

# Impact of savings groups on the lives of the poor

Dean Karlan<sup>a,b,1</sup>, Beniamino Savonitto<sup>b</sup>, Bram Thuysbaert<sup>b,c</sup>, and Christopher Udry<sup>a</sup>

<sup>a</sup>Department of Economics, Yale University, New Haven, CT 06511; <sup>b</sup>Innovations for Poverty Action, New Haven, CT 06510; and <sup>c</sup>Department of Economics, Ghent University, B-9000 Ghent, Belgium

Edited by Abhijit Banerjee, Massachusetts Institute of Technology and Jameel Poverty Action Laboratory, and accepted by Editorial Board Member Paul R. Milgrom February 10, 2017 (received for review July 22, 2016)

**Savings-led microfinance programs operate in poor rural communities in developing countries to establish groups that save and then lend out the accumulated savings to each other. Nonprofit organizations train villagers to create and lead these groups. In a clustered randomized evaluation spanning three African countries (Ghana, Malawi, and Uganda), we find that the promotion of these community-based microfinance groups leads to an improvement in household business outcomes and women's empowerment. However, we do not find evidence of impacts on average consumption or other livelihoods.**

savings | microfinance | rural development | resilience | empowerment

The poor make complex financial decisions and use the limited range of financial instruments available to them to address their varying needs. The available formal and informal tools, however, are often risky and expensive or lack necessary flexibilities (1). The quest to improve access to appropriate financial services for poor households has traditionally focused on providing credit via formal alternatives to informal moneylenders. The most relevant example is the growth of the Grameen microcredit model developed by Mohammed Yunus.

However, the limited participation and geographic reach of microcredit institutions, especially among the rural poor, have shifted efforts into expanding access to savings. A growing impact literature on microcredit and microsavings shows strong welfare impacts of the latter type of programs, suggesting that they might have more transformative impacts than credit programs as currently typically implemented (2, 3).

In parallel to the development of microcredit and microsavings products for the poor, many nongovernmental organizations (NGOs) have begun promoting informal savings-led microfinance groups that emulate and improve on the model of informal associations indigenous to many societies [often called Rotating Savings and Credit Associations (ROSCAs)]. Although often implemented as a stand-alone program, these savings groups are also often a component within multifaceted programs (4). The appeal of the savings-led microfinance approach is shown by the growth of these groups, which now reach over 10 million people in more than 70 countries\* after only a few years of significant expansion efforts. Many actors, including international donors, such as the Bill and Melinda Gates Foundation,<sup>†</sup> and nongovernment organizations,<sup>‡</sup> have pushed to create and expand such groups, viewing them as a grassroots and low-cost mechanism to provide (albeit informal) financial services to the poor. Similarly, the Andhra Pradesh Government in India promoted a self-help group model (Velugu) as an alternative to formal microcredit during a crisis in 2010 (5). The model, which is being explored by many NGOs, is a type of savings group with the added feature of explicit intentions to link groups to formal sector banks for credit after they mature enough [e.g., Plan and CARE (Cooperative for Assistance and Relief Everywhere) have partnered with Barclays for this purpose] (6).

Savings-led microfinance groups vary depending on the implementing organization and the context; however, the basic features follow the Village Savings and Loan Association (VSLA) model developed in the early 1990s in Mali by CARE, one of the leading promoters of these groups and the main partner in this study. The VSLA model was designed as an improvement of the local tontine,

a type of savings group where members gather at regular meetings to contribute a fixed amount of money and the total pot is assigned in full to each member in turn. VSLAs adopt the same meeting and contribution structure but introduce flexibility. At each group meeting, members can decide to contribute more than the agreed minimum and can take a loan from the group without having to wait for their preassigned turn. These loans are charged an interest rate, so that the money deposited by group members can earn interest. Savings and loan repayments are kept in a group lockbox that can only be opened at group meetings and “shared out” among members at the end of a predefined cycle. A VSLA may have an additional social or solidarity fund, which is an insurance fund managed by the group that can be accessed by members in the form of an interest-free loan or cash grant in case of an emergency. Disbursements for these purposes are assessed and determined by the group. VSLAs do not receive any capital through grants or external loans; the pots simply grow over time as individuals collectively accumulate more savings. We present the results from three randomized, controlled trials of the VSLA program implemented by CARE and its partners in Ghana, Malawi, and Uganda over a period of 22–30 mo.

First, we report evidence on the program's impact across eight outcomes. The growth in financial intermediation resulting from the program led to improved microenterprise outcomes and women's empowerment. The positive impact on female empowerment stands in contrast to the results of six of seven randomized trials on microcredit (3, 7–13). However, when examining average treatment effects, much like the microcredit studies, we do not find impacts on typical welfare indicators, such as household income, consumption, food security, asset ownership, or community participation.

## Significance

We conducted a large randomized evaluation of a savings-led microfinance program across three countries. This evaluation provides important evidence on the impact of a popular development intervention on the lives of low-income households in rural communities by looking at its effects on usage of financial services, microenterprise activity, income, female empowerment, consumption, and the ability to cope with shocks.

Author contributions: D.K., B.S., B.T., and C.U. designed research; D.K., B.S., B.T., and C.U. analyzed data; D.K., B.S., and C.U. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. A.B. is a Guest Editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in Dataverse, <https://dataverse.harvard.edu/dataverse/socialsciencercs> (doi no. 10.7910/DVN/EJTYHC).

<sup>1</sup>To whom correspondence should be addressed. Email: dean.karlan@yale.edu.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1611520114/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1611520114/-DCSupplemental).

\*Data on the growth of these groups can be accessed at [seepnetwork.org/savings-groups-global-outreach-pages-20015.php](http://seepnetwork.org/savings-groups-global-outreach-pages-20015.php).

<sup>†</sup>In 2009, the Bill and Melinda Gates Foundation awarded a grant to CARE “to develop sustainability models for savings-led financial services for the poor” ([www.gatesfoundation.org/How-We-Work/Quick-Links/Grants-Database/Grants/2008/09/OPP49313](http://www.gatesfoundation.org/How-We-Work/Quick-Links/Grants-Database/Grants/2008/09/OPP49313)).

<sup>‡</sup>Examples are available at [www.vsla.net](http://www.vsla.net) and [www.thesavix.org](http://www.thesavix.org).

Second, we examine the program's effect on resilience to shocks. In villages that suffered from drought (an aggregate shock to the community), we find a positive impact of the program on income (although significance does not hold up to multiple hypotheses correction). There is also some evidence (again, weak statistically) that the positive impact of VSLAs on female empowerment is reduced in villages subject to drought. We find no evidence of differential program effects for households with a self-reported bad harvest (which is a tentative indicator of idiosyncratic shocks after controlling for aggregate rainfall but subject to potential mismeasurement and selection concerns discussed below). Thus, in net, we find at best suggestive evidence that VSLAs may influence risk management capabilities, more so for aggregate shocks than for idiosyncratic shocks.

The results here can be compared with those found in two other randomized evaluations of savings-led microfinance groups in Mali (14) and Malawi (15). In Mali, researchers found positive impacts on food security and investment in cattle but found no impact on female empowerment. In Malawi, researchers found positive impacts on consumption on average (not merely during shocks), business income, and also, investment in housing structure (specifically, expansion in the number of rooms).

This study received approval from the Yale University Human Subjects Committee [Institutional Review Board (IRB) nos. 0805003819, 0904005015, and 0903004937] and the Innovations for Poverty Action Human Subjects Committee (IRB Protocols 010.08April-003, 109.09April-001, and 110.09March-001).

## Program

Implemented by CARE and 13 local partner NGOs in the study areas, the VSLA program has three main components: (i) a group-based commitment savings mechanism, (ii) a process for members to request loans from the group at any point, and (iii) a social or emergency fund financed by members with a regular contribution. In each site, after an initial community meeting to introduce the program, trained officers or agents form and guide VSLAs for an initial cycle (usually 8–12 mo) and provide oversight and support for a second cycle. Groups are comprised of 19–30 members, mostly women, who choose to come together as a result of an agent's promotional activity in the community or after having observed other groups.

Table S1 compares the implementation strategies and characteristics of VSLAs across the three sites. In each country, the role and mandate of the trainers were slightly different. In Ghana, officers from the NGO (called field officers) were in charge of creating VSLAs in each of their villages. In Uganda and Malawi, however, implementation was designed to facilitate scaling the program in a sustainable way: Local people were trained as village agents, who then trained other groups in the village and surrounding areas for a fee. The Malawi program was designed to rely heavily on these agents at the onset. Program data from the last 3 mo of the study (2nd quarter, 2011) show that 90.3% of the groups created by study partners were created by village agents.<sup>§</sup> In Uganda, the field officer was given the mandate to work in a group of villages and begin training agents only later on in the program's timeline. The difference in strategy is reflected in the fact that, in the same period, only 20.4% of the groups were formed by village agents. Notably, research in Kenya finds important differences in the efficacy of savings-led microfinance groups managed by NGO-paid agents and village agents paid by group members themselves: the village-paid agent model is less expensive; generates more borrowing, savings, and enterprise investment; and likely does this by attracting more business-oriented members (16).

<sup>§</sup>Information on breakdown by delivery agent, as well as a plethora of other data, are available at [www.thesavix.org/analysis](http://www.thesavix.org/analysis). However, this online tool does not allow one to see historical results on breakdown by delivery agent. The reported number was given by our study partners and the online tool reflects a similar breakdown for all projects in Malawi across time periods.

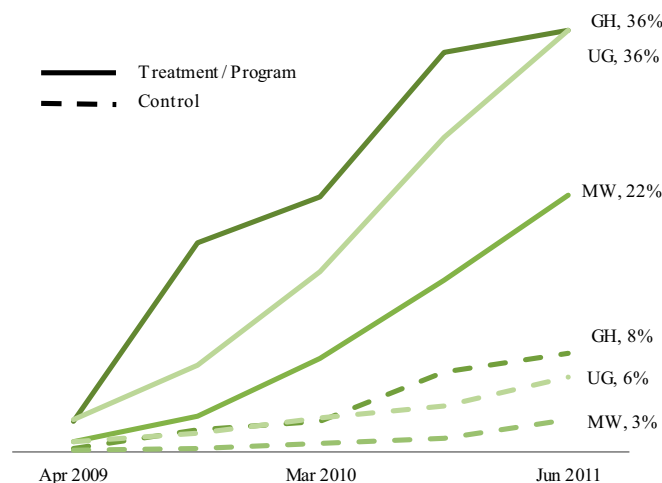


Fig. 1. Program uptake. Percentage of female primary respondents joining a VSLA in the study sample in treatment (solid lines) and control (dotted lines) clusters. GH, Ghana; MW, Malawi; UG, Uganda.

Take-up rates at the end of the study averaged 31.6% across the three sites, with a steady increase over time. We find evidence of the program successfully replicating VSLAs beyond the primary target village, with take up increasing over time and also reaching the control villages, which had a take-up rate of 6.2% by end line. Take-up trends and overall numbers varied slightly across sites as shown in Fig. 1.

At the time of the end-line survey, respondents who had joined a group had been members for a median value of 14 mo, with over 61% of them having completed a full savings cycle and receiving at least one share out. Members made small weekly savings contributions, with median values ranging between purchasing power parity (PPP) 2011 US\$0.66 and \$0.84 across the three countries.<sup>¶</sup> The main reported uses for savings share outs (median value of PPP 2011 US\$38.5) were agricultural investments (22%), food (16%), and education costs (16%).

At end line, 68% of members had received at least one loan from the group. The median loan was PPP 2011 US\$19.7, with a median interest charge of 10% flat. The main stated uses for loans were business (29%), food (13%), and education (13%).

## Research Design

We present the results from randomized, controlled trials of the VSLA program across a total of 561 clusters, 282 of which were randomly assigned to treatment and the remaining of which were randomly assigned to control. Treatment clusters received initial promotion and group formation activities by agents of the implementing organizations. Table S2 presents key features of the research design across the three research sites.

In Ghana, a cluster consisted of one community identified by the partnering NGO as a viable village for implementation of the program. In Malawi and Uganda, to measure organic replication activities of the program, a cluster consisted of two villages: one was identified as the primary village, and a second one was randomly chosen from other villages within a few kilometers of the primary village. The location of the secondary village was not disclosed to the partner NGOs to avoid targeting and, thereby, permit us to measure organic replication. Random assignment was

<sup>¶</sup>Throughout the study, we express currencies in terms of 2011 US dollars using the World Bank's purchasing power parity conversion factor for gross domestic product (GDP), expressed as local currency unit (LCU) per international dollar. For Ghana, the value of this factor is 0.7 Ghanaian cedi to the US dollar; for Malawi, it is 76.26 Malawian kwacha to the US dollar; and for Uganda, it is 833.54 Ugandan shillings to the US dollar.

stratified at the district level in all three countries and on other variables depending on the country, as specified in Table S2.

In each country, the village representatives were first approached to help identify adequate households for surveying. Surveyors then created a list of household heads in the village, from which they randomly selected participants. In the case of Uganda and Malawi, a majority was selected from the primary village, and a smaller number was selected from the secondary village. Overall, we collected panel survey data on 15,221 households from the three sites (survey instruments are available from the authors); 13,564 households were surveyed at baseline, 91.2% of which were resurveyed in the follow-up, and an additional 2,845 households were surveyed in Ghana at end line to mitigate concerns about lower than expected take-up rates in treatment villages and the proximity of control villages to treatment villages, which caused some of the former to adopt the VSLA program.

Four surveys were administered within each wave of data collection. First, a household survey collected information on indicators, such as agricultural production, income-generating activities, and economic shocks. Second, an adult survey collected information on the individual's experience of gender issues and community involvement and on their savings and loans activities. For the majority of our sample (13,502 households), the adult survey was administered to an adult woman in the household.<sup>#</sup> Third, a village survey gathered information about various characteristics of the community. Fourth, a market survey was used to record the market prices of a variety of staple foods grown by respondents.

## Analysis

As is the standard practice for multisite trials, we estimate a pooled model controlling for baseline values of our outcome variables, country, and district, which was used as a stratification variable in each site. SEs are clustered at the village level (comprised of a primary–secondary village pair for Uganda and Malawi). Each column in Table 1 is a result of an independent ordinary least squares regression modeled with the following specification:

$$Y_{ic}^k = \alpha + \beta_1 \text{Assignment}_c + \beta_2 Z_{ic}^k + T_{Country} + V_{Stratification} + \varepsilon_{ic},$$

where  $Y_{ic}^k$  is the index  $k$  of interest for either the household or the female primary respondent  $i$  in cluster  $c$ .  $\text{Assignment}_c$  is an indicator variable for whether the cluster was randomly assigned to receive the VSLA program or not,  $Z_{ic}^k$  is the household or female primary respondent baseline value of the outcome index  $k$  (if not available, the baseline value is replaced with  $-9$ , with a dummy variable indicating a missing value),  $T_{Country}$  is a vector of dummy variables for each of three countries in the study, and  $V_{Stratification}$  is a vector of dummy variables for districts.

The analysis used here is an “intent to treat” method, where we compare those who were randomly assigned to receive treatment with those who were not, irrespective of actual take up. This approach is important for interpretation and also in accordance with the program's promotion as a community-level intervention. There are two main reasons that we do not estimate the “treatment on the treated,” in which we would use random assignment as an instrumental variable for participation in a savings-led microfinance group to estimate the effect of participation on the outcomes of interest. First, participation in a VSLA is a continuous variable, with some just having started participation and others participating for a long time. Second, such a specification requires there to be no indirect effects on others in the community, which contradicts one of the main arguments for savings-led microfinance

programs: that they improve overall community wellbeing and social capital.

Because the program was multifaceted and we collected data on a large number of outcome variables, we are concerned with multiple hypothesis testing, which we address in two ways. First, we create indices for six of eight families of outcomes, which avoids interpreting one individual outcome measured with noise as indicative of a genuine change in an outcome. Second, we calculate  $q$  values across outcomes using the Benjamini–Hochberg step-up method (17) to control for false discovery rates. However, theory and evidence from elsewhere on impacts of improved financial inclusion (3) suggest that not all families of outcomes should be considered as equally likely to shift.

For the outcomes, we define six main indices and two aggregate measures: consumption and income. Each measure is formed by grouping together variables of related outcomes.

The financial inclusion index measures the extent of female primary respondents' involvement in the informal credit and savings sector and includes savings balances, loans taken in the last 12 mo, total amount of loans received, whether a person has savings, and whether the respondent is a member of a savings group.

The food security index uses metrics on meal patterns to determine food consumption during the 12 mo prior to surveying. It includes food intake reduction, days without eating for adults and children, and a variable indicating whether the household resorted to borrowing food.

Income is an aggregate measure of household income and revenues from all common sources: agriculture, livestock, business profits, and paid labor. All income measures refer to households' reported monthly values. Monthly household income is calculated as self-reported revenues minus expenses for all income-generating activities carried out by the household in the 12 mo before the survey. Annual microenterprise income is coded as zero for microenterprises operating less than 1 mo that year. Livestock income is the sales proceeds derived from livestock in a month minus the cost of livestock purchases. We do not deduct other input costs when calculating livestock income. Business profits are the profits earned through the households' enterprise(s) calculated as revenues minus costs. If a business operated for less than 1 mo, its profits were not included. Finally, income from paid labor is the income earned by a household member by being a paid employee, including agricultural day laborers.

The business outcomes index measures nonmonetary performance of microenterprises run by household members over the 12 mo preceding a survey. It includes the total number of businesses operated by household members, the sum of months that these businesses were in operation in the year preceding surveying, and whether any of the household's businesses employs labor from outside the household.

The assets index combines two standardized indices measuring household assets and productive assets held by the household. The calculation of values of assets was standardized across the three countries and time periods by expressing each asset value in relative terms: specifically as the number of bicycles needed to purchase one unit of the asset. More information on the construction of this index is provided in *SI Text*.

Consumption is an aggregate measure of monthly per capita expenditure on food and nonfood items. Food consumption expenditure used a 1-wk recall, and included only those items for which data were collected across the three sites (grains, tubers, nuts, and beans), meaning it presents only a partial view of the household's consumption basket. The nonfood consumption measure includes monthly per capita household expenditure on transport, clothes, electricity, and petrol.

The women's empowerment index captures women's self-reported influence on household decisions, particularly in relation to food expenses for the household, education and healthcare

<sup>#</sup>The sample of household female respondents at end line was 13,502, corresponding to approximately one woman per household in Ghana and Malawi and one woman per every two households in Uganda. In Uganda, we collected female primary respondent information from a randomly selected 50% of households in our sample.

expenses for the children, business expenses if the household operates a business, and the women's ability to visit friends.

Finally, the community participation index captures self-reported involvement in community affairs based on whether the female primary respondent has raised an issue before a village chief, government authority, or village council; whether she has attended a community meeting in the 12 mo preceding surveying; and whether she has participated in any social group within the community (a women's or workers' group, for instance).

Table 1 presents the main results of all eight outcome variables, and Tables S3–S10 show their breakdown by components. The indices are constructed as  $Y_{ifc}^k$ , interpreted as the outcome variable  $k$  for adult respondent (or household)  $i$  in family of outcomes  $f$  within a country  $c$ . All indices and variables have been defined, such that a higher value corresponds with improved outcomes. A detailed explanation on the method followed for index construction is provided in *SI Text*.

Information on drought and bad harvest interactions in Tables S3–S10 examines a critical heterogeneity: whether the program

had a differential impact during times of drought or among households reporting a bad harvest. We obtained rainfall data from the daily Africa Rainfall Climatology 2 dataset (18). Based on these data, we created an indicator variable for all villages that experienced a drought (defined as an annual rainfall less than 1 SD below a historic average) 12 mo before the end-line survey. In Uganda and Malawi, 34.6 and 15.4%, respectively, of villages experienced drought. Ghana is excluded from this analysis because the period of our study was exceptionally rainy; all but one village had a larger than average rainfall. In addition, we have end-line data from Uganda and Malawi on households' self-reported experience of a "bad harvest" over the previous year (80.1 and 33.2% of households, respectively).

## Results

At end line, we find substantial positive impacts on financial inclusion as a direct outcome of the program. Table S3 shows that overall participation in informal savings groups (including ROSCAs and other types of groups) is 17.4% points (SE = 0.015)

**Table 1. Impacts of the program on families of outcome variables**

Parameter	Financial inclusion index (adult female respondent)	Food security index (household)	Income and revenue (\$; household)	Business outcomes index (household)	Total asset index (household)	Monthly per capita consumption (\$; household)	Women's empowerment index (adult female respondent)	Community participation index (adult female respondent)
<b>All countries</b>								
Village offered VSLA	0.313***	0.024	3.717	0.057**	−0.005	−0.597	0.064*	0.035
SE	(0.028)	(0.024)	(3.886)	(0.023)	(0.019)	(0.745)	(0.034)	(0.025)
FDR-adjusted <i>P</i> value	[0.001]	[0.453]	[0.453]	[0.053]	[0.804]	[0.484]	[0.160]	[0.319]
Baseline mean	0.017	−0.013	58.626	0.002	−0.087	19.028	0.007	0.019
End-line mean for control group	0.000	0.000	123.102	0.000	0.000	35.716	0.000	0.000
End-line SD for control group	1.000	1.000	174.204	1.000	1.000	35.519	1.000	1.000
Observations	13,066	15,184	15,221	15,184	15,197	15,191	13,058	13,066
<b>Drought and bad harvest interactions</b>								
Village offered VSLA	0.337***	−0.013	−3.335	0.039	−0.015	−1.271	0.124***	0.015
SE	(0.044)	(0.036)	(7.460)	(0.037)	(0.025)	(1.010)	(0.043)	(0.046)
FDR-adjusted <i>P</i> value	[0.001]	[0.749]	[0.749]	[0.583]	[0.749]	[0.557]	[0.019]	[0.749]
Village offered VSLA × drought event	−0.122	0.092	26.402*	0.032	0.005	−1.197	−0.119*	−0.024
SE	(0.080)	(0.057)	(15.002)	(0.058)	(0.043)	(2.084)	(0.065)	(0.088)
FDR-adjusted <i>P</i> value	[0.260]	[0.260]	[0.260]	[0.772]	[0.916]	[0.772]	[0.260]	[0.898]
Drought event	0.010	−0.119***	2.535	−0.044	−0.008	3.153*	0.058	0.030
SE	(0.056)	(0.043)	(10.748)	(0.043)	(0.041)	(1.656)	(0.041)	(0.063)
FDR-adjusted <i>P</i> value	[0.861]	[0.045]	[0.861]	[0.611]	[0.861]	[0.231]	[0.434]	[0.861]
Village offered VSLA × reported a bad harvest	−0.007	0.022	4.668	0.080	0.017	1.564	−0.055	0.041
SE	(0.055)	(0.045)	(9.559)	(0.049)	(0.027)	(1.503)	(0.053)	(0.061)
FDR-adjusted <i>P</i> value	[0.898]	[0.723]	[0.723]	[0.723]	[0.723]	[0.723]	[0.723]	[0.723]
Reported a bad harvest	−0.007	−0.590***	−19.408***	−0.048	−0.072***	−1.343	0.003	−0.007
SE	(0.038)	(0.036)	(7.080)	(0.033)	(0.027)	(1.288)	(0.040)	(0.045)
FDR-adjusted <i>P</i> value	[0.949]	[0.001]	[0.022]	[0.300]	[0.022]	[0.477]	[0.949]	[0.949]
Baseline mean	0.109	0.111	61.325	0.161	0.024	19.960	−0.122	−0.029
End-line mean for control group	0.090	0.239	147.631	0.012	−0.182	37.810	0.214	0.288
End-line SD for control group	1.074	0.975	201.660	1.022	0.736	37.494	0.947	1.084
Observations	6,404	8,297	8,298	8,296	8,298	8,298	6,403	6,404

Results presented are mean standardized intent to treat estimates, including indicator variables for country, a control for the baseline value of the outcome variable, and a control for the geographic units used for stratification. All indices are standardized with respect to the control group in that wave of data. For both household- and adult-level outcome measures, SEs are clustered at the village level. Drought event is an indicator variable for all villages that experienced a drought in the 12 mo preceding the end line. More details are in *Analysis*. The bad harvest variable is an indicator of household self-reports on whether they had a bad harvest in the 12 mo preceding end line. All dollar values are in 2011 US dollars PPP. In Uganda, baseline values of the food security index present a standardized count of the number of meals eaten the day before surveying; in Ghana, neither survey included data on kids skipping meals or households borrowing food. At baseline, the business outcomes index for the three countries does not include whether the business has any employees; for Uganda, it does not include the number of months that the business has been in operation over the preceding year. At baseline, the assets index for Ghana is a simple standardized count of the categories of assets owned by the household. At baseline, total consumption per capita for Ghana and Uganda does not include food expenses. At baseline, the women's empowerment index for Ghana does not include data on women's influence on business decisions or visits to friends and family; for Uganda, it does not include data on women's decision-making power with respect to food. Significance levels are given (\*10%; \*\*5%; \*\*\*1%). False discovery rate (FDR), adjusted *p*-values, also known as *q*-values, were used to correct for multiple hypothesis testing. These were calculated following the Benjamini–Hochberg method (18).

greater for female primary respondents in program areas. Total reported savings are also significantly higher by PPP 2011 US\$13.7 (SE = 4.488), equivalent to a 34.5% increase relative to savings balances of respondents in control areas. As expected, the program improves access to credit as well: 42.1% of women obtained a loan in the year leading up to the end-line survey, an 11% point difference (SE = 0.012) from the control group. The average amount borrowed in a year is 16.2% higher or PPP 2011 US\$6.6 (SE = 3.851) in program communities. Drought and bad harvest interactions in Table 1 find no evidence that the positive impact on financial inclusion is any stronger for households in treatment villages who experienced a poor harvest.

Table 1, food security index (household) shows that the program, on average, had no significant positive impacts on food security. A closer look at bad shocks (Table 1, drought and bad harvest interactions) shows that drought has a strong negative impact ( $-0.119$  SD, SE = 0.043) on food security for households in control villages as does a poor harvest ( $-0.590$  SD, SE = 0.036). There is no evidence that the program improves food security for treatment households reporting a poor harvest or those who experienced drought [although the point estimate for the latter is large enough to be important (0.092 SD with SE = 0.057); thus, we cannot rule out positive and important impacts on food security].

Similarly, the program does not show a positive effect on income for the overall sample. We do find, however, a positive program effect (PPP US\$26.4, SE = 15.002) for treatment villages suffering from a drought event, although it is not statistically significant after correcting for multiple hypothesis testing ( $q$  value = 0.260). When we analyze components of the household income measure in Table S5, we find that the program has an overall positive impact on monthly household business profits of PPP US\$5.6 (SE = 1.857) or 24.4%, but it does not affect income for other household activities. There is no evidence that the program improves income or revenues for households reporting a poor harvest. These findings suggest that the VSLA program has an overall effect on business profits and also insulates members from adverse aggregate shocks on their economic activity more broadly.

Next, we find a 0.06-SD (SE = 0.023) positive impact of the program on an index of business outcomes. Table S6 shows that VSLAs lead to a slight increase in the total number of businesses operated by the household (0.024 businesses, SE = 0.014; equivalent to a 6.3% increase). Households in treatment communities operated businesses 0.20 mo (SE = 0.098) longer than the control group, equivalent to an 8.6% increase. It should be noted that businesses operated by households in the sample are mostly short-term seasonal businesses, with an average total of 2.33 mo of business operation per household in the control group. Although it is fairly uncommon for businesses to use outside labor, we find that VSLAs lead to a 1.0% point increase (SE = 0.004)—a 26.3% increase compared with a 3.8% control group average in the number of households with at least one employee in its business(es). Thus, the VSLA program stimulates investments to extend and expand businesses operated by the households. Information on drought and bad harvest interactions in Table S6 shows no evidence that these positive effects on business activities are any different among households that experience a bad harvest or in communities that experience drought.

Table 1, total asset index (household) and Table S7 find no impact on total assets owned by households nor any differential impacts for those experiencing bad shocks.

Table 1, monthly per capita consumption (\$; household) and Table S8 find no impact on consumption in both the aggregate or the component measures, nor do they find a differential impact for those experiencing drought. Here, it should be noted that the consumption measure refers only to 30 d preceding the interview and therefore, could miss consumption smoothing benefits.

Table 1, women's empowerment index (adult female respondent) shows a significant impact of the VSLA program on the

empowerment of female primary respondents. Surveyed women in program communities display a 0.06-SD (SE = 0.034) increase in their influence on household decision-making. Table S9 shows that this impact is strong across the board, with a 4.2% point (SE = 0.016) improvement in the share of women who have a high degree of control over household business decisions, a 3.7% point (SE = 0.014) increase in control over food expenses, and a 2.9% point (SE = 0.015) increase in women with influence on education expenses. The results in Table 1, drought and bad harvest interactions suggest that this increase in influence occurs only in communities that are not experiencing a drought: there is a  $-0.119$ -SD (SE = 0.065) decrease in female influence on household decision-making in treatment communities that suffer from drought.

Finally, we do not find evidence of a similar impact on female primary respondents' community participation. However, Table S10 does show a significant 2.3% points (SE = 0.011) increase in the number of women who attended a community meeting in the last 12 mo.

Datasets S1–S7 provide some additional results of interest, such as balance and attrition analyses, and breakdowns of impact by country and components.

## Discussion

We find important impacts after 2–3 y of treatment beginning: VSLAs facilitate investment and empower women. We find no evidence of differential impacts on any of the outcomes after an idiosyncratic agricultural shock (experiencing a bad harvest after controlling for bad rainfall). It is possible that the VSLA treatment could affect the likelihood of a poor harvest conditional on the aggregate shock by changing agricultural investment or technology. In this case, the unobserved characteristics of farmers reporting a poor harvest conditional on bad rainfall in a VSLA community might be different from those of farmers reporting the same in a control community, making the coefficient of this interaction difficult to interpret. However, we find no evidence of a VSLA impact on farming activities (Table S5, all countries), suggesting that any such selection effect is minimal.

The positive impact on income after bad rainfall (an aggregate community shock) could suggest that VSLAs may work through a savings mechanism, not merely an insurance mechanism that builds risk-sharing within the community. Naturally, the two are not mutually exclusive channels through which the program can generate impact. Given that these results do not maintain statistical significance after correcting for multiple hypotheses, we encourage replication of these evaluations.

Beyond establishing the base evidence for the average impact of these programs, three key and related questions remain. What exactly is new? Why does the specific pattern of impacts vary across countries? What is the mechanism through which these programs generate change? A VSLA requires no external capital and no legal infrastructure to issue and collect loans and is, in fact, similar to existing informal mechanisms, such as ROSCAs. However, average impact ensued. The average household in a community offered an opportunity to form a VSLA increased its financial savings by almost 35% or \$14. Average household monthly consumption expenditure is about \$200, and therefore, this change in savings balances is significant, albeit unlikely to be transformative. Therefore, it is unsurprising that there is no evidence of average impacts on household income, consumption, food security, or asset ownership. We instead find changes in the realms most directly linked to the VSLA: non-farm business operations and women's empowerment.

However, these effects vary across populations and countries. A detailed exploration of differences in implementation across the three countries may help us understand the mechanism through which it generates the average impacts that we have documented. A savings-led microfinance group might be thought of as a new "social" technology, one for which all of the necessary components existed beforehand but that still required some external agent to promote and train groups. Indeed, researchers in Mali found that, when villages

started these groups via word of mouth rather than through direct NGO promotion and training, they did not work as well (14).

In terms of mechanisms, we start with the simple: the near-zero transaction costs (i.e., no travel time required, merely time to attend a meeting) can explain its impact vis à vis formal banking. However, low costs do not explain the impact when compared with preexisting informal options, such as cash stored at home. Compared with these options, savings-led microfinance groups can be best thought of as a social commitment savings account. This logic relies on models of temptation, time-inconsistent preferences, or household bargaining inefficiencies to explain the demand for commitment. Furthermore, although some informal savings vehicles do have a form of commitment (e.g., livestock), such vehicles can be costly and not as effective as a commitment device that uses social capital to help ensure compliance.

Last, we consider the benefits relative to the costs. We find no evidence that the VSLA program changes average income or consumption, which poses complications for a benefit–cost analysis. The benefits of the program are subtle: possible improved income in the face of drought, empowerment of women within the treated communities, and improvements in a broad array of business outcomes. Asserting a monetary value for these improvements would require a series of strong modeling assumptions that would of necessity be tentative and are beyond the scope of this paper.

Despite the lack of a benefit-cost analysis, it is important to note that the program's cost per household is low. We obtained data on program costs from Malawi and Uganda as of June of 2011, but such data were unavailable for Ghana. The most conservative estimate yields an average cost per member of \$26 in Malawi and

\$20 in Uganda over 22 mo of program implementation (not including the opportunity cost of the participant's time).<sup>11</sup> Considering the take-up rates of 22 and 36% among female primary respondents in Malawi and Uganda, respectively, and the fact that 20 and 36% of VSLA members are men in Malawi and Uganda, respectively, the most conservative estimate of the costs of the program per household in the cluster would be based on the assumptions that all of the male members have the same take-up rate as that evidenced in female members and that the former come from different households than the latter. Under these assumptions, the average costs of the program per household in Malawi and Uganda are \$7 and \$11, respectively.\*\*

**ACKNOWLEDGMENTS.** We thank CARE and its partners for their collaboration and support in the study. We also thank all of the staff of Innovations for Poverty Action for fieldwork and research assistance. We thank S. Amadou, A. Bartik, M. Bucciarelli, E. Cupito, Y. Chuang, M. Dieci, S. Fontenay, A. Garcia Vargas, M. Ghahremani, M. Hoover, S. Khan, E. Koshy, R. Lokur, R. Naimpally, L. Pappagallo, E. Salgado, S. Stephen, R. Strohm, and N. Torres. We thank G. Grossman (University of Pennsylvania) and M. Ngatia (Tufts University) for their intellectual input on parts of this research. This research was funded by CARE International.

<sup>11</sup>The cost calculation took into account central management costs (project management and administration, learning, and indirect costs) as well as operational costs. Unfortunately, we did not receive any detailed information on these components and were, therefore, unable to carry out a more comprehensive analysis of program costs. For context, the website [www.vsla.net](http://www.vsla.net) indicates a \$22 global average cost per member.

\*\*The household uptake rate was first estimated by dividing the female uptake rate in each country by the ratio of women among VSLA members in each country. Then, estimated costs per household in the treatment cluster were calculated by multiplying the household uptake rate by the average cost per member in each country.

- Collins D, Morduch J, Rutherford S, Ruthven O (2009) *Portfolios of the Poor: How the World's Poor Live on \$2 a Day* (Princeton Univ Press, Princeton).
- Karlan D, Ratan AL, Zinman J (2014) Savings by and for the poor: A research review and agenda. *Rev Income Wealth* 60(1):36–78.
- Banerjee A, Karlan D, Zinman J (2015) Six randomized evaluations of microcredit: Introduction and further steps. *Am Econ J Appl Econ* 7(1):1–21.
- Banerjee A, et al. (2015) Development economics. A multifaceted program causes lasting progress for the very poor: Evidence from six countries. *Science* 348(6236):1260799.
- Mader P (2013) Rise and fall of microfinance in India: The Andhra Pradesh crisis in perspective. *Strat Change* 22(1–2):47–66.
- Allan A, Massu M, Svarer C (2013) *Banking on Change: Breaking the Barriers to Financial Inclusion (Plan, Barclays and CARE)*. Available at <https://www.home.barclays/content/dam/barclayspublic/docs/Citizenship/banking-on-change.pdf>. Accessed December 17, 2016.
- Angelucci M, Karlan D, Zinman J (2015) Microcredit impacts: Evidence from a randomized microcredit program placement experiment by Compartamos Banco. *Am Econ J Appl Econ* 7(1):151–182.
- Attanasio O, Augsburg B, De Haas R, Fitzsimons E, Harmgart H (2015) The impacts of microfinance: Evidence from joint-liability lending in Mongolia. *Am Econ J Appl Econ* 7(1):90–122.
- Augsburg B, De Haas R, Harmgart H, Meghir C (2015) The impacts of microcredit: Evidence from Bosnia and Herzegovina. *Am Econ J Appl Econ* 7(1):183–203.
- Banerjee A, Duflo E, Glennerster R, Kinnan C (2015) The miracle of microfinance? Evidence from a randomized evaluation. *Am Econ J Appl Econ* 7(1):22–53.
- Crépon B, Devoto F, Duflo E, Parienté W (2015) Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco. *Am Econ J Appl Econ* 7(1):123–150.
- Karlan D, Zinman J (2011) Microcredit in theory and practice: Using randomized credit scoring for impact evaluation. *Science* 332(6035):1278–1284.
- Tarozzi A, Desai J, Johnson K (2015) The impacts of microcredit: Evidence from Ethiopia. *Am Econ J Appl Econ* 7(1):54–89.
- Beaman L, Karlan D, Thuysbaert B (2014) *Saving for a (Not So) Rainy Day: A Randomized Evaluation of Savings Groups in Mali*. Available at [www.nber.org/papers/w20600](http://www.nber.org/papers/w20600). Accessed February 13, 2015.
- Ksoll C, Lilleør HB, Lønborg JH, Rasmussen OD (2016) Impact of village savings and loans associations: Evidence from a cluster randomized trial. *J Dev Econ* 120:70–85.
- Greaney B, Kaboski JP, Van Leemput E (2016) Can self-help groups really be “self-help”? *Rev Econ Stud* 83(4):1614–1644.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol* 57:289–300.
- Novella NS, Thiaw WM (2012) African rainfall climatology version 2 for famine early warning systems. *J Appl Meteorol* 52:588–606.
- Helmore K (2010) Village agents: Self-replication and sustainability of CARE's village savings and loan associations. *Access Africa Technical Learning Series*, ed Chidiac S (CARE USA, Atlanta), No 1.
- Burlando A, Canidio A (2017) Does group inclusion hurt financial inclusion? Evidence from ultra-poor members of Ugandan savings groups. *J Dev Econ*, in press.
- Dupas P, Karlan D, Robinson J, Ubfal D (2016) *Banking The Unbanked? Evidence from Three Countries*. NBER Working Paper No. 22463 (National Bureau of Economic Research, Cambridge, MA). Available at [www.nber.org/papers/w22463](http://www.nber.org/papers/w22463). Accessed February 22, 2017.