B × Fin. literate	-0.019	0.250*	-0.283*	0.060	0.188	0.102
	(0.154)	(0.146)	(0.143)	(0.124)	(0.166)	(0.160)
Obs.	1,162	1,114	1,160	1,108	1,007	1,111
R^2	0.052	0.115	0.157	0.010	0.150	0.136
Clusters	83	83	83	83	83	83
District FEs	yes	yes	yes	yes	yes	yes
$y_{(t-1)}$ controls	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

(online appendix not intended for publication)

to accompany

"**Active learning fosters financial behavior: Evidence from rural Uganda**"

Appendix A: Supplementary tables and figures

Appendix B: Robustness checks

Appendix C: Measuring financial literacy

Appendix A: Supplementary tables and figures

Figure A1: Timeline

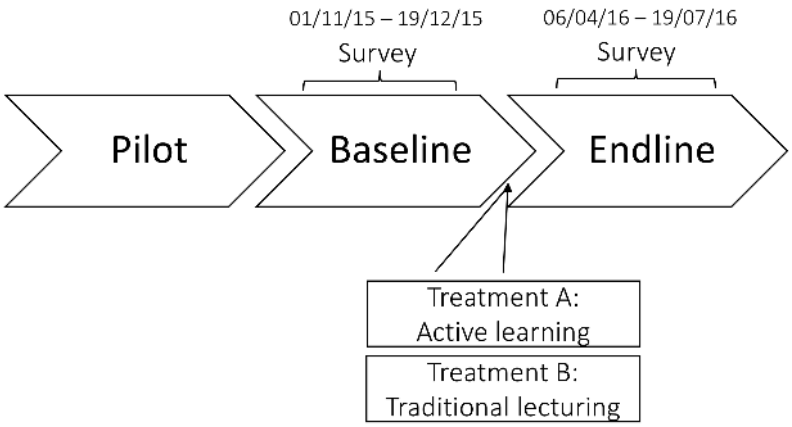

Table A1: Overview of financial education treatments

	Topic	Treatment A: Active learning		Treatment B: Community lecture	
		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 asks the participant to cross out wrong	5	The lecturer asks the participants whether they have questions or comments.

			statements. This results in a collection of ten statements on the benefits of savings which are subsequently repeated.		
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 institutions</u> 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.

Outcomes on further items within indices. The *financial literacy score* is a combination of five items. As the financial education trainings provided here do not primarily aim at improving financial literacy with regard to financial numeracy, it may be not surprising that progress in this domain is modest. Still, the coefficients are almost all positive and treatment A realizes significant treatment effects on 2 out of 5 items a, i.e. on items 3 and 5 (see Table A2) which are of medium and high difficulty (see Table C1 in Appendix C).

The *budgeting index* consists of five binary items capturing changes in budgeting and record keeping behavior (see Table A3 in Appendix A). 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.

The *borrowing index* consists of six items (see Table A4 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. 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.

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. 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).

Overall, even if the three aggregated indices on budgeting, borrowing 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.

Table A2: Analysis of 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 (0.477)	0.672 (0.470)	0.444 (0.497)	0.436 (0.496)	0.512 (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 2PML 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 A3: 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 A4: Borrowing index – individual components

	(1) Loans for consumption*	(2) Loans for productive investment	(3) Debt to asset ratio (z-score)*	(4) Would take loan if no clear plan*	(5) Can correctly identify a "bad" loan
Treatment A	2,142	-15,615	-0.089	-0.055**	-0.001
	(5,840)	(37,458)	(0.074)	(0.023)	(0.010)
Treatment B	12,855	-43,846	0.006	-0.038	0.004
	(9,345)	(36,785)	(0.090)	(0.023)	(0.010)
$A - B = 0$ (p-value)	0.231	0.4043	0.268	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 (98,688)	151,021 (387,398)	0.000 (1.000)	0.139 (0.346)	0.029 (0.167)
Observations	1,158	1,158	1,153	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 an 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 A5: 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 A6: Descriptive statistics for intermediating 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]
Fin literacy (z-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]
Fin. confidence (z-score)	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]
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]

Notes: For individual components of financial literacy score see Appendix C. Fin. confidence is an aggregate measure (sum and standardized to have a zero mean and a SD of one for the control group) of the following items: (1) "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). (2) "I am confident enough to approach a bank and ask questions to learn more about their products." (disagree strongly – agree strongly). (3) "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). (4) 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. (5) Self control: "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 A7: 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) 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) 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]
(3) Savings index	1,161	0.04	0.00	0.07	0.07	0.05	0.05
		(1.04)	(1.00)	(1.14)	[0.09]	(1.00)	[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+	1,162	267,290	239,830	302,031	63,201	262,057	22,227
		(786,995)	(770,826)	(848,947)	[60,234]	(736,135)	[73,996]
(4) Borrowing index	1,126	0.06	0.00	0.01	0.01	0.17	0.17
		(1.25)	(1.00)	(1.32)	[0.08]	(1.43)	[0.11]
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-score)	1,157	0.01	0.00	0.00	-0.00	0.02	0.02
		(0.93)	(1.00)	(0.85)	[0.07]	(0.92)	[0.07]
Would take a 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]
(5) Investments index	1,142	0.04	0.00	0.00	0.00	0.10	0.10
		(1.07)	(1.00)	(1.08)	[0.11]	(1.14)	[0.11]
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 registered	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]
(6) 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]

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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix B: Robustness checks

This appendix mainly deals with the issue of selective attrition. In the following, we analyze determinants of attrition and examine the implications of attrition on the estimated treatment effects by estimating bounds for the treatment effects (A) and (B) under a plausible scenario for outcomes of those who are not met at the follow-up survey. Finally, we show that correcting for differential attrition according to treatment groups and districts through inverse-probability weighting yields similar results to an unweighted analysis that ignores attrition.

Selective attrition. As mentioned in Section 3.3, attrition varies between the treatment and control groups. 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 to be driven by this exogenous shock.

To investigate further determinants of attrition, we estimate a linear probability model of being absent at follow-up on selected baseline covariates which have a relatively small number of missing values. In a second step, we test for differential determinants depending on the treatment group by interacting these baseline covariates with the treatment dummies.

< Table B1 about here >

Several characteristics are systematically related to dropping out from the survey. In general, being allocated to treatment B is a positive determinant of attrition, even when controlling for other baseline covariates. Column (1) of Table B1 shows that, in general, females are about 5 percentage points less likely to drop out. Also, a negative coefficient on age indicates that younger respondents are more likely to be absent at follow-up. However, the marginal effect of a one-year increase in age is only estimated to be a 0.3 percentage point increase in the probability to drop out. Finally, there appears to be a negative effect of the total monthly household consumption on the probability to be absent at baseline. However, this effect, while statistically significant at the 5%-level, is economically miniscule: A 100,000 UGX increase in household consumption (an almost 50% increase over the mean) would only result in a 0.4 percentage points decrease in the probability to be absent at baseline.

Turning to differential determinants for the treatment groups, it appears that none of the interaction terms between the indicator for being assigned to group A is statistically significant at the 5%-level. The interaction effects between the indicator for treatment B and baseline covariates indicate that most of the determinants of attrition appear to be driven by characteristics of this group. The linear predictors mostly lose explanatory power, and, despite effects of monthly income heading into opposite directions for the treatment groups and the control group (they are economically very small), heterogeneous effects on attrition may be present in group B. In group B, age is a strong negative predictor of being absent at follow-up, and those who are literate are also less likely to drop out. Thus, we may be missing young and less educated respondents in this group. Depending on the assumptions of the average treatment effect in this group this may lead to an upward (or downward) bias. Note, however, that these models explain only 5 and 7% of the variance in attrition. Nevertheless, there is

evidence for selective attrition. Especially, there seems to be a treatment effect of being assigned to treatment B on dropping out of the survey, even when controlling the exogenous district-level shock through district dummies and when controlling for other baseline covariates.

Inverse probability weighting. Next, we account for potential biases due to selective attrition. We weight all regressions with inverse probability of selection into endline tracking as estimated by a logit-model based in the specification in Column 1 of Table B1.

<Table B2 about here>

The results are very similar to the unweighted specification discussed in the main text.

Bounds estimates for the financial education treatments. To 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 lower bound, we impute the minimum value of each outcome variable in the observed distribution of y to the attriters in the treatment groups, and the maximum value of the observed control distribution to the attriters in the control group. For the upper bound, we impute the maximum value of each outcome variable in the observed treatment distributions to the attriters in the treatment groups, and the minimum value of the observed control distribution to the attriters in the control group (cf. Karlan and Valdivia, 2011, p. 522). Thus, column (1) in Table B2 shows the lower-bound estimate for the worst-case scenario of selection bias, and column (3) shows the upper-bound estimate for the worst-case scenario. Column (2) shows the unconditional treatment effect. All coefficients are estimated within a single-difference framework and include district-fixed-effects.

<Table B3 about here>

Turning to the worst case lower-bound scenario for treatment effects on the savings index, column (1) of *Panel A* shows that both treatments are estimated to have negative effect, which appears to be even significant at the 5-% level in the case of treatment B. Thus, equality of treatments can be rejected by an F-test (testing $A - B = 0$). Turning to the worst case upper bound, both coefficients are estimated to be positive but only the effect of treatment A is large and statistically highly significant at the 1%-level. The coefficient B is estimated to be roughly one third of the size of treatment A (0.7 vs. 0.21 standard deviation units) and statistically insignificant. Thus, the differential treatment effect on the savings index can be confirmed regardless of the extreme missing data scenarios. Treatment A is always estimated to have larger (or less adverse) effects on savings.

Panel B shows the lower and upper bounds for the worst-case scenario regarding the treatment effect on financial literacy. The lower bound worst case-scenario (column 1) shows large negative effect sizes, statistically significant only for treatment B. Again, treatment A is estimated to have less of an adverse effect in this scenario. The test of equality of coefficients is rejected. In the upper bound scenario both treatments are estimated to yield strong effects on financial behavior (around 0.4 standard deviation units), and the effect of treatment B is now even estimated to be stronger than the effect of treatment A. However, equality of treatment effect coefficients cannot be rejected in this scenario. Overall, these examinations show that selective attrition may have an impact on the estimated treatment effects. These scenarios are extreme by design, and thus, unlikely to be at work in the sample of attriters. Instead, the results are relatively reassuring, since the main results regarding the differential impact of the two treatments on savings appear to be robust to imputing missing values according to either one of these extreme scenarios.

Standard FL-score. Table B4 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 main result presented in the text, albeit with larger standard errors due to the relative imprecise measure of financial literacy.

References in Appendix B

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.

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

Table B1: Determinants of attrition

Baseline covariates	Binary for attrition		
	(1) Attrition	(2) Attrition	(3) Attrition
Treatment A	-0.020	-0.030	0.085
	(0.024)	(0.023)	(0.099)
Treatment B	0.055*	0.058*	0.388***
	(0.031)	(0.033)	(0.131)
Female		-0.053**	-0.012
		(0.024)	(0.027)
Age		-0.003***	-0.001
		(0.000)	(0.001)
Literate		-0.039*	0.014
		(0.020)	(0.028)
Monthly consumption+		-0.000**	-0.000
		(0.000)	(0.000)
Monthly income+		0.000	-0.000***
		(0.000)	(0.000)
Treatment A × Female			-0.053
			(0.041)
Treatment A × Age			-0.001
			(0.002)
Treatment A × Literate			-0.058
			(0.042)
Treatment A × Monthly consumption+			-0.000
			(0.000)
Treatment A × Monthly income+			0.000*
			(0.000)
Treatment B × Female			-0.075
			(0.070)
Treatment B × Age			-0.006***
			(0.002)
Treatment B × Literate			-0.105*
			(0.054)
Treatment B × Monthly consumption+			-0.000
			(0.000)
Treatment B × Monthly income+			0.000**
			(0.000)
Constant	0.061**	0.311***	0.163**
	(0.026)	(0.0622)	(0.0746)
District dummies	Yes	Yes	Yes
R^2	0.03	0.05	0.07
n (Individuals)	1,291	1,157	1,157
n (Clusters)	83	83	83

Notes: Results from linear probability models. + indicates that the currency denominated outcome (in Ugandan Shilling (UGX)) is winsorized at the 99th percentile. Standard errors, clustered at the market-level, 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

	(1) FK score (z)	(2) Budget index (z)	(3) Savings index (z)	(4) Debt index (z)	(5) Invest-ments (z)	(6) Fin. services index (z)
Treatment A	0.138	0.042	0.161**	0.104*	0.280***	0.109
	(0.084)	(0.086)	(0.072)	(0.057)	(0.096)	(0.077)
Treatment B	0.084	0.004	0.017	-0.045	0.168	0.158
	(0.075)	(0.086)	(0.076)	(0.076)	(0.116)	(0.095)
$A - B = 0$ (p-value)	0.521	0.672	0.094*	0.043**	0.354	0.617
R^2	0.052	0.109	0.151	0.009	0.142	0.131
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	1,160	1,114	1,160	1,108	1,007	1,136
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: 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) Worst case lower bound	(2) Unadjusted treatment effect	(3) Worst case upper bound
	Panel A: Impact on the savings index		
Treatment A	-0.037	0.121**	0.209***
	(0.059)	(0.055)	(0.066)
Treatment B	-0.196***	0.010	0.076
	(0.062)	(0.060)	(0.065)
Test: $A - B = 0$ (p-value)	0.02	0.09	0.06
Obs.	1,291	1,147	1,291
R^2	0.465	0.161	0.624
	Panel B: Impact on financial literacy scores		
Treatment A	-0.136	0.135*	0.396***
	(0.092)	(0.079)	(0.081)
Treatment B	-0.331***	0.080	0.468***
	(0.085)	(0.071)	(0.086)
Test: $A - B = 0$ (p-value)	0.04	0.49	0.45
Obs.	1,291	1,160	1,291
R^2	0.061	0.051	0.065
District dummies	yes	yes	yes
$y_{(t-1)}$ controls	yes	yes	yes

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

Table B4: Standard financial literacy 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 literacy 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. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

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

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

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 factor analysis (principal factors). 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 literacy 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, missing values, and response options Y) and Z) are coded to be equal to zero.

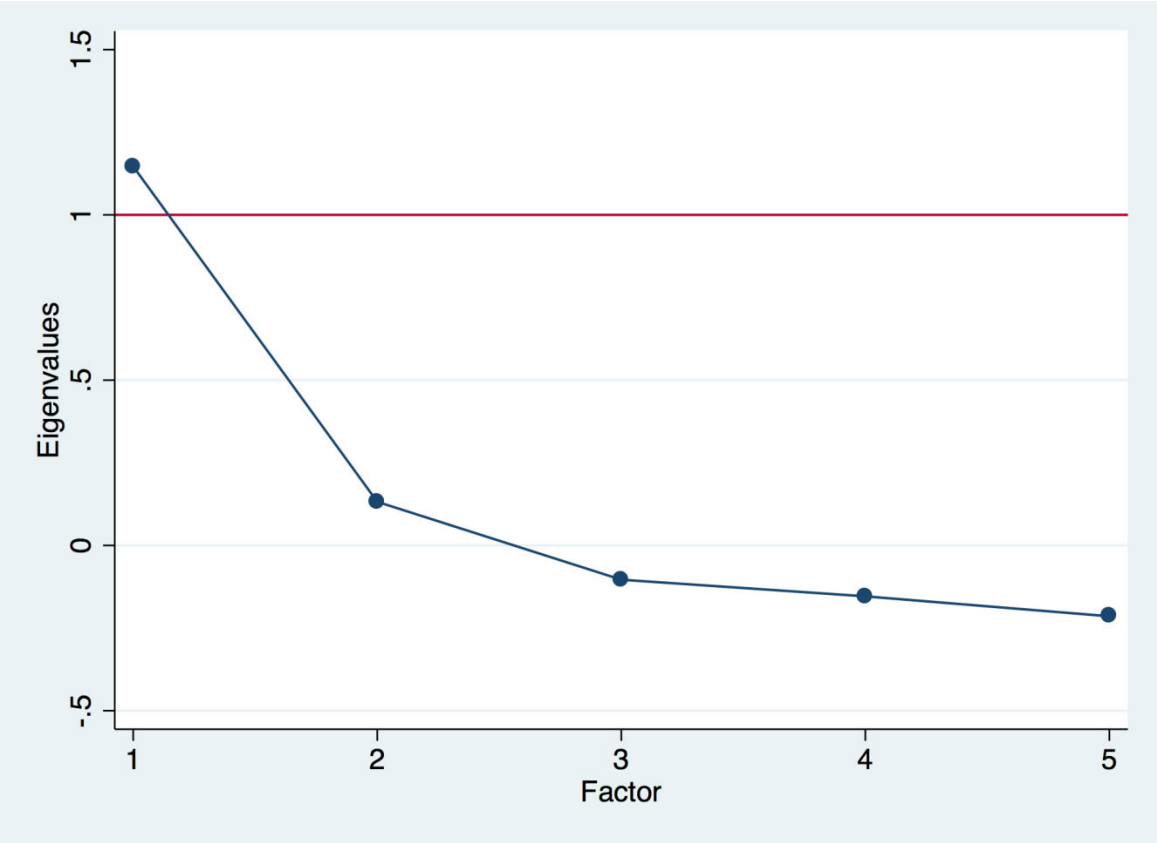

Figure C1: Screeplot of eigenvalues by factor after factor analysis (principal factors)

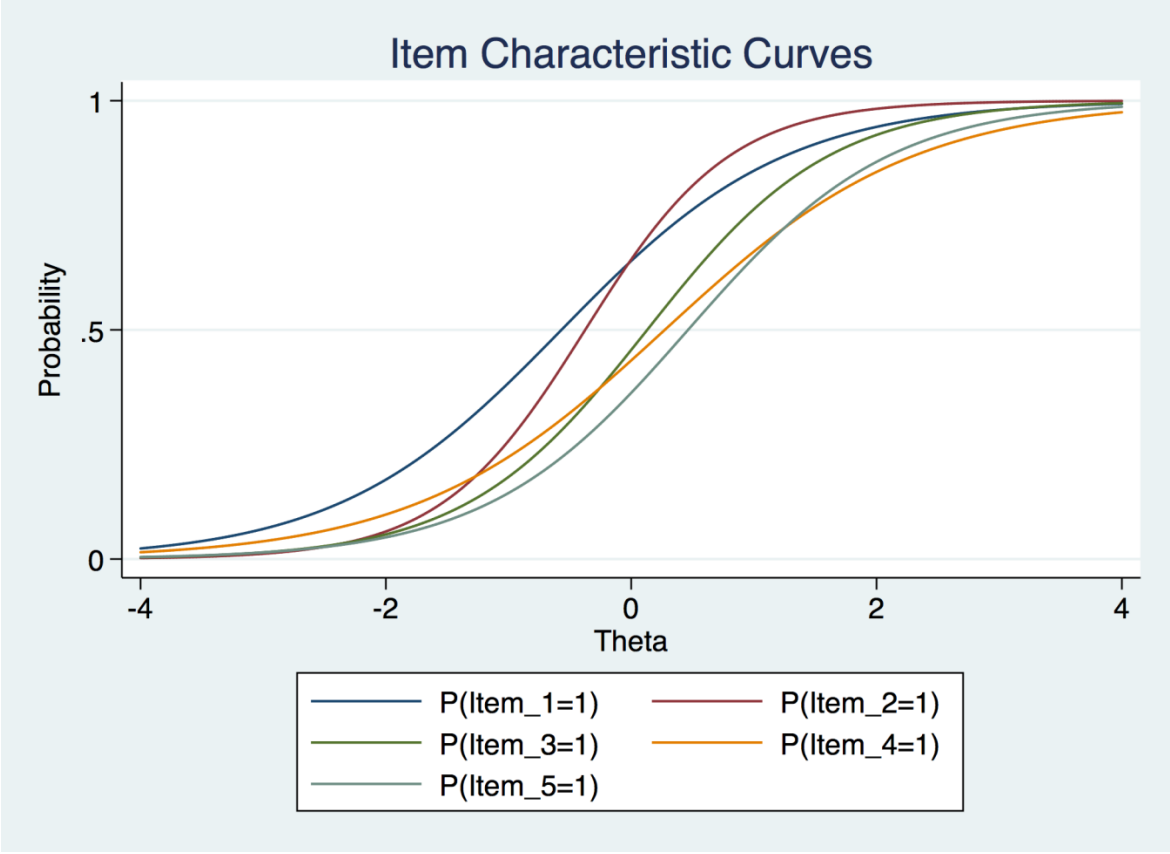

Figure C2: Item characteristic curves for the 2PLM financial literacy scale

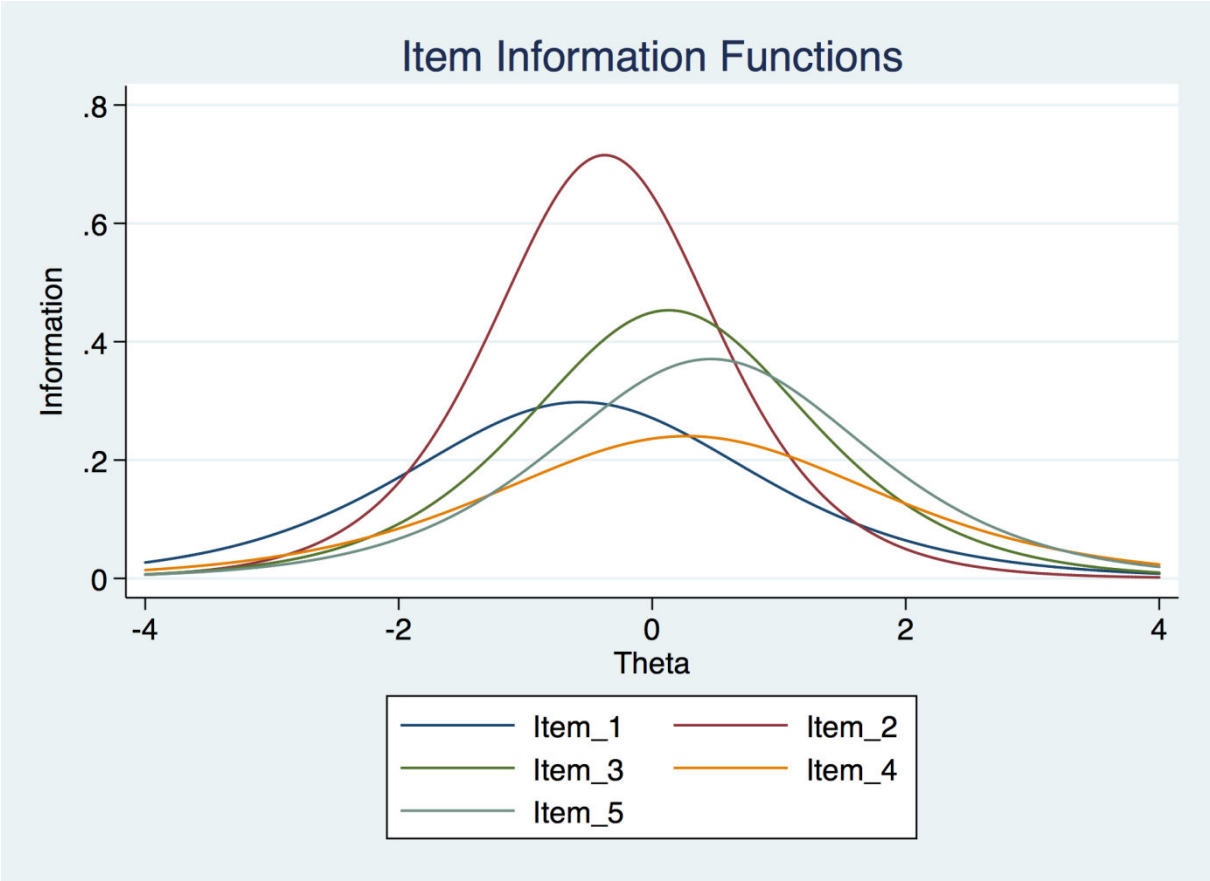

Figure C3: Item information functions for the 2PLM financial literacy scale

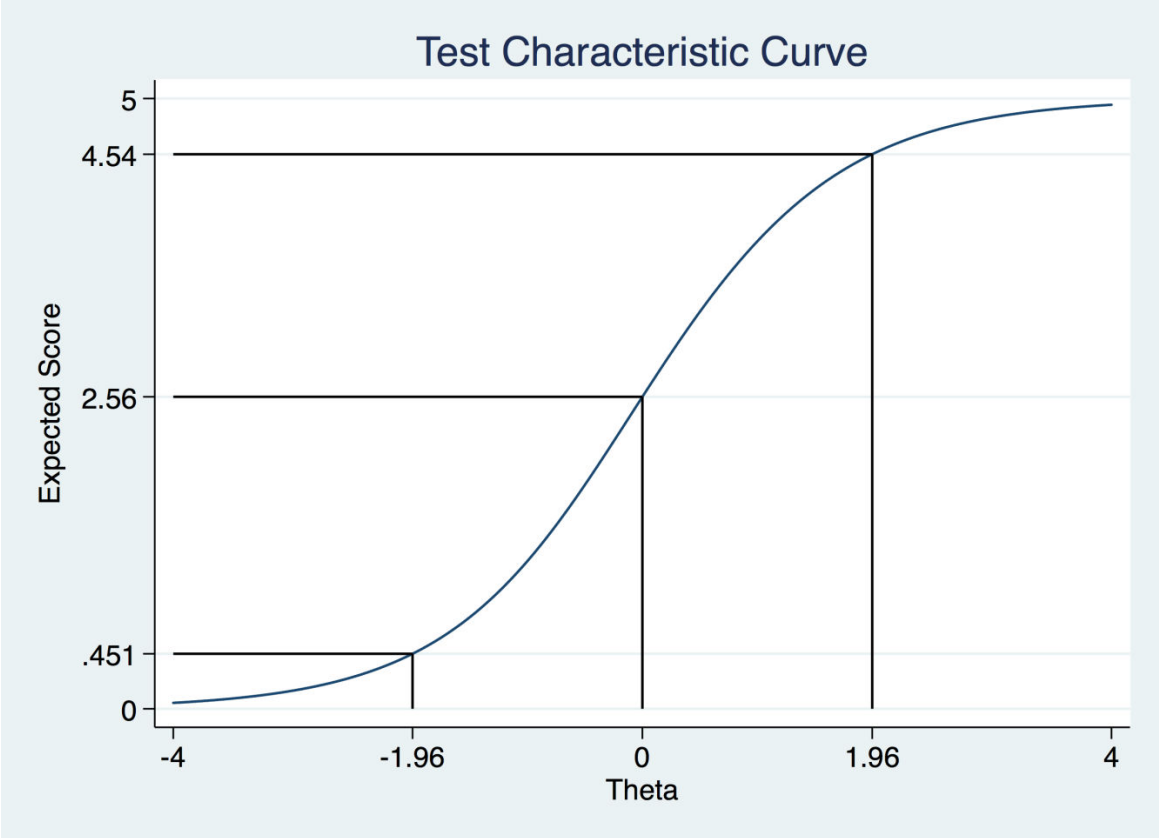

Figure C4: Test characteristic curve for the 2PLM financial literacy scale

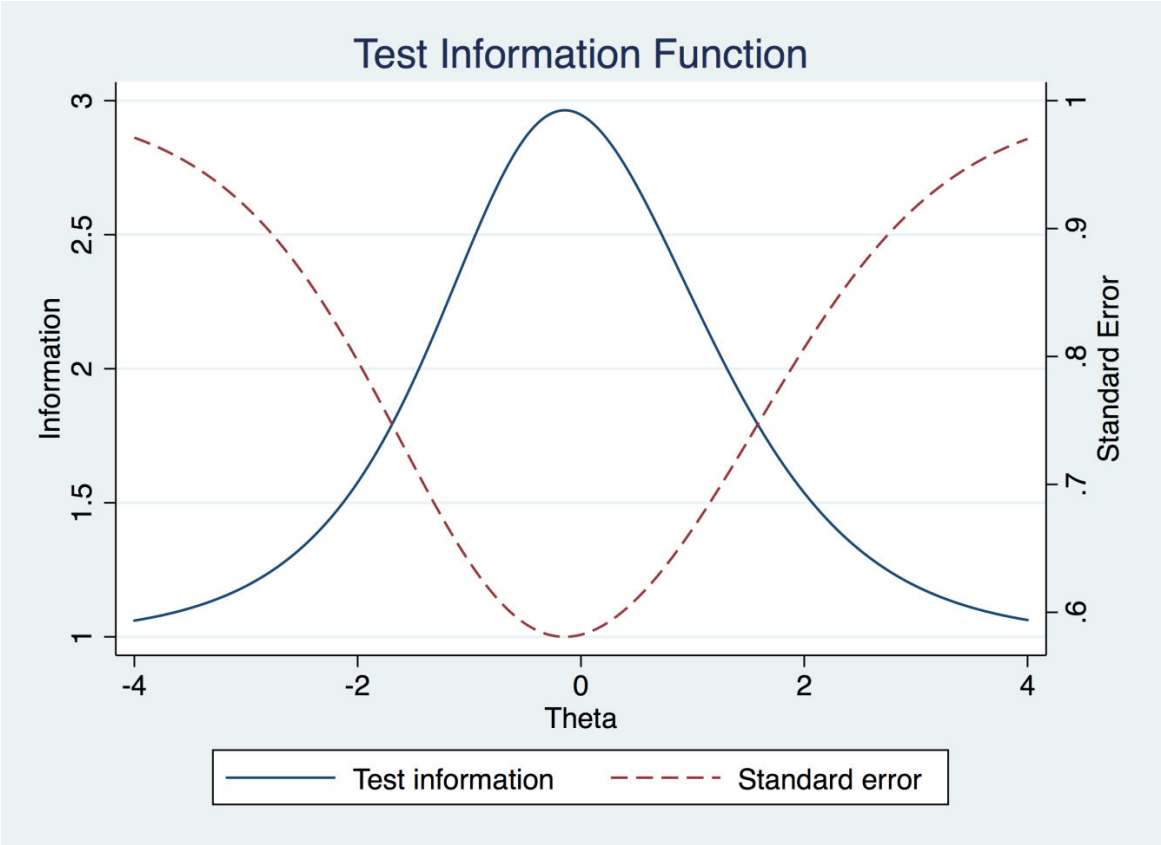

Figure C5: Test information function for the financial literacy scale

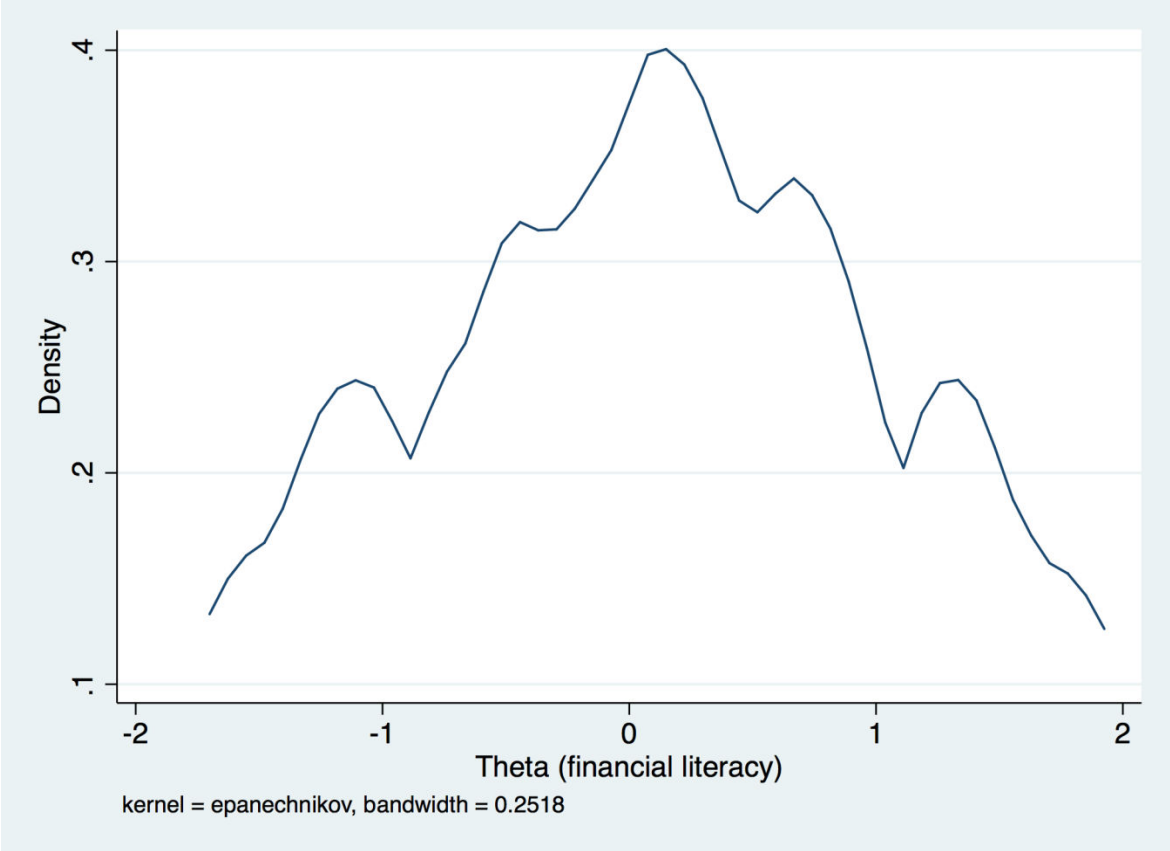

Figure C6: Distribution of standardized financial literacy IRT-scores at baseline