

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Hamdan, Jana S.; Kaiser, Tim; Menkhoff, Lukas; Xu, Yuanwei

Working Paper Scaling Financial Education Among Micro-Entrepreneurs: A Randomized Saturation Experiment

CESifo Working Paper, No. 11431

Provided in Cooperation with: Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Hamdan, Jana S.; Kaiser, Tim; Menkhoff, Lukas; Xu, Yuanwei (2024) : Scaling Financial Education Among Micro-Entrepreneurs: A Randomized Saturation Experiment, CESifo Working Paper, No. 11431, CESifo GmbH, Munich

This Version is available at: https://hdl.handle.net/10419/307361

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU



Scaling Financial Education Among Micro-Entrepreneurs: A Randomized Saturation Experiment

Jana S. Hamdan, Tim Kaiser, Lukas Menkhoff, Yuanwei Xu



Impressum:

CESifo Working Papers ISSN 2364-1428 (electronic version) Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute Poschingerstr. 5, 81679 Munich, Germany Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de Editor: Clemens Fuest https://www.cesifo.org/en/wp An electronic version of the paper may be downloaded • from the SSRN website: www.SSRN.com

- from the RePEc website: <u>www.RePEc.org</u>
- from the CESifo website: <u>https://www.cesifo.org/en/wp</u>

Scaling Financial Education Among Micro-Entrepreneurs: A Randomized Saturation Experiment

Abstract

We study the effects of scaling up a financial- and business education program in a randomized saturation experiment in Uganda. We randomly assign the program at the cluster-level, and then randomize the share of treated individuals within treated clusters. 15 months later, we find that treated entrepreneurs are more likely to use mobile money savings accounts and payments, increase their mobile money and bank savings at the intensive and extensive margins, and invest more. We find little evidence of spillovers on untreated peers, but as the share of treated entrepreneurs increases, beneficial effects on the treated decline.

JEL-Codes: C930, D140, G530, O120.

Keywords: scaling, business training, financial literacy, micro-entrepreneurs, mobile money, spillover effects, saturation effects.

Jana S. Hamdan German Institute for Economic Research (DIW) Berlin / Germany jhamdan@posteo.de

Lukas Menkhoff Kiel Institute for the World Economy Kiel / Germany lukas.menkhoff@ifw-kiel.de Tim Kaiser University of Kaiserslautern-Landau (RPTU) / Germany tim.kaiser@rptu.de

Yuanwei Xu Ruhr University Bochum / Germany yuanweii.xu@gmail.com

October 21, 2024

We thank Mountains of the Moon University in Fort Portal as well as Filder Aryemo (Gaplink Uganda) and Helke Seitz (DIW Berlin) for support in the field work. We further thank Arianna Antezza, Kaja Rupieper, and Markus Rafael Paffhausen for excellent research assistance. We also thank participants of the American Economic Association Meetings 2022, the Southern Economic Association Annual Conference 2022, the Bank of Finland Financial Literacy Conference 2023, the German Development Economics Conference 2024, the VfS Annual Conference 2024, and the Cherry Blossom Financial Education Institute 2024 at Stanford University for helpful comments, in particular Tessa Bold, Adam Leive, Annamaria Lusardi, and Matthias Schündeln. This research was funded by the German Research Foundation (DFG) through the Research Training Group (RTG) 1723 "Globalization and Development." An earlier version of this research circulated as "Financial education and spillover effects: Experimental evidence from Uganda." The field experiment was preregistered under AEARCTR-0006407. Declarations of interest: none.

1 Introduction

Business- and financial education for small scale entrepreneurs is a common type of development assistance. Financial education programs cause directional changes in financial behaviors in a wide variety of settings (see Miller et al., 2015; Kaiser et al., 2022, for metaanalyses) but evidence on the causal effects of financial education programs for microentrepreneurs *operated at scale* remains limited. Additionally, the extant literature has largely focused on direct causal effects of financial education programs on the beneficiaries with limited attention into potential externalities of treatment: While governments and non-governmental development organizations often implicitly assume positive spillovers from treated to untreated individuals and businesses (i.e., knowledge diffusion), recent literature in economics has more generally studied possible externalities on untreated businesses or households in response to the transfer of physical or human capital, i.e., also allowing for potentially negative externalities (e.g., Angelucci and Di Giorgi, 2009; Haushofer and Shapiro, 2016; McKenzie and Puerto, 2021; Egger et al., 2022). The questions whether treatment effects persist if these programs are operated at scale or whether positive or negative spillovers are present are crucial to assess the net effects of a policy intervention (see Giné and Mansuri, 2018).

To study these questions in context of the financial education literature, we conduct a largescale randomized saturation experiment with 1,975 micro-entrepreneurs in Western Uganda to estimate treatment effects for several important outcomes and estimands in the context of interference ignored in the extant literature on business- and financial education interventions (see Baird et al., 2018). For this research we collect information about the location and size of all permanent trading centers in the Kabarole district of Western Uganda with about 350,000 inhabitants. We identify 108 trading centers with almost 5,500 micro-enterprises and randomize clusters to one of four assigned treatment saturations, i.e., where 0% (pure control clusters), 50%, 75%, and 100% of the sampled individuals within clusters are then randomly invited to a financial education program previously shown to generate large shifts in financial behavior among the directly treated in an earlier field experiment (Kaiser and Menkhoff, 2022). Thus, we relax the assumption of no interference and set up the experiment to identify direct and spillover effects on treated and untreated entrepreneurs in the context of partial interference.

We first report results about pooled intention to treat and spillover effects before we come to

several analyses also considering saturation (i.e., scaling) effects. The results appear heterogenous by outcome type: for most outcomes related to savings, to the use of mobile money as a savings and payments technology and to investment, the direct effects on those assigned to treatment in treated clusters are considerable. However, the spillover effects of the intervention on the untreated individuals in treated clusters are mostly estimated with a negative sign, albeit none of them being statistically significant at conventional levels.

The analysis of potential saturation effects uses three approaches: first, we distinguish between the three levels of assigned saturation; second, we instrument the *true treatment saturation* (i.e., the product of the relative frequency of sampled households relative to the clusters size and the assigned treatment saturation) with the assigned saturation and then study three groups of different degrees of true saturation; third, we study a parametric analysis of treatment saturation and focus on the slope effects of increasing the share of treated individuals. For all three approaches we find consistently that there are positive treatment effects for the lowest degree of saturation only. For higher degrees of saturation, the effects become smaller and may become even negative.

Thus, our finding is two-sided: while financial education seems to benefit those directly exposed, and does not show significant spillover effects, the beneficial effects decline as a larger share of peers are treated, so the net effects may become very small in our setting and if this program is operated at a larger scale. We interpret this finding as a specific form of "voltage drop" (Al-Ubaydli et al., 2017, 2019, 2023) where treatment effects in small scale programs are reduced due to negative externalities on the treated as the program is scaled up. We then explore alternative interpretations for causal mechanisms of this drop, including differential take-up, class size, program quality, and competitive advantages. We find a clear difference of better venues in low saturation clusters that is specific to our study. We also find some suggestive evidence of treatment effects in low-saturation clusters being driven by increased business activity at the expense of business attracted by untreated peers ("business stealing") (see also Drexler et al., 2014). As the share of treated peers increase, the room for business stealing becomes smaller and treatment effects decline.

Literature. Our study complements three strands of literature, i.e., the analysis of (1) financial- and business education programs, (2) mobile money adoption, and (3) scaling and spillover effects of educational interventions. Financial education has been recognized as an

important element of education for the whole population as evidenced by respective programs supported by the OECD worldwide (OECD, 2020). Micro-entrepreneurs in developing countries are a particular target of these efforts as improved financial behavior may not only contribute to their family welfare but may also stimulate economic growth. While the use of improved financial behavior is undisputed, there was some discussion whether financial trainings are able to realize this ambition (Fernandes et al., 2014). However, most recent evidence clearly shows that financial education, even when evaluated via randomized controlled trials (RCTs) and when corrected for publication bias, has a positive causal effect on financial knowledge and downstream financial behaviors (Miller et al., 2015; Kaiser et al., 2022). However, the challenge remains to make best use of resources by increasing effectiveness of financial education. There are several ways that have been suggested, such as using (very costly) individual counseling (Carpena et al., 2019), to rely on goal setting (Carpena et al., 2019), to use an entertainment environment (Berg and Zia, 2017), to time education at a "teachable moment" (Doi et al., 2014; Kaiser and Menkhoff, 2017) or to rely on "active learning" as a teaching method (Freeman et al., 2014; Kaiser and Menkhoff, 2022). The intervention studied here relies on the two latter elements, i.e., the training is designed to meet the purposes of these rural micro-entrepreneurs, and the way of delivery follows the concept of "active learning." The treatment effect on the uniquely treated suggests relatively large effects on treated individuals absent any spillover effects. Thus, this paper shows that this treatment is – in principle – able to change individual financial behavior which is in line with evidence from recent meta-analyses and an earlier experiment in the same setting (Kaiser and Menkhoff, 2022).

This study also adds directly to the research on mobile money adoption. The intervention provides information about mobile money as a savings-technology and about mobile money transfers (see Suri, 2017). While there is lot of evidence showing the impact of mobile money on risk sharing, household welfare and financial inclusion (e.g., Jack and Suri, 2014; Munyegera and Matsumoto, 2016; Riley, 2018; Hamdan et al., 2022), and also some evidence on more recently introduced mobile money products (Bharadwaj et al., 2021; Annan et al., 2024), there is little explicit evidence on the success of addressing mobile money in a financial education training. Chiwaula et al. (2020) combine a financial literacy and mobile money training with reminders among community savers in Malawi, finding a positive effect on mobile money savings four months after the intervention. However, they cannot differentiate the impact of the training from

the reminders and focus on a selected group whose members already save. Our results show that financial education interventions can stimulate mobile money use as a transfer and savings technology both at the extensive and intensive margins but that the spillover dynamics need to be considered when rolling out education programs.

Finally, this study contributes to literature on scaling and spillover effects of financial literacy education and human capital programs more broadly. Al-Ubaydli et al. (2017, 2019, 2023; List 2024), discuss issues of scalability, i.e., whether the financial education program can be expected to yield a consistent treatment effect if operated at a larger scale (i.e., treating a larger share of the target population). A common finding in economics is described as a "voltage drop" (Al-Ubaydli et al., 2023), i.e., a reduction in effectiveness, if programs are scaled up (e.g., Alcot, 2015; DellaVigna and Linos, 2022). Our study design allows us to analyze externalities from treated to untreated peers (conventional spillover analysis), and in addition to estimate the spillover effects on the treated, i.e., the saturation-dependent spillover effect on those invited to attend the financial education program.

Relatedly, program evaluations often only focus on the direct effects on the treated while few studies investigate the potential of positive or negative externalities which may alter conclusions on the net effects of programs substantially. While our paper is the first to provide evidence on moving financial education programs from a smaller scale to a larger scale (in terms of share of the population treated), there have been earlier attempts to address the question of externalities of financial literacy and education. Positive peer effects have been found regarding the degree of stock market participation, channeled, for example, via the family (Black et al., 2017) and communication therein (Knüpfer et al., 2023) or via knowledgeable neighbors (Haliassos et al., 2020). The latter use a natural experiment, exploiting the variation of average financial literacy in Swedish neighborhoods, to study the long-term impact on refugees, who were randomly allocated to these neighborhoods. Bruhn et al. (2016) and Frisancho (2023, 2023a) find some evidence of intergenerational spillovers to parents when children are exposed to financial education in high school. In the context of business and financial education programs for entrepreneurs, Drexler et al. (2014) conduct a post-hoc analysis of potential spillovers from a financial- and business education program randomized at the individual-level by relying on geographic proximity of treated peers. They find suggestive evidence that some of the effects on revenues may be due to

the crowding-out of other businesses in the same markets.

Closer to our design is McKenzie and Puerto (2021) who do not find spillovers from business education in their study on small female-led businesses in Kenya. The positive effects of the assignment to a training do not come along with significant spillover effects on the untreated businesses in the same markets, so that markets appear to grow as a result. Similarly, Calderon et al. (2020) find no evidence of spillovers in a small-scale experiment in Mexico studying the effects of a business and financial-literacy program. Thus, our study is among the first to identify declining treatment effects on business and household financial outcomes when the share of treated individuals increases. This result has important implications for the design and evaluation of financial education programs operated at scale.

The paper is organized as follows. In the next Section 2, we describe our study sample, the research design, the financial education program, the empirical strategy, summary statistics, balance, program take-up and attrition. Results are discussed in Section 3 and we conclude in Section 4.

2 Sample and experimental design

2.1 Study setting, sample, and timeline

Uganda is similar to many developing countries in its stage of financial development and financial inclusion: mobile money services have boosted basic access to financial accounts significantly in recent years, however, the access to traditional formal financial institutions (banks) remains low and a gender gap persists (Demirgüç-Kunt et al., 2020). The 46-million population grows at an annual rate of 3.3 percent, 46 percent of Ugandans are below the age of 15 and 75 percent live in rural areas (WDI, 2020). Most Ugandans work for themselves or for their families in low-quality jobs, and less than 10 percent are in formal employment (Merotto, 2019). Thus, policy efforts to support the growth of small businesses are needed to create jobs and improve their quality – "whether as entrepreneurs or as employees, rural youth need access to finance, technology, skills, and assets" (Merotto, 2019). Meanwhile, the Ugandan National Financial Inclusion Strategy (NFIS) 2017-2022 unites efforts to improve financial security via financial inclusion. Its goals include increasing women's use of financial services, individual emergency

savings and, more generally, consumer protection (Bank of Uganda, 2017).

In this context, we implement our study in the rural Kabarole district in the Western region of Uganda. We take this district as it was used to be known to people during baseline and still continues in several government statistics; officially, the southern part of this district, i.e., the former county of "Bunyangabu," has become a new independent district in 2017. The majority of the population resides outside the main town, Fort Portal.

Our sample is composed of small business owners from this area. They typically run their family enterprises, usually in the form of small shops, in trading centers. These are shopping streets usually lined along the main road of villages and small towns. As there is no official documentation of the existing trading centers in the district, we mapped the 113 trading centers within Kabarole with the support of local government officials and university staff in November 2018, prior to the baseline survey. The baseline survey is implemented from February to April 2019 in all the mapped villages, and a total of 108 trading centers are included in the study. Five mapped trading centers are not included as no open business was found during the baseline visit. Figure 1 displays the locations of the 108 trading centers on a map.

< Figure 1 >

During the baseline survey, we conduct face-to-face interviews with a total of 2,223 small business owners, with an average of about 20 respondents in each trading center.¹ About 62 percent of interviewed small business owners have small retail or wholesale shops, and 28 percent run a service business, mostly hair dressing salons and restaurants. The remaining 10 percent are in manufacturing, mostly in furniture making or metal goods production. We estimate that this sample covers roughly 40 percent of all small businesses in the rural areas of Kabarole, as we count 5,478 shop units that seem to be active in some way. All those with an open small business are, in principle, eligible to participate in the study.

Our sampling procedure is designed to interview a sufficiently large share of small business owners in each trading center. As smaller trading centers are typically less busy, many business

¹ Prior to the baseline survey, we piloted the interviews in Rugombe, a community in the adjacent Kyenjojo district. There, we also piloted the financial education intervention, in August 2019.

owners also work in subsistence farming in proximity to their shop and close their business for part of the day. In order to interview enough business owners in smaller trading centers, we approach every open business in trading centers with up to 100 businesses and invite them to participate in our survey (which they basically always accept). This strategy results in covering 46 percent of shops in these areas, implying that most shops (54 percent) were not open during our visiting time.

In the nine largest trading centers, i.e., those with more than 100 small businesses, we randomly approach every third open business. As here about 75 percent of shops are open, we cover about a quarter of all these shops. In total, this brings us to the outcome of interviewing 36 percent of all small business owners. Thus, in a big (small) trading center, we under-(over-)sample. Moreover, the sampling rate ranges from 20 to 100 percent across the trading centers, with an overall average of 43.6 percent (SD of 12.5 percent). To control for these varying sampling probabilities by trading center size, we will incorporate inverse sampling weights in our later analysis.

On average, a baseline survey takes about 45 minutes and is compensated with 4,900 UGX (equivalent to 1.40 USD). It collects extensive information on the financial situation and type of the businesses and the financial decision-making of the owners, particularly their saving, borrowing, transfer behaviors and the use of mobile money. Two survey experiments are incorporated into the baseline survey as well, one on the willingness to pay with mobile money for making transfers, the other measuring the risk preferences; both are not further reported in this paper.

We re-visit the 54 treatment trading centers in August and September 2019 and implement our financial education intervention, thus four to six months after the baseline. Treated baseline participants are invited in advance via phone calls. On the day of the intervention, those participants are called again prior to the beginning of the training. In addition, they are informed of and encouraged to take part in the training by the local council, elected officials who run villages (the lowest political administrative units in Uganda), on the day of the training as well. During the training, other participants who are not invited show up in the training and request to participate. We allow them to attend the training and give them the cash compensation at the end. In the ITT analysis, these participants belong to the spillover group because they are not invited. This may tentatively increase positive spillovers. The follow-up survey is conducted in two stages, about 13-19 months after the intervention (timeline in Appendix <u>Table A1</u>). The first round of collection is done using phone surveys from October to December 2020, reaching 72.9 percent of the baseline sample. Afterwards, in April 2021, we follow up with hitherto unreached participants via face-to-face interviews. On average, the follow-up survey takes on average about 30 minutes and is compensated with 6,400 UGX (equivalent to 1.80 USD 2019 PPP). Overall, we reach 90.7 percent (N = 1,975) of baseline participants.

In the baseline and the intervention phases, the study is implemented in cooperation with our mentioned partner, the MMU. The follow-up data collection is implemented by Gaplink Uganda, a Kampala-based independent research company. Their team is kept unaware of the experimental design and is provided with only the contact details of the participants and local councils. All surveys are conducted in the local language Rutooro or optional in the official language English. The responses are recorded on tablets using SurveyCTO.

2.2 Financial education program

We evaluate the effectiveness of a five-hour financial education program. It includes a total of six sessions – one introductory session and five further sessions with the following topics: (1) personal financial management (budgeting and record keeping), (2) saving, (3) debt management, (4) investment and (5) money transfer. Participants always start with the introductory session, and then immediately move from session (1) until session (5), spending up to one hour per station.

In all training sessions, the trainers employ an active learning method (Kaiser and Menkhoff, 2022), where numerous interactive activities are included in the curriculum to engage the participants in the learning process and to encourage discussions. They include, for example, picture stories, real-life case studies as exercises, small-group discussions and the evaluation of individual situations. These features promote paying attention and learning from peers. The goals of the training include, for example, learning how to budget and financially plan for themselves and their business, being able to distinguish needs from wants, promoting savings and knowing the different ways of saving, selecting and managing debt, making an investment plan and increasing business investments, and understanding the risks and costs associated with several ways to transfer money, in particular how to use mobile money (further details in Appendix <u>Table A2</u>).

The program is implemented in function halls, school classrooms or suitable outside areas in or near the center of the trading centers (or villages) in our study region. On average, the training lasts five hours, with a group consisting of 15 participants, on average (with a range between two and 26, SD of 6 participants). The training is delivered by one licensed financial literacy trainer, called the master trainer, as well as five local undergraduate students from the public Mountain of the Moon University (MMU), who act as trainers, being responsible for one session each. All trainers take part in a three-days training, where they become familiar with the curriculum, the learning goals, and the active learning method. All trainers' teaching qualifications are verified by holding mock training sessions and a pilot training. The entire team that partly alternated during the intervention includes two female and two male master trainers. It further consists of 16 female trainers, 11 male trainers, five female research assistants, and one male research assistant.

Participants of the training receive a cash payout of 10,000 UGX (equivalent to 2.80 USD) after completion of all training sessions. According to the baseline data, the average monthly income from the respondent's business and other sources is about 447,000 UGX. Thus, they earn an average of 17,880 UGX per day when assuming 25 working days per month. Therefore, the opportunity cost of attending a training is likely compensated for with the disbursed expense allowance.

2.3 Two-stage randomized saturation experiment

The randomization is conducted in two stages (see Figure 2). In the first stage, 54 out of a total of 108 clusters (trading centers) are randomly allocated to the treatment group and the other 54 are allocated to the control group. The randomization is conducted via a stratification strategy based on the mobile money account ownership rates and financial literacy levels at baseline:

< Figure 2 >

We first create three strata based on the cluster-level average share of mobile money account owners, splitting the 108 clusters into low, middle and high mobile money ownership rates, each with 36 clusters. Afterwards, within each stratum, we further create six substrata depending on the average financial literacy score.² We obtain a total of 18 strata, with six clusters within each stratum. We perform the randomization into control and treatment group in a 1:1 ratio, using a random seed in Stata.

In the second step, among the 54 treated clusters, we further randomize the cluster-level treatment intensity with probability 1/3. In 18 clusters we invite 100 percent of the baseline respondents to participate in the financial education program, in another 18 clusters we invite 75 percent, and in the remaining 18 clusters we invite 50 percent. Overall, 879 respondents are directly targeted to the take part in the training, thus 39.5 percent of the baseline sample. The treatment intensity is varied to assess possible impacts from saturation.

We register our trial and pre-analysis plan in the American Economic Association's registry for randomized controlled trials (ID AEARCTR-0006407). Blinding of participants and the trainers was not feasible. However, the endline survey was conducted by a firm being new to this study and we did not inform them about the allocation of survey participants to the different groups, so that there should be no distortion from this side.

2.4 Empirical models

As described in Section 2.3, our randomized saturation experiment allows for partial interference between units, i.e., interference of individuals within clusters but stable unit treatment values between clusters (see Baird et al., 2018). This section describes the main estimands discussed in this paper. We focus on reduced form results (Intention to Treat Effects), i.e., the effect of being assigned to treatment and spillover groups and linear spillover analyses. We estimate all models with weighted least squared using saturation and sampling weights: Observations in the pure control clusters are weighted by the inverse of the sampling probability for each cluster ($w_{i,c,t} = p_c^{-1}$). Treated individuals in treated clusters are weighted by the inverse of the product of the assigned treatment saturation (π) and the sampling probability $w_{i,c,t} = (\pi_c p_c)^{-1}$. Untreated individuals in treated clusters are weighted by $w_{i,c,t} = ((1 - \pi_c) p_c)^{-1}$.

² The two variables are chosen because participants with different baseline financial literacy levels and mobile money adoption may be significantly differently impacted by the financial education program. In order to understand possible heterogeneity in treatment effects, it is, for instance, necessary to have a sufficient number of observations of low financial literacy respondents in every group.

2.4.1 Pooled intention to treat and spillover effects

Our first specification reports pooled intention to treat and spillover effects. We relax the assumption of no interference (stable unit treatment values (SUTVA) to an assumption allowing partial interference, i.e., assuming spillovers on the treated and non-treated within clusters but not between treatment and pure control clusters. We define the following equation with the saturation weights defined above to estimate the pooled intention to treat and spillover effects, i.e., the average ITT and spillover on the non-treated effects across all saturations.

$$y_{i,c(t)} = \alpha_0 + \beta_1 T_{i,c} + \beta_2 S_{i,c} + \delta_1 y_{ic(t-1)} + \delta_2 m_{ic(t-1)} + \lambda X_{s,i,c} + \varepsilon_{ict}$$
(2)

Here, $y_{i,c(t)}$ is the outcome y of individual i in cluster c at the time of follow-up t. α_0 denotes the intercept. $\hat{\beta}_1$ captures the pooled intention to treat effect of being assigned to treatment T and $\hat{\beta}_2$ captures the pooled spillover on the non-treated within treated clusters. We control for the lagged outcome at baseline $y_{ic(t-1)}$, an indicator for missing outcome values at baseline $m_{ic(t-1)}$ as well as randomization strata dummies, enumerator dummies and a field interview dummy contained in the vector $X_{s,i,c}$. ε_{ict} denotes the error term. Standard errors are clustered at the market-level, i.e., the highest level of random assignment.

2.4.2 Non-parametric saturation analysis

To capture potential effects from varying saturation, however, we need a different analysis than before. We present in the following three kinds of analysis. First, we just differentiate between the three levels of assigned saturation, i.e. the levels of 50%, 75%, and 100%, and report effects for the treatment groups and the non-treated in treated trading centers, i.e. the spillover group.

$$y_{i,c(t)} = \alpha_0 + \beta_1 T_{i,c} \times 1(\pi_c = 50\%) + \beta_2 S_{i,c} \times 1(\pi_c = 50\%) + \beta_3 T_{i,c} \times 1(\pi_c = 75\%) + \beta_4 S_{i,c} \times 1(\pi_c = 75\%) + \beta_5 T_{i,c} \times 1(\pi_c = 100\%) + \delta_1 y_{ic(t-1)} + \delta_2 m_{ic(t-1)} + \lambda X_{s,i,c} + \varepsilon_{ict}(3)$$

Because spillovers can occur in the larger population within the cluster, we instrument the true treatment saturation (Π_c , i.e., the actual share of treated individuals within the cluster which is

partly endogenous due to non-random variation ins cluster size) with the assigned treatment saturation (π_c) (see Baird et al., 2018). In the non-parametric analysis, this results in three bins of true saturations, i.e., 1($\Pi_c \ge 12\% < 30\%$), 1($\Pi_c \ge 30\% \le 45\%$), and 1($\pi_c = 100\%$).

2.4.3 Linear saturation analysis

Because the ITT and spillover effects can be heterogenous at any given saturation and the experiment was explicitly designed to estimate slope rather than pooled effects (see Baird et al. 2018 for a discussion of this trade-off in experimental design), we next move to a discussion of linear spillover effects. Thus, we impose the functional form:

$$y_{i,c(t)} = \alpha_0 + \beta_1 T_{i,c} + \beta_2 S_{i,c} + \beta_3 (T_{i,c} \times \Pi_c) + \beta_4 (S_{i,c} \times \Pi_c) + \delta_1 y_{ic(t-1)} + \delta_2 m_{ic(t-1)} + \lambda X_{s,i,c} + \varepsilon_{ict}$$
(4)

Here, $\hat{\beta}_1$ captures the intercept of the treatment group or the *treatment effect on the uniquely treated*, i.e., the intention to treat effect on a hypothetical sole individual invited to treatment within a cluster. $\hat{\beta}_3$ is the slope effect capturing how the spillover on the treated changes at a given true saturation (Π_c) (i.e., combining the estimates of $\hat{\beta}_1$ and $\hat{\beta}_3$ yields the intention to treat at a given saturation which is the sum of the treatment effect on the uniquely treated and the spillover effect on the treated). $\hat{\beta}_2$ captures the spillover effect at saturation zero (which should be zero if spillovers have a linear relationship with the saturation).³ $\hat{\beta}_4$ is the slope effect capturing how spillovers on the non-treated within treated clusters change with increasing saturation. Due to our experimental design, only individuals sampled at baseline can be invited to participate in the program, i.e., random assignment is conditional on being sampled at baseline. Since the sampling probabilities vary with cluster size in our setting, this creates a 'gateway to treatment'. Because spillovers can occur in the larger population within the cluster, we again instrument the true treatment saturation (Π_c , i.e., the actual share of treated individuals within the cluster which is partly endogenous due to non-random variation ins cluster size) with the assigned treatment saturation (π_c) (see Baird et al., 2018).

³ Testing $\hat{\beta}_2 = 0$ serves as a test for the linearity of the spillover effect (Baird et al., 2018).

2.5 Summary statistics, attrition, and randomization balance

Overall, we evaluate the impact of our intervention on a total of 20 variables. Table 1 displays summary statistics of these outcome variables as well as selected socio-demographic characteristics to verify the orthogonality of randomization at baseline. Table 1 reveals that 64% of our sampled micro-entrepreneurs are female. Moreover, the respondents are on average 34 years old and have some secondary education.

< Table 1 >

As can be seen from Panel A of Table 1, for most variables there are no baseline differences across the three different treatment arms. Specifically, there are no significant differences in the areas of household characteristics, and there are no significant differences between targeted respondents compared to the control group across the observed characteristics. The only statistically significant imbalance appears to occur in the propensity to own bank loans where a larger share of control group respondents appear to have taken up bank loans at baseline.

Accordingly, as can be seen in Panel B of Table 2, the slope balancing tests also suggests balance for the standardized outcome indices at baseline for all outcome variables except the borrowing index. Here, there appears to be a statistically significant difference between pure control individuals and the uniquely treated (i.e., at saturation zero) in the treatment group. Yet, these differences appear to disappear with increasing saturation, indicating the pooled imbalance to be driven by few observations in low-saturation clusters. Overall, these imbalances are what can be expected to occur by chance.

Next, we note that the compliance rate upon invitation to the training among our baseline sample was 67.02%. Overall, few variables correlate with take-up of the training. Within the treated clusters, participation rates are lower on average among the more educated compared to those with fewer years of education. This may reflect that they think there is less to learn. Among those who are untargeted and thus serve as the spillover group, we find that wealth proxied by number of assets correlates negatively with training participation. This shows a tendency that those less well-off are more interested to participate in a financial training. Like McKenzie and Puerto (2021), we

find a positive coefficient of previous training participation for take-up, however, our coefficient is not statistically significant.

The endline survey tracks 90.7% of the baseline sample, making attrition relatively low between the two waves, apart by about 15 months on average. However, respondents in the treatment clusters are 3.34% less likely to attrite than respondents in the control clusters (92.2% tracking rate in the treatment clusters against 88.87% in the control clusters). To understand why differential attrition occurs, we exploit the different reasons for attrition that have been documented in the endline survey, and we find suggestions that the differential attrition is due to the differences in business survival. There are five reasons for attrition: relocation, death, decline, sickness and imprisonment, plus cases where the reason is unknown. Of these, relocation causes about 71% of the attrition cases, and only ten participants declined to participate in the endline survey (further details in Appendix <u>Table A3</u>). Those participants who relocated closed the original business as interviewed in the baseline and migrated to another region. According to the baseline data, those who relocated had lower sales initially, compared to those participated in the endline survey (details in Appendix <u>Table A4</u>).

When we regress the experimental design indicators on the attrition status and control for strata fixed effects, we find that participants in the control clusters are 3.8 percentage points more likely to attrite than those in the treatment clusters. Moreover, there is no difference in coefficients between being directly targeted or in the spillover group. Thus, the differential attrition is not driven by treatment intensity but appears to be occurring at the trading center level. Rather, the difference in attrition is almost completely driven by the difference in relocation (or business closure) by treatment status of the trading centers. Those in the treated clusters were 3.0 percentage points less likely to close their business and relocate between baseline and endline, compared to those in the control clusters. There are only minor differences in the probability of declining to participate in the endline survey across different arms, or other reasons (full regression results in Appendix Table A5). Due to the overall small differences in attrition, we do not further consider this issue in the following analysis.

3 Results

In line with the training curriculum, we present outcomes in the five areas of the treatment,

i.e., the use of mobile money, savings, investment, budgeting, and borrowing. For each treatment area we form an index, representing the outcomes of the specific area. Section 3.1 covers results on the five outcome indices plus the summary, Section 3.2 presents results on mobile money use in general and regarding transfers in particular, reflecting that this focus is relatively new in the literature. Section 3.3 shows results on savings and investment because the kind of training analyzed here was quite successful before in these respects (Kaiser and Menkhoff, 2022), and shortly reports further results on budgeting and borrowing. A discussion is provided in Section 3.4.

The tables presenting results in Sections 3.1 to 3.3 cover information in up to four panels. While Sections 3.2 and 3.3 rely on a subset of the panels (the remaining information is provided in the Appendix), Section 3.1 shows all four panels. The first panel, i.e. Panel A informs about treatment effects in treated clusters relative to control clusters, and splits covered entrepreneurs in treated clusters into those being invited to the training (the "ITT effect") and those not being invited, i.e. the potential spillover group that we are particularly interested in. The three further panels report different analyses of the potential treatment effects from varying saturation while always also considering possible spillover effects. Panel B presents the non-parametric analysis of assigned saturation, i.e. it distinguishes between the three levels of assigned saturation as planned and implemented from the beginning, at 50%, 75%, and 100%. Panels C and D focus, by contrast, on the true saturation which is a combination of the assigned saturation, the level of participation and the size of the trading center (as large centers were sampled only at one third). Effectively, true saturation ranges from 12% to 64%, so that we form three groups of about equal range, starting with 12% to 30% true saturation, etc., and analyze these in Panel C. In Panel D we apply a linear analysis of true saturation, i.e. allowing for a broad band from 1% saturation to 100%.

3.1 Effects on outcome indices

The five outcome indices, reflecting the five training areas, are formed by equally weighting the respective single outcome items. These single items are introduced in further subsections; each of them is standardized, as is the index.

Panel A. Starting with Table 2, Panel A on the conventional ITT and spillover analysis, four out of five treatment coefficients are positive; two of them, the coefficients on savings and investment, are statistically significant. By contrast, the coefficient on borrowing is negative which

partially reflects that the intended impact of the training is on responsible borrowing, so that decreasing loan amounts can be desirable. This pattern is in line with earlier experiences from a similar training in Uganda (Kaiser and Menkhoff, 2022) as well as with results from a meta-study covering RCTs on financial education (Kaiser et al., 2022). Taking all indices together in a summary index, the effect of 0.092 standard deviation units is sizable in this literature but insignificant. The effects on the spillover group tend to have negative coefficient sign, one of them (on mobile money) being marginally significant. Due to this contradiction in signs between training and spillover group, the tests of equality are mostly rejected. Overall, the training seems to work in specific outcome areas and there are no positive spillover effects. In the following panels these results are further disaggregated by considering potential effects from varying saturation.

< Table 2 >

Panel B shows results, when three treatment groups are differentiated, depending on their degree of assigned saturation. Interestingly, treatment coefficients of the first group, i.e. with an assigned saturation of 50% are larger than those of the two other groups. The spillover coefficients for this group fluctuate around zero. The treatment coefficients are larger than those from Panel A, indicating that this level of saturation provides results above average. By contrast, treatment coefficients of the second group (assigned saturation at 75%) are relatively smaller and consistently insignificant; also the spillover coefficients are smaller, that means here more negative, and one of them turns significantly negative. Relative to the results of these two groups, the results with an assigned saturation of 100% are more similar to the first group, but the summary index has about half size, i.e. effects are much smaller. Overall, treatment effects are highest with the relatively lowest degree of saturation, i.e. here 50%. For this group, effects are much stronger than on average. However, there is no clear pattern regarding the two further levels of saturation, i.e. 75% and 100%, which motivates an analysis with a finer measure of true saturation. Moreover, we instrument true saturation with assigned saturation to control for possible selection effects.

Panel C shows the results of this analysis, where we again form three groups of true saturation to allow for non-parametric effects. The treatment coefficients for the first group, representing a true saturation of between 12% and 30%, are very similar to those of the first group

shown in Panel B. The respective spillover coefficients are positive here although mostly insignificant. Very different from Panel B is the clear pattern to be recognized now for the two other groups. The group with medium true saturation shows smaller positive, but insignificant coefficients on the treatment. The last group with highest true saturation has even negative treatment coefficients, most of them being insignificant. Spillover effects are more negative in group two than in group three, but coefficients are almost completely insignificant. Overall, results mirror those of the non-parametric analysis of assigned saturation in that there is a clear positive treatment effect for the smallest degree of saturation; results differ in that there is now a clear order of effect sizes across the three groups. This motivates to complement the non-parametric with a linear analysis of true saturation.

Panel D. These results are provided in Panel D. Due to the estimation approach, the total effect – depending on the degree of saturation – is the sum of the basis effect plus the interaction term. For example, taking the case of mobile money, the calculated effect of the first person to be trained is 0.432 (i.e., with a minimal true saturation), while the effect at average true saturation, i.e. 30%, is about 0.432 minus 0.39, i.e. rather small. Higher degrees of true saturation lead to even smaller effects. While the spillover effects look also promising, coefficients are insignificant and turn already negative for true saturation of less than 20%.

Overall, the four forms of empirical analysis converge to three main messages: first, there are positive treatments effects but no positive spillover effects. Second, the pattern of treatment effects across training subjects follows the literature, with the clearest positive effects on savings and investment. Third, increasing saturation has a negative impact on treatment effects, so that – in our case – degrees of true saturation going beyond 30% may diminish effectiveness to a degree that it becomes insignificant.

3.2 Effects on mobile money

We report several outcomes of the training on the use of mobile money (MM). In general, there are at the time of our study four areas of MM use, i.e., for transfers, for payments, for saving and for borrowing. In this section we focus – in line with the organization of the training – on the general use and the two areas of transfers and payments, while we leave saving and borrowing to the respective further sections.

MM activities. Regarding general use of MM, we report the incidence and number of activities of MM use. MM is very widespread as 91.6% of the control group have an own account at endline, see column (1) in Panel A of Table 3, and the average number of MM activities used is 1.866 out of 4 (column 2). The (ITT) treatment effects are about 2% and, thus, not significant. Spillover effects are also small and negative, the one on the number of activities is significant.

< Table 3 >

Turning now to Panel B, informing about effects in three groups of true saturation, results confirm what we had seen in Section 3.1 in general, that the treatment effect is higher with low saturation: the use of MM increases by 5.5 percentage points for true saturation up to 30% (although not to a significant degree), but does not change for medium saturation and has even a negative coefficient of 5.0 percentage points for high saturation. Coming to column (2), the point estimate of the training effect for the low saturation group is 0.189, implying a marginally significant increase by about 10%, while the increase is much smaller for medium saturation and turns significantly negative for high saturation. The linear spillover analysis in Panel C confirms large effects on the uniquely treated at 12.6 percentage points and a linear slope effect of a reduction in treatment effects of 0.3 percentage points for each increasing percentage point of treatment saturation (column 1). The same is true for column 2 where we find large effects on the number of MM activities (a 23 percent increase) but a linear slope effect of 1.2 percentage points reducing the treatment effect with increasing saturation.

MM transfers. Transfers via mobile money to individuals in other places have been the historical starting point of mobile money and are the single most often used MM-product. The training discusses the costs and safety of the four main forms of transfers, i.e., via mobile money, informal transfers (friends, busses etc.), semiformal (via semiformal financial institutions) and formal transfers via banks. A training aspect of costs is, if possible, avoiding transferring very small amounts where minimum fees can amount to 20% of the transfer amount.

Comparing endline with baseline reveals an enormous increase in the use of MM-transfers. 78.5% of the control group use this service at endline, while just 30.5% did so at baseline. Moreover, MM-transfers have become the favorite form of transfer with a share of 46%, ahead of informal transfers with 40% and the two other forms mentioned above, which are relatively negligible. This big push in using MM-transfers is due to the Corona-pandemic with its contact restrictions, favoring the use of MM, but it is not due to the treatment as columns (3) and (4) in Table 3 show. Looking at the conventional ITT-effects documented in Panel A, coefficients are negative, a bit larger than 2% and far away from statistical significance, indicating that the training does not change behavior. Surprisingly, the MM-transfers also decline in the spillover group, relative to control, even to a significant extent.

Looking at Panels B and C columns (3) and (4), it is somewhat surprising that all coefficients are negative, indicating that the training may have led to a more conscious use of MM transfers. However, this reasoning does not explain why the spillover coefficients are even more negative and for the low saturation group significant. It may also be possible that these results mainly indicate zero effects, partially created by a generally large increase of MM transfers due to the Covid-19 pandemic.

MM payments. The case is different again for payment services. While MM transfers to other private individuals have been the starting point for MM services in general, payment services were introduced later, and thus some of these services are less widespread in use (Annan et al., 2024). They always refer to various business transactions and come at lower transaction costs. Examples include paying for water, deliveries at businesses, or customers paying their goods and services at shops. Any use of payment services is documented in Table 3 column (5) and customers paying with MM, i.e. the perspective from the micro-entrepreneurs, is shown in column (6). Treatment effects are positively significant in general (at the 10% level) and positive but insignificant for the spillover group (see Panel A). Known patterns apply to the saturation groups (see Panel B). Effects are relatively strong for the low saturation group, with increases by more than 10% each. Coefficient sizes decrease by roughly one third in the medium saturation group and turn negative in the high saturation group. Again, this pattern is confirmed by a linear model of treatment saturation (Panel C, column 1) where treatment effects on the uniquely treated are relatively large but linearly decreasing with increasing saturation.

Overall, the treatment does increase the intensified use of MM services, not at least for the relatively new service of MM payments, while MM transfers seem to remain unaffected (at a high level). Significant positive effects stem only from the low saturation group, effects become much

smaller for medium saturation and tend to be negative for high saturation. Spillover effects mainly mirror this pattern, so that there is nothing specific to learn from them.

3.3 Effects on savings and investment

Meta-studies have shown that financial education tends to be successful in the domain of savings behavior but produces quite mixed results in the domain of borrowing behavior (Kaiser et al., 2022). This pattern is confirmed by an earlier RCT with the same content and approach, also in Western Uganda, but with a different target group, i.e. vendors offering mainly agricultural products in on-off-markets. Kaiser and Menkhoff (2002) find positive treatment effects on savings and investment behavior, but no effects on three further domains, i.e., record keeping / budgeting, borrowing, and transfers. Therefore, we present further results split into these two groups, starting with savings and investment.

Savings. The treatment may create slight increases in having any savings, see Table 4, Panel A, column (1), and in the amount of savings (column 2), as the positive but insignificant coefficients indicate. The increases are much stronger for MM savings (columns 3 and 4) and bank savings (columns 5 and 6) only, both at the intensive and extensive margin. These latter increases are in the order of magnitude of more than 20%. Specifically, MM savings incidence increases by 31%, MM savings amounts by 27%, bank savings by 29% at the extensive margin and about 27% at the intensive margin relative to control. These large increases of MM and bank savings in combination with small general saving increases imply that there occurs a shift in relative importance towards formal savings, which is intended by the treatment. Panel B of Table 4 shows that these positive effects are again driven by the group of low saturation, while effects at the group with high saturation are tiny in general. Spillover effects have typically negative coefficient signs but remain insignificant. The linear saturation analysis confirms the non-parametric results: We find very strong effects on the uniquely treated and a strong linear decline of MM and bank savings with increasing saturation (columns 3 to 6, Panel C).

< Table 4 >

Investment. In line with an emphasis on stimulating longer-term economic success, one

treatment area addresses the productive role of increased investments. The treatment indeed causes increased incidence and amounts of investments. The effects of the treatment are statistically significant and economically substantial, as incidence increases by 6% and amount by 5%, respectively, at a time of economic crisis (Table 4, Panel A, columns 7 and 8). As with MM savings or other savings, the spillover effects are insignificant with a negative coefficient sign.

The results of the saturation analysis show a similar pattern as with savings (see Table 4, columns 7 and 8, Panel B and C). The coefficients of those assigned to training in the low and the saturation groups are positive and significant. It is only the high saturation group where treatment coefficients are estimated with a negative sign. The coefficients of the spillover groups decline with increasing saturation: they are small and positive for low saturation, small and negative for medium saturation, and larger and significantly negative for high saturation. The linear spillover analysis confirms this pattern qualitatively, albeit with large standard errors.

3.4 Effects on further outcomes

As effects on transfers, savings and investment have been shown in Sections 3.2 and 3.3 above, we now turn to the remaining training areas of budgeting (personal financial management) and borrowing. As there are no clear pattern to report, we present results in the Appendix <u>Tables A6 and A7</u>. The point estimates on "keeping records" and "separating money for business and personal finance" are positive, but small and not statistically significant, aside from rather marginal issues. Probably the intended change in behavior needs repetition, and not just a one-hour training, to exercise the lessons learned so that a repeated training might be able to create sustained change. Regarding the debt area, the focus is mainly on avoiding unproductive and too expensive debt. Thus, the intention is neither a general increase or decrease of debt, but a qualitative change to responsible borrowing behavior, such as a shift towards cheaper formal loans. Coefficients point indeed into this direction, as total loans (incidence and amounts) slightly decrease but formal loans increase, with even higher point estimates for the spillover group; however, again the effects are too small to be likely precisely estimated in our sample with limited power.

3.5 Discussion

We structure this discussion according to the three effects of the financial education training

being distinguished so far.

Direct effects. The direct effects of this training can be compared to the literature. It is wellknown that some training areas of financial education generate better results than others. In particular, the saving education works well, while training on borrowing behavior provides mixed results (Kaiser and Menkhoff, 2017). This is largely, what we find, too. Moreover, the training curriculum here is very similar to the one of Kaiser and Menkhoff (2022), and the pattern of outcome effects is also very similar. Still, the degree of effectiveness may be somewhat smaller, possibly because the economic environment was so volatile due to Covid-19 and the related long business and school closures, which may overlay treatment effects and increase heterogeneity (standard errors). New is the focus on mobile money, where training effects on use of MM, saving via MM, and MM payments work well, while the area of transfer behavior does not create clear effects.

Spillover effects. Coming to the second group of effects, i.e. spillovers in treated clusters from the invited to the non-invited, results are largely neutral, in a few cases they are even slightly negative and significant. This seems also to be in line with the literature. The positive spillovers being documented in the literature seem to apply to studies covering the private domain of individuals. For example, individuals buy stocks because they learn this from neighbors, or parents learn from their children receiving financial education at school. However, if the business domain is covered in studies, there may be even negative spillover effects as the success of one small enterprise may come at the expense of a competing entrepreneur who does not receive a training. Our study clearly falls in the business domain.

Saturation effects. Finally, the analysis of saturation effects appears to be (relatively) new to the literature on financial education. Here we find consistently negative effects which may be explained analogously to the spillover discussion. If a single entrepreneur is financially educated this creates a competitive advantage, but if all are educated, this increases the level of education for all but does not create by itself an expansion of the market or of productivity. We will test this and other potential mechanisms in the subsequent Section 4.

This result is not in contradiction to Angelucci and De Giorgi (2009) or Egger et al. (2022) examining general equilibrium effects of large cash transfers. These transfers are large enough to increase demand, and therefore to increase jobs. In our case, however, the increase is in knowledge,

and it is not easy to transform this into real output.

4 Exploring mechanisms behind saturation effects

We explore four potential mechanisms that may contribute to the negative saturation effects we find in our study. Three of them argue that increasing saturation may decrease various aspects of the quality of the training so that its effectiveness declines, while the fourth mechanism argues that training effects are an individual phenomenon in the first place that does not necessarily aggregate at the market level. Increasing saturation may reduce the competitive advantage one can gain from participating at the training. Overall, the decreasing effectiveness of the training coming with higher saturation may be (partially) caused by (i) negatively selected take-up, (ii) larger class sizes giving less room to micro-entrepreneurs to bring in their individual issues in the training, (iii) less quality of the trainings due to various channels, and (iv) declining competitive advantage from superior knowledge and behavior.

Differences in take-up. In low saturation clusters, only few are invited and those will attend the training who have an interest in doing so. However, in high saturation clusters, some may takeup the program only because their peers are attending the session. They will attend the sessions, but this group may be relatively more inattentive. We test this hypothesis in two ways: first, we analyze take-up across the three saturation levels. Evidence presented in Table 5, column (1), shows that take-up is basically the same for all saturation levels. Second, we check a possible later reaction, i.e. whether there are differences in the degree that participants leave the training before it is finished. Results in column (2) provide a slight hint that something may be less attractive in trainings with high saturation, but this is not due to a higher take-up. This difference is also not due to a different composition of participants vs. non-participants (details in Appendix <u>Table A8</u>). Overall, differential take-up does not seem to be a mechanism that could explain the lower effectiveness of treatments with higher saturation.

< Table 5 >

Differences in class size. It has been often argued in the education literature that class size

has a tentatively negative effect on teaching effectiveness (Chetty et al., 2011). The reason is straight forward that the teacher is less able in larger classes to address needs of individual participants so that they learn less. An increasing saturation may lead to larger class size if the increasing number of training participants is not fully compensated by an increasing number of training groups. Thus, we test whether there is a relation between class size and saturation. Results shown in Table 5, column (3) show that there is indeed some variation in class size but there is no significant relation to the saturation levels, so that class size does not explain declining effects with higher saturation.

Differences in training quality. Another mechanism could be that trainings occur with less quality in higher saturation locations which will tentatively reduce the effectiveness of learning. We test this potential mechanism by a respective classroom-observation. There are ten items asking for various quality aspects of the training, such as the overall quality of the stations, the motivation of the teacher, and participant engagement. Moreover, another item asks specifically whether the venue of training is fully adequate. Regarding quality measures we aggregate the ten items into a single measure, because the single item responses are always and consistently very positive, i.e. at a level of about 9 on a scale from 0 to 10. Table 5 column (4) shows that there is no relation between this variable of quality rating and the degree of saturation, indicating that a general quality mechanism does not apply here.

However, there is a clear relation between venue adequacy and saturation as adequacy is much lower rated in higher saturation groups. While 75% of participants rate the venue as fully adequate in group one, this share declines by 30 percentage points to about 45% only in the two other groups. As there are no deeper questions asking about reasons for inadequacy, we have geo-coded all training locations and see that outdoor venues may be less appreciated: their share is 56% in group one, but 66% in group two and even 87% in group 3. Finally, it seems plausible that a dissatisfaction with venue quality is a reason why 4% of group 3 left early, relatively to 1% for group one. In summary, while the choice of venue did not appear to influence the quality rating of the training, our results suggest that part of the declining treatment effects at higher saturations may be due to institutional constraints in our specific context (i.e., difficulty of finding adequate indoor venues to operate the financial education program at scale).

Differences in competitive advantage. A final mechanism we analyze is the potential effect

of the training on the treated and the resulting competitive advantage over other microentrepreneurs. While it seems desirable that the treated improve their behavior and thus contribute to overall economic development, it seems also possible that treatment effects in low saturation clusters are large, because they are driven by "business stealing." The argument is that the treated do indeed improve behavior by more mobile money use, saving and investing, which – for example – may increase their stock of products in a retail shop and thus may attract additional customers. However, if total demand for these products does not increase, the success of some shops comes at the expense of those shops that do not improve. Thus, seen from the pure entrepreneurial perspective, it seems crucial that "the market" increases with the training.

Due to data limitations, we do not know market sizes nor more details about the three sectors of micro-enterprises. Still, it seems plausible that enterprises in retailing, which form the largest sector with a share of 60%, and which often offer very similar products (homogenous goods), have more problems to gain a competitive advantage and increase a market than enterprises that offer more distinct services and products (differentiated products). It follows that trainings improving behavior of retail business may be expected to not be much affected by the degree of saturation. The situation may be very different in the service and manufacturing sectors: there are generally less enterprises of these sectors in one trading center due to their smaller numbers, and these enterprises are more differentiated, as they cover a broad range. Consequently, they have a very limited number of competitors in a specific trading center (while retailers always have many competitors). A low degree of saturation then translates into a situation where possibly just one specific enterprise gets treated, whereas a single or a few competitors remain non-treated. If saturation increases, competing enterprises get treated too, so that a competitive advantage from exclusive treatment in low saturation trading centers disappears with high saturation.

We test this hypothesis by distinguishing treatment effects between retail enterprises and others (enterprises from services and manufacturing are combined because of small numbers). Results in Table 6 show a difference between the two groups: saturation effects are indeed strong in services and manufacturing, because only group one profits from the training and the coefficients for groups two and three are consistently negative and coefficient sizes in group 3 are quite large, although always insignificant for groups two and three. The situation is very different in retailing because eight out of nine coefficients have a positive sign and the decline in effect size from group

one to group three is moderate. Thus, compared to enterprises in services and manufacturing, the effects for group one are smaller, but remain more stable for higher degrees of saturation. These results provide evidence being consistent with a competition mechanism.

< Table 6 >

Discussion. Overall, evidence presented above sheds some light on mechanisms that can or cannot explain (part of) the saturation pattern in this study. Evidence indicates that neither differential take-up nor class size or general training quality help to understand the saturation effects we find. However, there are two mechanisms that seem to be relevant: first, venue quality in higher saturation groups is assessed less than in group one potentially reducing training effectiveness. Second, the declining effectiveness in trading centers with higher saturation seems to be driven mainly by enterprises in services and manufacturing. These enterprises may profit more from financial education if they are treated exclusively, implying that they profit less if competing enterprises get the same training. This effect of competitive advantage shows up as high difference between positive effects in low saturation locations and hardly any effects in high saturation locations.

5 Conclusion

The lack of financial knowledge and skills is an obvious constraint in the development of micro-enterprises. Consequently, policy invests into trainings to (partially) overcome this constraint, and financial education is typically part of these efforts. The literature on financial education has focused on the direct effects of these treatments and finds significant positive effects in general (Miller et al., 2015; Kaiser et al., 2022).

These reduced form effects are usually estimated in small-scale settings. However, policy is not only interested in effects on the treated (and invited non-participants) but also on potential effects if the training is scaled up. Thus, our research studies two potential effects which may occur in such a situation: first, spillovers may occur from treated to non-invited in treated areas which would ease the scaling in case of positive spillovers. Second, there may be saturation effects, i.e. the dependence of treatment effectiveness on the share of treated to the total population in the treatment area, and scaling up would be eased if such effects are either positive or neutral. Thus, we analyze in total three estimands, i.e., the conventional direct effect of financial education, a spillover effect on the non-invited in treated areas, and a saturation effect estimating whether an increasing share of the treated impacts the estimated treatment effects on treated entrepreneurs. We find that the direct effects are often positive, the spillover effects seem to be largely neutral, and the saturation effects tend to be negative in our setting.

This extends the literature in several ways. Our underlying RCT provides for another case that even a relatively short financial training can change behavior, also covering the use of mobile money. This is the basis to show that scaling this treatment – and keeping the degree of effectiveness – is not an easy task. While there is hardly any spillover on the non-treated in the general setting, the case is different when we consider effects on the treated at different degrees of saturation. We find that higher saturation can have a negative effect on behavior change.

Regarding policy implications of these results, we first note that the evidence comes from one RCT, and thus would obviously profit from further investigations. Moreover, we analyze impacts among micro-entrepreneurs' wo are competing against each other to some extent, so that positive spillovers reported in the literature when treating individuals in their private domain are no contradiction to the results here. The saturation result seems to fit into the literature, that any positive average effect estimated in a small-scale RCT cannot be assumed to be fully scalable. Thus, as a first policy implication, it seems useful to measure and then consider treatment effects beyond the direct effects.

This holds in particular regarding entrepreneurs where it could be important to focus on trainings that do not only improve behavior of single entrepreneurs but improve the situation of the overall economy in treated regions. We do not want to downplay the useful effects trainings such as the one analyzed here can have just at the individual level, when safety of savings and transactions goes up due to a higher share of formal financial services or when higher savings allow a better buffering of adverse shocks. This tends to improve these entrepreneurs' situation. Still, it seems important to also have the objective of increasing markets in mind. Then financial trainings should also contribute to transforming additional savings and investments into growth in treated regions in the one or other way.

References

Al-Ubaydli, O., Lai, C.-Y., and List, J.A. (2023). A simple rational expectations model of the voltage effect. NBER WP 30850.

Al-Ubaydli, O., List, J.A., LoRe, D., and Suskind, D. (2017). Scaling for economists: Lessons from the non-adherence problem in the medical literature. *Journal of Economic Perspectives*, 31 (4):125–144.

Al-Ubaydli, O., List, J.A., and Suskind, D. (2019). The science of using science: Towards an understanding of the threats to scaling experiments. NBER WP 25848.

Angelucci, M. and De Giorgi, G. (2009). Indirect effects of an aid program: How do cash transfers affect ineligibles' consumption? *American Economic Review*, 99(1):486–508.

Angelucci, M. and Di Maro, V. (2016). Programme evaluation and spillover effects. *Journal of Development Effectiveness*, 8(1):22–43.

Annan, F., Cheung, C., and Giné, X. (2024). Digital payments. *Oxford Review of Economic Policy*, 40(1):118–128.

Apiors, E.K. and Suzuki, A. (2022). Effects of mobile money education on mobile money usage: Evidence from Ghana. *European Journal of Development Research*, doi:10.1057/s41287-022-00529-x.

Baird, S., Bohren, J.A., McIntosh, C., and Özler, B. (2018). Optimal design of experiments in the presence of interference. *Review of Economics and Statistics*, 100(5):844–860.

Banerjee, A.V., Cole, S., Duflo, E., and Linden, L. (2007). Remedying education: Evidence from two randomized experiments in India. *Quarterly Journal of Economics*, 122(3):1235–1264.

Bank of Uganda (2017). National Financial Inclusion Strategy 2017-2022. https://www.bou.or.ug/bou/bouwebsite/bouwebsitecontent/publications/ special_pubs/2017/National-Financial-Inclusion-Strategy.pdf [Online; accessed 30-November-2020].

Bank of Uganda (2021). Mobile Money Statistics. Data and Statistics. https://www. bou.or.ug/bou/bouwebsite/PaymentSystems/dataandstat.html [Online; accessed 1- October-2021].

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.

Bharadwaj, P., Jack, W., and Suri, T. (2021). Fintech and household resilience to shocks: Evidence from digital loans in Kenya. *Journal of Development Economics*, 153, 102697.

Black, S., Devereux, P., Lundborg, P., and Majlesi, K. (2017). On the origins of risk-taking in financial markets. *Journal of Finance*, 72(5): 2229–2277.

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

Calderon, G., Cunha, J.M., and De Giorgi, G. (2020). Business literacy and development: Evidence from a randomized controlled trial in rural Mexico. *Economic Development and Cultural Change*, 68(2): 507–540.

Carpena, F., Cole, S., Shapiro, J., and Zia, B. (2019). The ABCs of financial education:

Experimental evidence on attitudes, behavior, and cognitive biases. *Management Science*, 65(1):346–369.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., and Yagan, D. (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *Quarterly Journal of Economics*, *126*(4), 1593-1660.

Chiwaula, L., Matita, M., Cassim, L., Agurto, M., et al. (2020). Combining financial-literacy training and text-message reminders to influence mobile-money use and financial behavior among members of village savings and loan associations: Experimental evidence from Malawi. Partnership for Economic Policy Working Paper No. 2020-10.

Demirgüç-Kunt, A., Klapper, L., Singer, D., Ansar, S., Hess, J., et al. (2020). The Global Findex Database 2017: Measuring financial inclusion and opportunities to expand access to and use of financial services. *World Bank Economic Review*, 34(Supplement 1):S2–S8.

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.

Eggers, D., Haushofer, J., Miguel, E., Niehaus, P., and Walker, M. (2022). General equilibrium effects of cash transfers: Experimental evidence from Kenya. *Econometrica*, 90(6): 2603–2643.

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.

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.

Frisancho, V. (2023). Is school-based financial education effective? Immediate and long-lasting impacts on high school students. *Economic Journal*, 133(651):1147–1180.

Frisancho, V. (2023a). Spillover effects of financial education: The impact of school-based programs on parents. *Journal of Financial Literacy and Wellbeing*, 1(1):138–153.

Giné, X. and Mansuri, G. (2018). Together we will: Experimental evidence on female voting behavior in Pakistan. *American Economic Journal: Applied Economics*, 10(1):207–235.

GSMA (2021). State of the Industry Report on Mobile Money 2021. https://www.gsma. com/sotir/.

Haliassos, M., Jansson, T., and Karabulut, Y. (2020). Financial literacy externalities. *Review of Financial Studies*, 33(2):950–989.

Hamdan, J.S., Lehmann-Uschner, K., and Menkhoff, L. (2022). Mobile money, financial inclusion, and unmet opportunities. Evidence from Uganda. *Journal of Development Studies*, 58(4):671–691.

Haushofer, J. and Shapiro, J. (2016). The short-term impact of unconditional cash transfers to the poor: Experimental evidence from Kenya. *Quarterly Journal of Economics*, 131(4):1973–2042.

Haushofer, J. and Shapiro, J. (2018). The long-term impact of unconditional cash transfers: Experimental evidence from Kenya. *Busara Center for Behavioral Economics, Nairobi, Kenya*.

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.

Kaiser, T., Lusardi, A., Menkhoff, L., and Urban, C. (2022). Financial education affects financial knowledge and downstream behaviors. *Journal of Financial Economics*, 145:255-272.

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

Kaiser, T. and Menkhoff, L. (2022). Active learning improves financial education. *Journal of Development Economics*, 157, 102870.

Knüpfer, S., Rantapuska, E., and Sarvimäki, M. (2023). Social interaction in the family: Evidence from investors' security holdings. *Review of Finance*, 27(4):1297–1327.

List, J.A. (2024). Optimally generate policy-based evidence before scaling. *Nature*, 626(7999), 491–499.

McKenzie, D. and Puerto, S. (2021). Growing markets through business training for female entrepreneurs: A market-level randomized experiment in Kenya. *American Economic Journal: Applied Economics*, 13(2):297–332.

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

Munyegera, G.K. and Matsumoto, T. (2016). Mobile money, remittances, and household welfare: Panel evidence from rural Uganda. *World Development*, 79:127–137.

OECD (2020). Recommendation of the Council on Financial Literacy, OECD/LEGAL/0461.

Riley, E. (2018). Mobile money and risk sharing against village shocks. *Journal of Development Economics*, 135:43–58.

Suri, T. (2017). Mobile money. Annual Review of Economics, 9:497–520.

UBOS (2021). The population projection of Kabarole district from 2015-2020. Ugandan Bureau of Statistics, Kampala, Uganda. http://catalog.data.ug/ dataset/uganda-s-district-population-projections-2015-2020/resource/ 904a83cf-3735-4d98-ac02-8fd6188b25a4[Online; accessed 1-December-2021].

WDI (2020). World Development Indicators. DataBank. Washington, D.C.: The World Bank. https://databank.worldbank.org/reports.aspx?source=2&country=UGA [Online; accessed 1-December-2021].

Figure 1: Study setting in Western Uganda



Notes: The figure shows in (a) the location of the Kabarole district in Western Uganda, and in (b) the distribution of the 108 sampled trading centers in this district by assigned treatment saturation.





Notes: This figure describes the levels of randomization two-stage randomized saturation experiment showing the Number of clusters (trading centers, TCs) and individuals (micro-entrepreneurs, MEs) in each treatment arm. We provide details on the share of MEs relative to the entire population within each arm and the variation (standard deviation, minimum and maximum) of this metric across clusters within each arm.

Table 1: Balance at baseline in the endline sample (n=1,975)

Panel A: Pooled means and	standard deviations				
		Pure control group (n=862)	Spillover group (n=320)	Assigned to training (n=793)	Equality of means (p)
Female (0/1)		0.638	0.661	0.615	0.702
Age		33.681	34.886	34.554	0.108
		(11.390)	(11.507)	(11.803)	
Years of Education		8.957	9.132	8.826	0.332
H 1.110		(4.551)	(4.650)	(4.613)	0.774
Household Size		3.992	4.065	4.104	0.776
No of assets		(2.424)	(2.555)	(2.479)	0.847
10. 01 assets		(18 213)	(17,179)	(17 558)	0.047
Household Consumption (UGX)	506.097	517.887	505.608	0.665
1	,	(352,976)	(342,138)	(333,849)	
Mobile Money Index					0.998
Any Mobile Money (MM) a	ctivity (0/1)	0.508	0.306	0.518	0.611
# MM Activities (0-4)		0.705	0.433	0.696	0.964
		(0.807)	(0.731)	(0.782)	
Any MM Transfers (0/1)		0.305	0.177	0.292	0.799
MM Transfers Amount (IHS	ST)	3.716	2.102	3.474	0.613
		(5.674)	(4.570)	(5.465)	0.450
Any MM payment service	used (0/1)	0.300	0.311	0.334	0.473
Customers paying with MM	in past 3m (share)	0.007	0.008	(0.007)	0.627
Saving Index		(0.057)	(0.000)	(0.040)	0.834
Any Savings (0/1)		0.802	0.456	0.783	0.736
Savings Amount (IHST)		10.688	6.078	10.498	0.780
5		(5.507)	(6.733)	(5.686)	
Any MM Savings (0/1)		0.100	0.063	0.073	0.191
MM Savings Amount (IHS)	Γ)	1.188	0.713	0.905	0.290
		(3.607)	(2.835)	(3.243)	
Any Bank Savings (0/1)		0.154	0.108	0.176	0.373
Bank Savings Amount (IHS	T)	2.152	1.561	2.479	0.401
		(5.151)	(4.512)	(5.403)	0.524
Investment Index		0.872	0.521	0.860	0.324
Any Investment $(0/1)$		12 372	7.407	12 354	0.703
Investment Amount (IHS1)		(4 890)	(7.174)	(4 963)	0.090
Budgeting Index		(1.050)	(/.1/1)	(11905)	0.524
Keeps records (0/1)		0.233	0.140	0.211	0.377
Separate records business ar	nd personal (0/1)	0.248	0.138	0.233	0.796
Borrowing Index					0.771
Any Loan (0/1)		0.356	0.196	0.343	0.818
Any MM Loan (0/1)		0.001	0.000	0.004	0.244
Any Bank Loan (0/1)		0.054	0.037	0.041	0.041**
Total Loan Amount (IHST)		10.669	10.017	10.175	0.746
Panel B: Slone balance test		(5.529)	(5.843)	(5.824)	I
1 aner D. Stope bulance test	(1)	(2)	(3)	(4)	(5)
	(1) Mo Money	(2) Savings	(J)	(+) Budgeting	Borrowing
Assigned to Training	0.061	0.042	0.140	0.062	0.262***
Assigned to Training	0.001	0.042	-0.140	-0.062	-0.362
	(0.1/1)	(0.109)	(0.202)	(0.155)	(0.122)
Assigned to Training ×	-0.002	-0.002	0.004	0.001	0.011
Π_c (% True Saturation)	(0.005)	(0.003)	(0.006)	(0.004)	(0.004)
Spillover Group	-0.245	-0.288	-0.186	-0.142	-0.060
	(0.297)	(0.210)	(0.217)	(0.297)	(0.249)
Spillover Group \times	0.010	0.013	0.005	0.004	-0.000
Π_c (% True Saturation)	(0.012)	(0.009)	(0.009)	(0.013)	(0.010)
/					

Notes: Panel A displays the summary statistics for the control group and by individual treatment status for the endline estimation sample (N=1,975) at baseline in 2019. In addition, p-values for the equality of means are shown. Panel B shows slope balance tests probing saturation-dependent balance at baseline for the standardized outcome indices. The true saturation (Π_c) is instrumented with the assigned saturation (π_c). All regressions are estimated with design weights, and strata fixed effects. Standard errors are clustered at the trading center level (i.e., level of random assignment). *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Treatment effects on standardized outcome indices

	(1)	(2)	(2)	(4)	(5)	(6)		
	(1) Mo Money	(2) Savings	(<i>3)</i> Investment	(4) Budgeting	Borrowing	(0) Summary		
Panel A: Pooled Intention to Trea	and Spillover Fi	fects	mvestment	Dudgeting	Donowing	Summary		
1 unei 11. 1 obieu intention to 11 eu	una spinover <u>D</u>	jeets						
Assigned to Training	0.040	0.149***	0.103^{*}	0.037	-0.045	0.092		
	(0.073)	(0.054)	(0.054)	(0.061)	(0.046)	(0.058)		
Spillover Group	-0.135*	-0.041	-0.021	-0.037	-0.077	-0.101		
	(0.076)	(0.060)	(0.063)	(0.065)	(0.053)	(0.064)		
Test of equality $(p - value)$	0.031**	0.004^{***}	0.034**	0.183	0.499	0.002***		
Panel B: Non-Parametric Analysis of Assigned Saturation								
Assigned to Training X	0.183*	0 277***	0.131*	0.053	-0.048	0 193**		
$1(\pi = 50\%)$	(0.103)	(0.078)	(0.078)	(0.080)	(0.060)	(0.080)		
$r(n_c = 5070)$ Spillover Group X	-0.054	0.008	0.067	0.045	-0.063	0.000		
$1(\pi - 50\%)$	(0.094)	(0.008)	(0.068)	(0.043)	-0.003	(0.001)		
$1(n_c = 50\%)$	(0.084)	(0.070)	(0.008)	(0.002)	(0.039)	(0.009)		
Assigned to Training \times	-0.054	0.027	0.046	0.006	-0.006	0.006		
$1(\pi_c = 75\%)$	(0.095)	(0.068)	(0.055)	(0.066)	(0.059)	(0.074)		
Spillover Group X	-0.223*	-0.084	-0.131	-0.118	-0.084	-0.207**		
$1(\pi = 75\%)$	(0.125)	(0.089)	(0.090)	(0.097)	(0.075)	(0.101)		
	(0.120)	(0.005)	(0.03 0)	(0.037)	(0.070)	(01101)		
Assigned to Training \times	-0.014	0.151**	0.131*	0.084	-0.055	0.097		
$1(\pi_c = 100\%)$	(0.082)	(0.062)	(0.077)	(0.081)	(0.052)	(0.071)		
Panel C: Non-Parametric Analysis	s of True Saturati	ion						
	0.102	0 01 4**	0 104**	0.0(0	0.002	0 222**		
Assigned to Training X	0.193	0.214	0.194	0.069	-0.002	0.232		
$1(\Pi_c \ge 12\% < 30\%)$	(0.128)	(0.087)	(0.080)	(0.084)	(0.069)	(0.102)		
Spillover Group X	0.061	0.067	0.207	0.108	-0.052	0.145		
$1(\Pi_c \ge 12\% < 30\%)$	(0.135)	(0.114)	(0.115)	(0.097)	(0.083)	(0.118)		
Assigned to Training X	0.033	0.183	0.144	0.138	-0.106	0.132		
$1(\Pi_{2} > 30\% < 45\%)$	(0.143)	(0.121)	(0.143)	(0.170)	(0.119)	(0.133)		
Spillover Group X	-1 535	-1 243	-1 767	-1 523	-0.292	-2 223		
$1(\Pi > 30\% < 45\%)$	(1419)	(1.089)	(1.343)	(1.141)	(0.680)	(1.498)		
$1(11_{c} \ge 30.00 \ge 13.00)$	(1.11))	(1.00))	(1.5 15)	(1.111)	(0.000)	(1.190)		
Assigned to Training \times	-0.235*	-0.025	-0.043	-0.012	-0.121*	-0.174		
$1(\Pi_c > 45\% \le 64\%)$	(0.126)	(0.116)	(0.124)	(0.141)	(0.071)	(0.134)		
Spillover Group ×	-0.250	0.191	-0.432**	-0.056	0.209	-0.059		
$1(\Pi_c > 45\% \le 64\%)$	(0.281)	(0.249)	(0.193)	(0.178)	(0.188)	(0.252)		
Panel D: Linear Analysis of True	Saturation			\$ 7	```			
Assigned to Training	0.432*	0 447***	0 220	0.071	-0.029	0.370**		
Thoughou to Trunning	(0.229)	(0.159)	(0.159)	(0.158)	(0.120)	(0.180)		
Assigned to Training X	-0.013*	-0.010**	-0.003	-0.001	-0.001	-0.009*		
Π_{-} (% True Saturation)	(0.006)	(0.005)	(0.005)	(0.001)	(0.003)	(0.005)		
n _c (/o mue Suturation)	(0.000)	(0.005)	(0.000)	(0.005)	(0.003)	(0.000)		
Spillover Group	0.295	0.227	0.532**	0.393	-0.055	0.451*		
	(0.339)	(0.263)	(0.228)	(0.246)	(0.190)	(0.258)		
Spillover Group ×	-0.016	-0.010	-0.022**	-0.017*	-0.001	-0.021**		
Π_c (% True Saturation)	(0.014)	(0.010)	(0.009)	(0.011)	(0.008)	(0.011)		

Notes: The dependent variables in column 1-6 are all standardized to have a mean of 0 and a standard deviation of 1 in the pure control group. Panel A shows intention to treat, and spillover effects pooled across all saturations (see Figure 2 for the distribution and support of the saturations). Panel B shows a non-parametric analysis of intention to treat and spillover effects at each randomly *assigned saturation*. Panel C shows a non-parametric analysis of intention to treat and spillover effects for three bins of *true saturations* ($\Pi_c \ge 12\% < 30\%$), ($\Pi_c \ge 30\% \le 45\%$), ($\Pi_c > 45\% \le 64\%$). In this analysis, the true saturation (Π_c) is instrumented with the randomly assigned saturation (π_c). Panel D shows results of a parametric analysis of treatment saturation, i.e., an affine model of how intention to treat and spillover effects, a phone-survey (relative to field) indicator, and controls for sectors. Standard errors are clustered at the trading-center level, i.e., the level of random assignment. N=1,975. *** p<0.01, ** p<0.05, * p<0.1.

	(1) Any MM activity	(2) Number of MM activities in (max=4)	(3) Any MM transfers (0/1)	(4) MM Transfers Amt. †	(5) Any MM payment service used (0/1)	(6) Customers paying with MM in past 3m (share)
Control group mean (SD)	0.916	1.866 (0.884)	0.785	10.031(5.501)	0.797	0.037 (0.099)
Panel A: Pooled Intention to	o Treat and Sp	illover Effects				
Assigned to Training	0.020 (0.018)	0.063 (0.056)	-0.024 (0.030)	-0.278 (0.390)	0.048* (0.026)	0.012* (0.007)
Spillover Group	-0.012	-0.123*	-0.090**	-1.068**	0.012	0.003
Test of equality (r. value)	(0.020)	(0.064)	(0.035)	(0.468)	(0.026)	(0.007)
Panel B: Non-Parametric A	0.112 nalysis of True	0.000 Saturation	0.031	0.077	0.204	0.226
Tunet D. Non-Turumetric A		0.100*	0.007	0.105	0 10 1**	0.00**
Assigned to Training \times	0.055	0.189	-0.006	-0.105	0.104	0.020
$1(\Pi_c \ge 12\% < 30\%)$	(0.035)	(0.102)	(0.044)	(0.586)	(0.053)	(0.010)
Spillover Group \times	0.035	0.068	-0.076	-1.059	0.114	0.012
$1(\Pi_c \ge 12\% < 30\%)$	(0.043)	(0.124)	(0.046)	(0.612)	(0.058)	(0.011)
Assigned to Training X	-0.001	0.065	-0.054	-0.695	0.080	0.012
$1(\Pi_c \ge 30\% \le 45\%)$	(0.036)	(0.124)	(0.065)	(0.848)	(0.057)	(0.023)
Spillover Group ×	-0.386	-1.675	-0.152	-0.359	-0.836	-0.066
$1(\Pi_c \ge 30\% \le 45\%)$	(0.456)	(1.338)	(0.468)	(6.680)	(0.701)	(0.079)
Assigned to Training \times	-0.050	-0.230	-0.057	-0.589	-0.096	-0.009
$1(II_c > 45\% \le 64\%)$	(0.036)	(0.115)	(0.052)	(0.697)	(0.051)	(0.009)
Spillover Group ×	-0.023	-0.232	-0.109	-1.213	-0.086	-0.014
$1(\Pi_c > 45\% \le 64\%)$	(0.089)	(0.202)	(0.139)	(1.832)	(0.087)	(0.016)
Panel C: Linear Analysis of	True Saturati	on				
Assigned to Training	0.126 ^{**} (0.060)	0.432*** (0.158)	0.056 (0.078)	0.764 (1.000)	0.195 ^{**} (0.089)	0.039* (0.021)
Assigned to Training \times	-0.003***	-0.012***	-0.003	-0.035	-0.005*	-0.001
Π_c (% True Saturation)	(0.002)	(0.005)	(0.002)	(0.029)	(0.003)	(0.001)
Spillover Group	0.114	0.396	-0.058	-1.128	0.294**	0.028
1 · · · · · · · · · · · · · · · · · · ·	(0.107)	(0.285)	(0.130)	(1.773)	(0.125)	(0.023)
Spillover Group ×	-0.005	-0.020*	-0.001	0.006	-0.011**	-0.001
Π_c (% True Saturation)	(0.004)	(0.012)	(0.006)	(0.078)	(0.005)	(0.001)

Table 3: Treatment effects on mobile money index components

Notes: Panel A shows intention to treat, and spillover effects pooled across all saturations. Panel B shows a non-parametric analysis of intention to treat and spillover effects for three bins of *true saturations* ($\Pi_c \ge 12\% < 30\%$), ($\Pi_c \ge 30\% \le 45\%$), ($\Pi_c > 45\% \le 64\%$). In this analysis, the true saturation (Π_c) is instrumented with the randomly assigned saturation (π_c); Panel C shows results of an affine model of how intention to treat and spillover effects change linearly with increasing true saturation. The true saturation (Π_c) is instrumented with the assigned saturation (π_c). All regressions included the dependent variable at baseline (ANCOVA), randomization strata fixed effects, enumerator fixed effects, a phone-survey (relative to field) indicator, and controls for sectors. Standard errors are clustered at the trading-center level, i.e., the level of random assignment. N=1,975. *** p<0.01, ** p<0.05, * p<0.1.

			Saving	g Index			Investm	ent Index
-	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Any	Savings Amt.	Any MM	MM Savings	Any	Bank	Any Investment	Invest. Amt. †
	Savings	Ť	Savings	Amt. †	Bank Savings	Savings Amt.†		
Control group mean (SD)	0.853	11.526 (5.247)	0.192	2.243 (4.705)	0.181	2.332 (5.222)	0.763	10.698 (6.068)
Panel A: Pooled Intention to	Treat and Spill	over Effects						
Assigned to Training	0.021	0.228	0.059***	0.597**	0.046**	0.581*	0.043**	0.567^{*}
	(0.020)	(0.286)	(0.020)	(0.249)	(0.022)	(0.297)	(0.021)	(0.300)
Spillover Group	-0.010	-0.377	-0.021	-0.249	0.000	-0.004	-0.025	-0.370
	(0.021)	(0.329)	(0.026)	(0.316)	(0.023)	(0.318)	(0.028)	(0.386)
Test of equality (p-value)	0.1161	0.0618	0.0016	0.0065	0.1118	0.1295	0.0075	0.0071
Panel B: Non-Parametric An	alysis of True S	Saturation						
Assigned to Training \times	0.022	0.177	0.093***	0.953***	0.066^{*}	0.926**	0.074**	0.980^{*}
$1(\Pi_c \ge 12\% < 30\%)$	(0.031)	(0.482)	(0.030)	(0.369)	(0.035)	(0.466)	(0.038)	(0.532)
Spillover Group ×	0.001	0.018	0.036	0.410	0.004	0.200	0.080	1.089
$1(\Pi_c \ge 12\% < 30\%)$	(0.039)	(0.610)	(0.044)	(0.535)	(0.033)	(0.452)	(0.050)	(0.740)
Assigned to Training \times	0.059	1.433*	0.055	0.625	0.006	-0.084	0.114*	1.492
$1(\Pi_c \ge 30\% \le 45\%)$	(0.054)	(0.740)	(0.043)	(0.519)	(0.054)	(0.701)	(0.069)	(0.958)
Spillover Group ×	-0.178	-4.849	-0.560	-6.413	-0.045	-1.903	-0.933	-13.020
$1(\Pi_c \ge 30\% \le 45\%)$	(0.339)	(5.908)	(0.427)	(4.927)	(0.340)	(4.663)	(0.657)	(9.275)
Assigned to Training ×	-0.004	-0.258	-0.010	-0.128	0.007	0.021	-0.022	-0.224
$1(\Pi_c > 45\% \le 64\%)$	(0.033)	(0.517)	(0.050)	(0.591)	(0.031)	(0.435)	(0.055)	(0.800)
Spillover Group ×	0.046	-0.561	0.031	0.095	0.130	1.530	-0.228**	-3.247**
$1(\Pi_c > 45\% \le 64\%)$	(0.065)	(1.536)	(0.103)	(1.166)	(0.142)	(1.836)	(0.116)	(1.598)
Panel C: Linear Analysis of	True Saturation							
Assigned to Training	0.044	0.317	0.185***	1.945***	0.145***	1.919***	0.082	1.184
	(0.056)	(0.739)	(0.056)	(0.718)	(0.053)	(0.721)	(0.056)	(0.822)
Assigned to Training \times	-0.001	-0.002	-0.004**	-0.043**	-0.003**	-0.044**	-0.001	-0.016
Π_c (% True Saturation)	(0.002)	(0.021)	(0.002)	(0.021)	(0.001)	(0.020)	(0.002)	(0.026)
Spillover Group	0.027	0.891	0.140	1.614	0.002	0.433	0.281***	3.885***
	(0.096)	(1.505)	(0.090)	(1.090)	(0.094)	(1.298)	(0.090)	(1.367)
Spillover Group ×	-0.001	-0.051	-0.006*	-0.070^{*}	0.000	-0.013	-0.012***	-0.170***
Π_c (% True Saturation)	(0.004)	(0.061)	(0.003)	(0.042)	(0.004)	(0.057)	(0.004)	(0.055)

Table 4: Treatment effects on savings- and investment index components

Notes: Panel A shows intention to treat, and spillover effects pooled across all saturations. Panel B shows a non-parametric analysis of intention to treat and spillover effects for three bins of *true saturations* ($\Pi_c \ge 12\% < 30\%$), ($\Pi_c \ge 30\% \le 45\%$), ($\Pi_c > 45\% \le 64\%$). In this analysis, the true saturation (Π_c) is instrumented with the randomly assigned saturation (π_c); Panel C shows results of an affine model of how intention to treat and spillover effects change linearly with increasing true saturation. The true saturation (Π_c) is instrumented with the assigned saturation (π_c). All regressions included the dependent variable at baseline (ANCOVA), randomization strata fixed effects, enumerator fixed effects, a phone-survey (relative to field) indicator, and controls for sectors. Standard errors are clustered at the trading-center level, i.e., the level of random assignment. N=1,975. *** p<0.01, ** p<0.05, * p<0.1.

	(1) Take- Up	(2) Left early (conditional on take-up)	(3) Class-Size	(4) Quality rating (training)	(5) Fully adequate venue (binary rating)	(6) Outdoor venue
Reference group mean (SD)	0.000	0.013	15.977 (6.996)	9.135 (0.202)	0.752	0.563
$1(\pi_c = 50\%)$. ,			
1 ($\pi_c = 75\%$)		-0.006	-2.240	0.053	-0.303^{**}	0.094
$1 (\pi_c = 100\%)$		(0.013) 0.027* (0.016)	2.121 (2.166)	(0.030) -0.042 (0.084)	(0.113) -0.298^{**} (0.123)	(0.187) 0.310^{*} (0.178)
Assigned to Training (T) × $1(\pi_c = 50\%)$ Spillover Group (S) × $1(\pi_c = 50\%)$	0.708*** (0.053) 0.302*** (0.057)					
Assigned to Training (T) × $1(\pi_c = 75\%)$ Spillover Group (S) × $1(\pi_c = 75\%)$	0.705 ^{***} (0.039) 0.393 ^{***} (0.061)					
Assigned to Training (T) × $1(\pi_c = 100\%)$	0.729*** (0.033)					
Tests of equality (p-values): $T \times (\pi_c = 50\%) = T \times (\pi_c = 75\%)$	0.954					
$T \times (\pi_c = 50\%) = T \times (\pi_c = 100\%)$	0.742					
$T \times (\pi_c = 75\%) = T \times (\pi_c = 100\%)$	0.648					
$S \times (\pi_c = 50\%) = S \times (\pi_c = 75\%)$	0.266					
$1(\pi_c = 50\%) = 1 \ (\pi_c = 75\%)$ $1(\pi_c = 50\%) = 1 \ (\pi_c = 100\%)$ $1(\pi_c = 75\%) = 1 \ (\pi_c = 100\%)$		0.675 0.094 0.069	0.154 0.332 0.046	0.293 0.622 0.315	0.011 0.018 0.972	0.618 0.088 0.233
N (individuals) N (clusters)	1,975 108	651 54	1,113 54	1,113 54	1,113 54	1,113 54

Table 5: Probing mechanisms behind saturation effects

 $\frac{100}{Notes}$ All regressions included randomization strata fixed effects, enumerator fixed effects, a phone-survey (relative to field) indicator, and controls for sectors. Standard errors are clustered at the trading-center level, i.e., the level of random assignment. N=1,975. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Heterogeneity by business type

		Retail		Service and Manufacturing				
		(n=1,380))		(n= 595)			
	(1)	(2)	(3)	(4)	(5)	(6)		
	Mo.	Savings	Investment	Mo. Money	Savings	Investment		
	Money							
Control group mean (SD)	0.004	0.054	-0.027	-0.009	-0.125	0.062		
	(0.991)	(1.031)	(1.026)	(1.021)	(0.913)	(0.937)		
Panel A: Pooled Intention to Treat and Spillover Effects by Business Type								
Assigned to Training	0.016	0.093	0.170^{***}	0.166^{*}	0.263***	-0.008		
	(0.080)	(0.060)	(0.062)	(0.097)	(0.091)	(0.082)		
Spillover Group	-0.143	-0.113*	-0.041	-0.101	0.082	0.012		
	(0.097)	(0.066)	(0.094)	(0.108)	(0.116)	(0.090)		
Test of equality (p-value)	0.132	0.008^{***}	0.034**	0.010^{**}	0.142	0.836		
Panel B: Non-Parametric Analysis of	of Assigned Se	aturation by I	Business Type					
	0.001	0.070	0.00/**	0.000*	0 400***	0.001		
Assigned to Training X	0.081	0.072	0.296	0.382	0.488	0.081		
$1(\Pi_c \ge 12\% < 30\%)$	(0.127)	(0.104)	(0.126)	(0.210)	(0.137)	(0.101)		
Spillover Group ×	-0.081	-0.053	0.286	0.347	0.335	0.043		
$1(\Pi_c \ge 12\% < 30\%)$	(0.131)	(0.121)	(0.193)	(0.266)	(0.214)	(0.102)		
Assigned to Training \times	0.100	0.302**	0.233	-0.114	-0.005	-0.069		
$1(\Pi_{c} \geq 30\% \leq 45\%)$	(0.151)	(0.131)	(0.163)	(0.259)	(0.189)	(0.185)		
Spillover Group ×	-0.313	-1.043	-2.824	-5.171	-2.134	0.756		
$1(\Pi_c \ge 30\% \le 45\%)$	(1.200)	(1.006)	(2.009)	(3.929)	(2.265)	(1.028)		
			, , , , , , , , , , , , , , , , , , ,			, , , , , , , , , , , , , , , , , , ,		
Assigned to Training \times	-0.244*	0.085	0.123	-0.473	-0.387	-0.257		
$1(\Pi_c > 45\% \le 64\%)$	(0.130)	(0.101)	(0.133)	(0.392)	(0.262)	(0.159)		
Spillover Group ×	-0.449	0.255	-0.214	0.357	0.016	-1.146***		
$1(\Pi_c > 45\% \le 64\%)$	(0.366)	(0.284)	(0.277)	(0.321)	(0.522)	(0.316)		
Panel C: Linear Spillover Analysis l	by Business T	vpe						
Assigned to Training	0.282	0.039	0.170	0.745^{**}	0.892***	0.243		
6 6	(0.212)	(0.196)	(0.145)	(0.300)	(0.236)	(0.183)		
Assigned to Training \times	-0.009	0.002	-0.000	-0.018**	-0.021***	-0.009		
Π_c (% True Saturation)	(0.006)	(0.005)	(0.004)	(0.008)	(0.006)	(0.006)		
Spillover Group	0.019	-0.004	0.722***	1.268**	0.773	-0.015		
-t-roter crowh	(0.332)	(0.259)	(0.340)	(0.552)	(0.518)	(0.271)		
Spillover Group X	-0.006	-0.003	-0.028**	-0.053**	-0.026	0.002		
Π (% True Saturation)	(0.014)	(0.000)	(0.014)	(0.023)	(0.020)	(0.002)		
Spillover Group × $1(\Pi_c \ge 12\% < 30\%)$ Assigned to Training × $1(\Pi_c \ge 30\% \le 45\%)$ Spillover Group × $1(\Pi_c \ge 30\% \le 45\%)$ Assigned to Training × $1(\Pi_c > 45\% \le 64\%)$ Spillover Group × $1(\Pi_c > 45\% \le 64\%)$ Panel C: Linear Spillover Analysis to Assigned to Training × Π_c (% True Saturation) Spillover Group × Π_c (% True Saturation)	$\begin{array}{r} -0.081 \\ (0.131) \\ 0.100 \\ (0.151) \\ -0.313 \\ (1.200) \\ -0.244^* \\ (0.130) \\ -0.449 \\ (0.366) \\ \hline by Business T \\ 0.282 \\ (0.212) \\ -0.009 \\ (0.006) \\ 0.019 \\ (0.332) \\ -0.006 \\ (0.014) \end{array}$	$\begin{array}{c} -0.053 \\ (0.121) \\ 0.302^{**} \\ (0.131) \\ -1.043 \\ (1.006) \\ 0.085 \\ (0.101) \\ 0.255 \\ (0.284) \\ \end{array}$	$\begin{array}{c} 0.286\\ (0.193)\\ 0.233\\ (0.163)\\ -2.824\\ (2.009)\\ 0.123\\ (0.133)\\ -0.214\\ (0.277)\\ \end{array}$ $\begin{array}{c} 0.170\\ (0.145)\\ -0.000\\ (0.004)\\ 0.722^{**}\\ (0.340)\\ -0.028^{**}\\ (0.014)\\ \end{array}$	$\begin{array}{c} 0.347\\ (0.266)\\ -0.114\\ (0.259)\\ -5.171\\ (3.929)\\ -0.473\\ (0.392)\\ 0.357\\ (0.321)\\ \end{array}$ $\begin{array}{c} 0.745^{**}\\ (0.300)\\ -0.018^{**}\\ (0.008)\\ 1.268^{**}\\ (0.552)\\ -0.053^{**}\\ (0.023)\\ \end{array}$	$\begin{array}{c} 0.335\\ (0.214)\\ -0.005\\ (0.189)\\ -2.134\\ (2.265)\\ -0.387\\ (0.262)\\ 0.016\\ (0.522)\\ \end{array}$	$\begin{array}{c} 0.043\\ (0.102)\\ -0.069\\ (0.185)\\ 0.756\\ (1.028)\\ -0.257\\ (0.159)\\ -1.146^{***}\\ (0.316)\\ \end{array}$ $\begin{array}{c} 0.243\\ (0.183)\\ -0.009\\ (0.006)\\ -0.015\\ (0.271)\\ 0.002\\ (0.011)\\ \end{array}$		

Notes: The dependent variables in column 1-6 are all standardized to have a mean of 0 and a standard deviation of 1 in the pooled pure control group sample. Panel A shows intention to treat, and spillover effects pooled across all saturations. Panel B shows a non-parametric analysis of intention to treat and spillover effects for three bins of *true saturations* ($\Pi_c \ge 12\% < 30\%$), ($\Pi_c \ge 30\% \le 45\%$), ($\Pi_c > 45\% \le 64\%$). In this analysis, the true saturation (Π_c) is instrumented with the randomly assigned saturation (π_c); Panel C shows results of an affine model of how intention to treat and spillover effects change linearly with increasing true saturation. The true saturation (Π_c) is instrumented with the assigned saturation (π_c). All regressions included the dependent variable at baseline (ANCOVA), randomization strata fixed effects, enumerator fixed effects, a phone-survey (relative to field) indicator, and controls for sectors. Standard errors are clustered at the trading-center level, i.e., the level of random assignment. *** p<0.01, ** p<0.05, * p<0.1.

ONLINE-APPENDIX

to accompany

Scaling financial education among micro-entrepreneurs: A randomized saturation experiment

(Jana Hamdan, Tim Kaiser, Lukas Menkhoff and Yuanwei Xu)

- Intervention and field experiment supplementary information
- Attrition and balance
- Field experiment auxiliary results

INTERVENTION AND FIELD EXPERIMENT SUPPLEMENTARY INFORMATION

Table A1: Timeline

	Q1/′19	Q2/'19	Q3/'19	Q4/′19	Q1/'20	Q2/'20	Q3/'20	Q4/'20	Q1/'21	Q2/'21
Baseline (field)	Х	Х								
Treatment			Х							
Follow-Up (phone)								Х		
Follow-Up (field)										Х

	Topic	Learning Objectives	Duration
Session	Introduction –	~ ~ ~	30-40
#1	Financial Service	- Participants know at least three Financial	minutes
	Providers	Service Providers	
		- Participants know at least 3 rights of	
		Financial Service users	
		- Participants are able to choose a financial	
~ .		service provider that suits their demands	<i>.</i> .
Session	Personal Financial	- Participants have the ability to differentiate	60 minutes
#2	Management	between financial inflows and outflows	and more
		- Participants have the ability to differentiate	
		wants and making a simple budget for	
		somebody else	
		- Participants are able to budget for themselves	
		and their (family) businesses	
		- Participant know the importance of personal	
		financial management and motivated to start	
		financial planning and record keeping in their	
		business	
Session	Saving	- Participants know the meaning of saving	60 minutes
#3		- Participants reflect on the reasons for saving	and more
		- Participants are able to evaluate saving	
		methods by their accessibility, return and	
		security and making educated choices of	
		financial products for saving	
Session	Debt management	- Participants know what a loan is	up to 60
#4		- Participants can distinguish productive and	minutes
		non-productive loans	
		- Participants know at least three costs of	
Section	Investment	Dorrowing Derticinents Image and reflect shout common	um to 60
5ession #5	Investment	- Participants know and reflect about common muths about investing	up to oo
#3		- Participants can analyze different investment	minutes
		ontions	
		- Participants are able to draft an investment	
		plan and to anticipate possible risks.	
Session	Money transfer	- Participants know what money transfer is	up to 60
#6	5	- Participants know about mobile money	minutes
		- Participants are able to compare the costs	
		between different means of money transfer	
		options	
		- Participants are able to compare the risk	
		associated with some informal ways of money	
		transfer.	

Table A2: Overview of financial education treatment

Figure A1: Example Learning Material



ATTRITION AND BALANCE

Reason	Control	Assigned to training	Spillover	Total
Relocated	92	61	24	177
Unknown	21	18	7	46
Died	7	2	1	10
Declined	6	4	0	10
Sick	2	0	0	2
Imprisoned	1	1	1	3
Total	129	86	33	248

Table A3: Attrition by group

	Table A	4: Corre	lates of	relocation
--	---------	----------	----------	------------

	Relocation (=1)	Relocation (=1)	Relocation (=1)
	(1)	(2)	(3)
Female	0.01	0.01	0.01
	(0.012)	(0.011)	(0.012)
Age	-0.00***	-0.00***	-0.00***
	(0.001)	(0.001)	(0.001)
Work Experience (Years)	-0.00	-0.00	-0.00
	(0.001)	(0.001)	(0.001)
Sales (UGX)	-0.00***	-0.00***	-0.00***
	(0.000)	(0.000)	(0.000)
Constant	0.18***	0.18***	0.17***
	(0.024)	(0.031)	(0.035)
Observations	2,165	2,165	2,165
R^2	0.020	0.033	0.033
Strata FE	no	yes	yes
Inverse Sampling Weights	no	no	yes

Note: The table shows linear regression results with the binary dependent variable being relocation outside of the study region as the reason for attrition. From column (2) onward, the regression equations include strata fixed effects and, in column (3), is weighted by sampling weights and experimental design weights. Sales refer to average daily sales in UGX in the past month and are top coded at 99%. Standard errors are clustered at the trading center level and displayed in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1

	Attrit	ion (=1)	Relocation Reason (=1)		Refusal Reason (=1)		Other Reason (=1)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated TC's	-0.038** (0.016)		-0.030* (0.017)		-0.003 (0.003)		-0.005^{*} (0.003)	
Assigned to Training	. ,	-0.038 ^{**} (0.017)		-0.030* (0.017)	()	-0.001 (0.003)	()	-0.006 ^{**} (0.003)
Spillover Group		-0.038 [*] (0.019)		-0.029 (0.019)		-0.006 ^{**} (0.003)		-0.003 (0.004)
Observations R^2	2,177 0.015	2,177 0.015	2,177 0.017	2,177 0.017	2,177 0.008	2,177 0.009	2,177 0.009	2,177 0.009
Strata FE	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark

Table A5: Attrition and its reasons by treatment status

Note: The table shows linear regression results with the binary dependent variable attrition in columns (1)-(2) and reasons for attrition in columns (3)-(8). The regression models include strata fixed effects. Weighted by sampling weights and experimental design weights. "Other" reasons include illnesses, imprisonment and death. Robust standard errors, clustered at the trading center (TC) level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

FIELD EXPERIMENT AUXILIARY RESULTS

Table A6: Pooled intention to treat and spillover effects, non-parametric saturation analysis, and linear saturation analysis for budgeting index components

	(1)	(2)
	Keeps business log/record	Separate records: business
	1 8	and personal
Pure control group mean	0.485	0.562
Panel A: Pooled ITT and Spillover		
Ĩ		
Assigned to Training	0.008	0.021
	(0.030)	(0.027)
Spillover Group	-0.004	-0.033
	(0.038)	(0.026)
Test of equality (p-value)	0.7207	0.0240
Panel B: Non-Parametric Analysis of True Saturation		
Assigned to Training \times	0.018	0.039
$1(\prod_{c} \ge 12\% < 30\%)$	(0.039)	(0.042)
Spillover Group ×	0.038	0.046
$1(\Pi_c \ge 12\% < 30\%)$	(0.055)	(0.043)
Assigned to Training \times	0.038	0.066
$1(\Pi_c \ge 30\% \le 45\%)$	(0.077)	(0.082)
Spillover Group ×	-0.440	-0.806
$1(\Pi_c \ge 30\% \le 45\%)$	(0.564)	(0.544)
Assigned to Training \times	-0.015	-0.004
$1(\Pi_c > 45\% \le 64\%)$	(0.066)	(0.060)
Spillover Group ×	0.002	-0.084
$1(\Pi_c > 45\% \le 64\%)$	(0.108)	(0.116)
Panel C: Linear Analysis of True Saturation		
Assigned to Training	0.005	0.047
	(0.065)	(0.077)
Assigned to Training \times	0.000	-0.001
Π_c (% True Saturation)	(0.002)	(0.002)
Spillover Group	0.121	0.201^{**}
	(0.153)	(0.099)
Spillover Group ×	-0.005	-0.009**
Π_c (% True Saturation)	(0.007)	(0.004)

Notes: Panel A shows intention to treat, and spillover effects pooled across all saturations. Panel B shows a non-parametric analysis of intention to treat and spillover effects for three bins of *true saturations* ($\Pi_c \ge 12\% < 30\%$), ($\Pi_c \ge 30\% \le 45\%$), ($\Pi_c > 45\% \le 64\%$). In this analysis, the true saturation (Π_c) is instrumented with the randomly assigned saturation (π_c); Panel C shows results of an affine model of how intention to treat and spillover effects change linearly with increasing true saturation. The true saturation (Π_c) is instrumented with the assigned saturation (π_c). All regressions included the dependent variable at baseline (ANCOVA), randomization strata fixed effects, enumerator fixed effects, a phone-survey (relative to field) indicator, and controls for sectors. Standard errors are clustered at the trading-center level, i.e., the level of random assignment. N=1,975. *** p<0.01, ** p<0.05, * p<0.1.

 Table A7: Pooled intention to treat and spillover effects, non-parametric saturation analysis, and linear saturation analysis for borrowing index components

	(1)	(2)	(3)	(4)	
	Any loan	MM loan	Bank Loan	Loan Amt.†	
Control group mean (SD)	0.792	0.091	0.090	10.669	
				(5.529)	
Panel A: Pooled ITT and Spillover				· · ·	
Assigned to Training	-0.018	-0.018	0.019	-0.321	
e e	(0.021)	(0.015)	(0.015)	(0.265)	
Spillover Group	-0.030	-0.023	0.019	-0.587*	
	(0.023)	(0.020)	(0.020)	(0.303)	
Test of equality (p-value)	0.6149	0.8360	0.9784	0.3321	
Panel B: Non-Parametric Analysis of Tru	e Saturation				
Assigned to Training \times	-0.004	-0.003	0.005	0.012	
$1(\Pi_c \ge 12\% < 30\%)$	(0.032)	(0.025)	(0.030)	(0.392)	
Spillover Group ×	0.020	-0.008	-0.054	0.126	
$1(\Pi_c \ge 12\% < 30\%)$	(0.039)	(0.031)	(0.038)	(0.509)	
Assigned to Training \times	0.003	-0.024	-0.025	-0.412	
$1(\Pi_c \ge 30\% \le 45\%)$	(0.053)	(0.045)	(0.050)	(0.731)	
Spillover Group ×	-0.640	-0.074	0.724	-8.309	
$1(\Pi_c \ge 30\% \le 45\%)$	(0.408)	(0.281)	(0.522)	(5.235)	
Assigned to Training \times	-0.037	-0.057**	0.034	-0.749	
$1(\Pi_c > 45\% \le 64\%)$	(0.035)	(0.022)	(0.031)	(0.480)	
Spillover Group ×	0.145^{**}	-0.096***	0.098	1.062	
$1(\Pi_c > 45\% \le 64\%)$	(0.067)	(0.025)	(0.071)	(1.070)	
Panel C: Linear Analysis of True Saturation					
Assigned to Training	-0.033	0.005	0.034	-0.072	
	(0.056)	(0.043)	(0.047)	(0.617)	
Assigned to Training \times	0.001	-0.001	-0.001	-0.006	
Π_c (% True Saturation)	(0.002)	(0.001)	(0.001)	(0.018)	
Spillover Group	0.116	0.012	-0.198**	1.432	
-	(0.086)	(0.078)	(0.092)	(1.103)	
Spillover Group ×	-0.006^{*}	-0.001	0.009^{**}	-0.081*	
Π_c (% True Saturation)	(0.004)	(0.003)	(0.004)	(0.045)	

Notes: Panel A shows intention to treat, and spillover effects pooled across all saturations. Panel B shows a non-parametric analysis of intention to treat and spillover effects for three bins of *true saturations* ($\Pi_c \ge 12\% < 30\%$), ($\Pi_c \ge 30\% \le 45\%$), ($\Pi_c > 45\% \le 64\%$). In this analysis, the true saturation (Π_c) is instrumented with the randomly assigned saturation (π_c); Panel C shows results of an affine model of how intention to treat and spillover effects change linearly with increasing true saturation. The true saturation (Π_c) is instrumented with the assigned saturation (π_c). All regressions included the dependent variable at baseline (ANCOVA), randomization strata fixed effects, enumerator fixed effects, a phone-survey (relative to field) indicator, and controls for sectors. Standard errors are clustered at the trading-center level, i.e., the level of random assignment. N=1,975. *** p<0.01, ** p<0.05, * p<0.1.

	Participation Among All in Treated TCs	Participation Among Targeted in Treated TCs	Participation Among Untargeted in Treated TCs	
	(1)	(2)	(3)	
Female	-0.000	-0.001	0.067	
	(0.037)	(0.039)	(0.051)	
Age	0.001	0.001	-0.000	
	(0.001)	(0.002)	(0.002)	
Married	-0.002	-0.030	0.012	
	(0.032)	(0.030)	(0.045)	
Years of Education	-0.008**	-0.007	-0.007	
	(0.004)	(0.005)	(0.004)	
Previous FL Training	0.079	0.049	0.107	
-	(0.066)	(0.099)	(0.094)	
Financial Literacy (0-7)	0.018	0.020	-0.004	
	(0.013)	(0.016)	(0.012)	
Risk Tolerance (0-10)	0.001	0.003	-0.000	
	(0.005)	(0.007)	(0.008)	
Household Size	-0.001	-0.003	0.001	
	(0.009)	(0.009)	(0.010)	
# Assets	0.001	0.002	-0.003*	
	(0.001)	(0.002)	(0.001)	
Household Consumption (UGX)	-0.000	-0.000	0.000	
· · · · · ·	(0.000)	(0.000)	(0.000)	
Constant	0.891***	0.818***	1.082***	
	(0.085)	(0.098)	(0.125)	
Observations	1,203	858	345	
R^2	0.082	0.062	0.302	

Table A8: Demographic correlates of take-up

Note: The table shows linear regression results with the binary dependent variable participation in the training. Column (1) includes all respondents in treated clusters. Column (2) includes targeted respondents only, column (3) only untargeted respondents within the treated clusters. The regression models include strata fixed effects and is weighted by sampling weights and experimental design weights. Standard errors are clustered at the trading center (TC) level and displayed in parentheses. *** p<0.01, ** p<0.05, * p<0.1