Journal of Development Economics 120 (2016) 70–85

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Impact of Village Savings and Loan Associations: Evidence from a cluster randomized trial☆

Christopher Ksoll [a], Helene Bie Lilleør [b], Jonas Helth Lønborg [c], Ole Dahl Rasmussen [c,d,*]

[a] 120 University (8007), Ottawa, ON, K1N6N5, Canada
[b] Rockwool Foundation, Sølvgade 10, 2nd floor, DK-1307 K Copenhagen, Denmark
[c] University of Southern, Denmark, Denmark
[d] DanChurchAid, Nørregade 15, 1165 Copenhagen K, Denmark

ARTICLE INFO

Article history:
Received 9 September 2014
Received in revised form 17 December 2015
Accepted 17 December 2015
Available online 29 December 2015

Keywords:
Savings groups
Village Savings and Loan Associations
VSLA
Malawi
Impact evaluation
Cluster-randomized controlled trial

ABSTRACT

The vast majority of the world's poor live in rural areas of developing countries with little access to financial services. Setting up Village Savings and Loan Associations (VSLAs) has become an increasingly widespread intervention aimed at improving local financial intermediation. Using a cluster randomized trial we investigate the impact of VSLAs in Northern Malawi over a two year period. We find evidence of positive and significant intention-to-treat effects on several outcomes, including the number of meals consumed per day, household expenditure as measured by the USAID Poverty Assessment Tool, and the number of rooms in the dwelling. This effect is linked to an increase in savings and credit obtained through the VSLAs, which has increased agricultural investments and income from small businesses.

1. Introduction

The vast majority of the world's poor live in rural areas of developing countries. There they endure long periods of time between inputs into and outputs from agricultural production, uncertainty about harvest outcomes, and dependency on the weather. This makes it crucial that they are able to smooth consumption, to access credit and to employ strategies for coping with risk (Conning and Udry, 2007). However, the history of rural financial intermediation is not encouraging, and even the explosive global growth in microfinance has been concentrated in urban and semi-urban areas (Allen and Panetta, 2010; Daley-Harris, 2009; Demirgüç-Kunt and Klapper, 2012).

When formal financial institutions are not available, households rely instead on networks, moneylenders and informal financial mechanisms, which exist in a myriad of forms (Collins et al., 2009; Rutherford, 2001). Some of the most well-known are perhaps ROSCAs (rotating savings and credit associations) (Besley et al., 1994).[1] In its simplest form, members contribute to a common pot of money that is awarded to a different member at each meeting. Apart from the first and last members, each member will be both a saver and a borrower in the course of a cycle (Anderson and Baland, 2002; Besley et al., 1994; Bouman, 1995; Klonner, 2008). Village Savings and Loan Associations (VSLAs) – a type of local financial institution that has been standardized by CARE and is increasingly being promoted by a host of different organizations – take many of the elements of ROSCAs and add more flexibility in savings and loans, standardize the governance structure and reinforce accountability elements. Each VSLA relies on its members'

☆ The authors wish to thank the implementing partners DanChurchAid and Livingstonia Synod Development Department (SOLDEV), in particular the field team led by Waluza Munthali, the Invest in Knowledge Initiative for data collection, the Rockwool Foundation for financial support for the implementation and evaluation of the VSLAs, and several research assistants, especially Morten Harding Jensen and Rachel Cassidy. We also thank participants at the CSAE 2012, NOVAFRICA 2012, and NEUDC 2013 conferences, UC Davis Agricultural and Applied Economics and DGPE workshops as well as Thomas Barnebeck Andersen, Nikolaj Malchow-Møller, Måns Söderbom, Peter Sandholdt Jensen, Henrik Hansen, Søren Leth-Petersen, Per Andersen, Roy Mersland, Dean Karlan and David McKenzie for useful comments. All errors are our own.

* Corresponding author.
E-mail addresses: cksoll@uottawa.ca (C. Ksoll), hbl@rff.dk (H.B. Lilleør), jonas.helth@gmail.com (J.H. Lønborg), odrasmussen@gmail.com (O.D. Rasmussen).

[1] Other examples are ASCAs (accumulating savings and credit associations), and Susu-collectors originating from Ghana ('Susu' meaning 'small small' in the Akan language in Ghana).

http://dx.doi.org/10.1016/j.jdeveco.2015.12.003

savings to provide credit (as well as small contributions to a limited fund which can be accessed in times of funerals or illnesses of household members). A minimum level of savings is compulsory at the weekly meetings. Typically once a month loans are made available to members from the VSLAs, and a stated aim of many VSLA projects is to encourage small businesses. VSLAs thus create a local financial market that allocates local savings to debtors who are group members. Moreover, VSLAs implement a number of accountability and governance features that are described in more detail below. These features lead to low cases of default or elite capture.

Once a year, all the savings and the accumulated interest are paid out. In our context, rural Malawi, a large number of groups plan this "share-out" just before the agricultural planting season, when seeds and fertilizer are bought. This also coincides with the beginning of the school year when education related expenses are high. There are few formal alternatives for saving and receiving credit in the villages we are studying: at baseline only 0.6% of households are participants in a formal microfinance institution, and 5 to 6% in a VSLA group for example.

VSLAs have now been introduced in 72 countries and have 11 million active participants worldwide.[2] However, despite this wide distribution and the increasing popularity among donors as a means of improving the financial infrastructure in remote rural areas, very little is known about the impact of VSLAs on household welfare. Unlike other types of microfinance for which several randomized impact studies have been made showing limited effects (see Banerjee (2013) for a review of microfinance, and Banerjee et al. (2015c) for a review of microcredit), rigorous impact evaluations of VSLAs that address endogenous program placement and selection bias have only recently been completed.[3]

We provide a rigorous impact assessment of VSLAs on pre-defined household outcomes. We do this through a cluster randomized control trial carried out in forty-six villages in northern Malawi. Among these forty-six villages, twenty-three villages were randomly selected to participate in a VSLA project implemented by a local NGO, Soldev, from 2009 onwards. The remaining twenty-three villages were used as a control group, before the project was finally extended to them in 2011. We conducted household surveys of 1775 households in 2009, just before the project was implemented in the treatment villages, and again in 2011, just before the intervention was introduced in the control villages. We carefully tracked households that moved between the 2009 and 2011 surveys, resulting in a low attrition rate of only 3%. By 2011, 45% of the households in treatment villages had joined VSLAs, while almost 21% of households in control villages had also joined. Despite the implementing NGO adhering strictly to the randomization design, there was spontaneous spill-over into control villages — typically because control village inhabitants decided to join VSLA groups in treatment villages.

We assess the impact of introducing VSLAs at village level on seven general indicators of household welfare. These were pre-defined as targets by the implementing partner as part of the project design. We operationalize the seven indicators by one or two outcome variables each (see Table 2) leading to a total of nine final outcome variables. We also provide multiple hypothesis corrected p-values (denoted by q in the following).

We find evidence of positive impacts on four out of nine outcome variables. Food security was improved in that the number of meals consumed per day increased by 0.13 (s.e. = 0.05, p = 0.02, q = 0.09)[4], corresponding to one more meal per week. Households were found to hold larger savings in total (p = 0.02, q = 0.07). Increases in households' scores on the USAID's Poverty Assessment Tool (PAT)[5] suggest that expenditures increased by about 4.2% (s.e. = 0.02, p = 0.05, q = 0.11). Finally, the average number of rooms per dwelling rose by 0.16 (s.e. = 0.06, p = 0.02, q = 0.09). Concurrently, there is some evidence that the number of income generating activities decreased, which could be consistent with increased specialization.

Overall, the results provide evidence that VSLAs can improve household outcomes, even in as short a period as two years, without any injection of outside capital. The impacts on participating households are likely to be larger, as our impact estimates are given by the intention-to-treat estimator, i.e. they are average impacts across both participating and non-participating households. Furthermore, there has been a non-negligible degree of spill-over into control villages and we may therefore be underestimating the actual impact size. The results of positive impacts are consistent with results in Beaman et al. (2014), who study VSLA groups in Mali. They also report positive impacts on savings (in particular livestock) and food security.[6]

In the literature, there are several suggestions for possible channels through which improved financial intermediation and participation in savings groups can have a positive impact on the participating households and local communities. For VSLAs to attract members, they need to offer higher interest rates or more secure repayments than alternative forms of savings. This is possible if groups are better at monitoring loans than individuals are, if groups have stronger enforcement mechanisms than individuals do, or simply because groups pool the risk on individual loans (Diamond, 1984).[7] In turn, the larger pool of savings coupled with better loan monitoring (and presumably lower interest rates on loans than those offered by money-lenders) expands access to credit within each VSLA, though only to the extent that savings have accumulated.[8,9] If households are constrained in their access to

[2] Based on information on the website www.vsla.net, maintained by VSL Associates, an organization that provides training in setting up VSLAs (site accessed on September 9th 2015).

[3] We know of a number of project reports that report on VSLA interventions. Three of these use randomized control trials. Annan et al. (2013) is an evaluation report on a study of VSLAs in Burundi; Beaman et al. (2014) study VSLAs in Mali. A larger international study covers Ghana, Malawi and Uganda, but at the time of writing, is not yet available for citation.

[4] We present the multiple-hypothesis corrected equivalent of p-values, the sharpened q-values (Anderson, 2008; Benjamini et al., 2006) alongside the p-values. See the methodology section below.

[5] The USAID Poverty Assessment Tool, usually abbreviated as PAT, is a list of twenty simple indicators that are chosen so as to measure poverty in expenditure as accurately as possible around the 1.25 USD a day poverty lines (see USAID, 2009 and www.povertytools.org/countries/Malawi/Malawi.html). Indicators included in the PAT for Malawi are chosen based on the ability to predict household poverty by predicting household expenditure using the Second Integrated Household Survey (a survey in the World Bank's Living Standards' Measurement Surveys series). The indicators include household member characteristics (the age and marital status of the household head, the size of the household, literacy and educational level of household members), household characteristics (the number of rooms, the type of floor, the type of toilet facility as well as electricity in the household), assets (ownership of bed, iron, radio, and bicycle respectively as well as raising or owning livestock in general and goats in particular) as well as other indicators (whether the household purchased any soap in the past month and whether the household grows a vegetable garden). The typical use of the tool is to measure average poverty status in a population. The calculation of poverty status involves as a first step predicting expenditure levels. We use the estimated expenditure levels directly, noting that the predictions are most accurate around the 1.25 USD poor/non-poor cutoff. Since the average (predicted) level of expenditures at baseline is around 1.17 USD per person per day, the prediction is likely to be reasonable good. While there is considerable documentation on how good PAT is at measuring expenditure-based assessments of poverty, as of now, there is no evidence on how good it is at measuring *changes*. It is likely that the measurement error introduced by the prediction would reduce the power to detect effects.

[6] Annan et al. (2013) document substantial effects of the VSLA program, though they note that their outcomes are mostly based on self-reports.

[7] By reducing the risk involved in savings, the not quite "risk-free" but lowest-risk option can also enable households to adjust their portfolio of activities and undertake more risky investments elsewhere, for example in businesses or agriculture.

[8] As such, this is quite different from a micro-finance institution expanding access in villages and providing new resources for credit, as studied in Fulford (2013). He finds that when household have precautionary saving motives and gain access to credit, there can be an initial consumption boom at the expense of precautionary savings. In VSLAs, this cannot be true on average.

[9] In addition, VSLAs typically offer a simple and small insurance component. The groups themselves decide on the exact nature of the insurance, but it almost always involves insurance against the illness or death of household members. That is, the insurance might be an additional device for coping with risk that can encourage households to discard inefficient ex-ante coping strategies, such as low risk/low return activities. If households choose activities involving a higher level of risk but higher expected returns, this can lead to increased consumption in the long run even if the insurance itself never has to pay out (Carter and Zimmerman, 2003; Dercon, 1996).

credit for profitable business purposes, an expansion in access to credit increases household income.

Furthermore, since VSLAs entail a minimum level of compulsory savings, participants with time-inconsistent preferences who realize their self-control problems may use VSLAs as a commitment device (Ashraf et al., 2006; O'Donoghue and Rabin, 1999). Even if participants do not have time-inconsistent preferences, VSLAs might allow them to commit the rest of their households in accordance with their own preferences (Anderson and Baland, 2002; Jackson and Yariv, 2014).[10] VSLAs disburse savings once a year, and typically time these share-outs to coincide with periods when most households need lump sums for investments in agriculture and education related expenses, facilitating such commitments.[11]

Although we do not aim to investigate all of these possible channels through which the observed impacts may have come about, we do investigate the impact on economic activities that VSLAs are meant to enable, i.e. where implementing NGOs think that there are large potential returns that households do not reap because of credit constraints. Beyond the increase in total savings, we find significant increases in the use of credit, including credit used for investment purposes. Likewise, respondents report that they use their savings upon share-out for investments, primarily in agriculture and business. We follow this flow of money into agriculture and small-scale businesses, finding a statistically significant increase in the use of fertilizers for maize. There is an increase in the value of maize sold, though no increase in maize production overall. We also find a significant increase in the total income from enterprises. Overall, the main channel through which the observed positive impacts on consumption seem to occur is through providing households with a means to save for investments, with the investments into businesses paying off.

2. The intervention and research design

2.1. The VSLA intervention

The VSLA intervention encourages the formation of savings groups with typically fifteen to twenty-five members each, and trains each group in the management of their own village savings and loan association. As no external capital is provided, the groups are essentially self-managed financial intermediaries.

Within the larger microfinance sector, community-managed microfinance belongs to the category of member-based, community-managed, accumulating microfinance institutions (see Table 1). Typically, microfinance impact evaluations have focused on professionally-managed microfinance institutions. Unlike such institutions, VSLAs do not depend on the injection of external funds, just as they do not depend upon the sustainability of a professionally-managed institution, but rather on the sustainability of the group formed within the local community.

The inspiration for VSLAs came from rotating savings and credit associations (ROSCAs) and was developed by CARE international and VSL Associates during the 1990s (Ashe, 2002). The aim was to improve on ROSCAs in two respects, namely by making the groups more sustainable and more flexible. Increased sustainability comes from a series of accountability features that prevent both the theft of funds and elite capture. Flexibility is increased in that members can at any time borrow any amount they wish up to three times the level of their own savings, provided that funds are available. Whereas ROSCAs multiply without external facilitation, VSLAs only do so to a much lesser extent, and thus require facilitation by, for instance, an NGO in order to spread rapidly. This is perhaps due to the rather complex governance features. VSLAs are implemented as described in Allen and Staehle (2007). After conducting awareness meetings in every targeted village, a local NGO facilitates the formation and training of groups. Participation in groups is free.[12] Initially, groups are visited every week in the first three months to set up the procedures. Groups work as member-owned financial intermediaries with three products: savings, credit and a smaller insurance component. Savings are compulsory; the amounts are collected at the weekly meetings and are conceptualized as buying shares. Every week, a member must buy at least one share and is permitted to buy up to five. The share value is set by the group and written into the group's constitution. It varied between MWK 50 and 100 in the groups in our study.[13]

Loans are available at every fourth meeting. If the funds requested by members exceed the amount of funds saved, the group decides who will be given loans by following a predetermined list of criteria written into the group's constitution. This typically assigns funds on the basis of the stated planned use of the loan. The interest rate on loans is set by the group, and can thus be used to regulate excess supply or excess demand in the medium term, making VSLAs less flexible in that regard than "bidding ROSCAs", which use a process of bidding for loans (Klonner, 2008). Usually, the nominal interest rate on loans is set at between 5% and 20% per month, but extensions in repayment schedules and inflation make the real interest rate considerably lower (Rasmussen, 2012). Loan contracts run for three months, with a grace period of one month. Rules for loan approval are set down in the group's constitution, but often focus on the requirement that the money be used for productive (income-generating) uses.

At the end of a cycle, which is around 12 months in duration, all savings and interest payments are divided by the number of shares and paid out — the "share-out". The average value of savings during the first share-out from all groups in the area was 6367 Kwacha (or 69.27 USD PPP-adjusted) which earned an average of 1801 MWK (19.59 USD PPP) in interest. The interest rate on savings typically works out at 4% to 5% per month, or about 60% per year (Rasmussen, 2012).[14]

The actual date of the final share-out is set by the group and tends to be chosen according to when most households need cash. In Malawi, this is most frequently at the end of the calendar year, when seeds and fertilizer need to be financed, as well as educational materials purchased. The NGO highlighted the sharing-out at the end of the year as a benefit from the participation in VSLAs.

At the end of a cycle, members decide whether to leave or remain in the VSLA group, and whether the group should accept new members. Any impact found in our analysis below will thus be the impact of at most two full cycles of collecting savings, distributing loans and returning the savings with interest among the 100 groups established by mid-2011. Usually after one year of initial training and monitoring, groups "graduate", which means that they cease to be supervised by the NGO that helped set them up.

[10] Indeed, Jackson and Yariv (2014) show that, even when an *individual's* preferences are not time- inconsistent, all *households* will exhibit time-inconsistent preferences, *unless* discount rates are equal across household members or decisions are dictatorial (i.e. one household member's preferences decide). Cultural norms would suggest that this would usually be the husband's preferences, which might be another reason for most VSLA members being women.

[11] If, for example, the share-out is at the beginning of the agricultural season, investment in agriculture might increase. Duflo et al. (2011) find that providing farmers with an opportunity to pay for fertilizer just after the harvest, when money is available, and having it delivered at the time of planting, increases fertilizer use dramatically. In addition, the timing of the share-out might serve as a "label", fixing the minds of the participants on using the money from the share-out on a particular asset — in this case, fertilizer or other lump-sum investments in agriculture (Thaler, 1990). These increased investments in, for example, agriculture can increase productivity and in turn lead to higher household income and an increase in household welfare.

[12] Greaney et al. (2013) provide evidence that charging for participation might lead to improved outcomes in a similar type of savings group, as instructors have an incentive to provide better training.

[13] Throughout the paper we report monetary values in Malawian Kwacha (MWK). In 2009, 1 USD corresponded to 91.91 MWK using poverty-adjusted PPP exchange rates in 2009.

[14] It is lower than the interest on loans primarily due to the facts that not all the funds are lent out all the time and that savings accumulate over time.

Table 1
Microfinance overview.

Ownership	Management	Fund Accumulation	Examples
Member-based institutions	Professionally-managed microfinance	Large accumulating savings and credit associations (ASCAs)	FECECAM, Benin Savings and Credit Cooperatives (SACCOS)
	Community-managed microfinance	Rotating savings and credit associations (ROSCAs)	Tontines, susu, upatu, merry-go-rounds, chit, pasanakus
		Small accumulating savings and credit associations (ASCAs)	Village savings and loan associations, Savings for Change, the WORTH model, savings and internal lending communities
For-profit institutions	Professionally managed		FINADEV, Benin Equity Bank, Kenya SKS, India
Non-profit institutions	Professionally managed		Grameen Bank, Bangladesh BRAC, Bangladesh Opportunity International

Apart from savings and loans, VSLAs also offer a smaller insurance component, financed by a small premium paid by each member each week separately from the savings and credit activities. The insurance is paid out as a transfer or an interest-free loan when certain events occur that are outlined in the constitution, usually the death of family members, death of cattle, sudden illness, or other emergencies.

A number of governance features guard against theft and elite capture. With the help of the facilitator, each group develops a constitution that describes areas of potential conflict and their solution, for example lending rules, election procedures, exclusion of members, and fines for delays and non-attendance. To achieve transparency in transactions, all transactions take place in the presence of all members at the weekly meetings, and funds are counted independently and in public by two elected money counters. Between meetings, funds are stored in a cashbox, locked with three padlocks. Three different members each hold a key to the box and the box is stored in the house of a fourth, so that the possibility of collusion for purposes of theft is greatly reduced. The identity of the "guardian" is kept secret from people outside the group, providing a degree of protection against theft by outsiders.

2.2. The design of the experiment

The crucial challenge for an impact evaluation is to construct a credible counterfactual that is not sensitive to selection bias, arising typically due to non-random program placement and self-selection into program participation (Angrist and Pischke, 2009; Banerjee and Duflo, 2009; Duflo et al., 2007).

We addressed the problem of non-random program placement by randomizing the roll-out of the VSLA intervention at the village level, i.e. a cluster randomized control trial. Out of 46 villages in the program area, we randomly chose 23 villages for implementation in the first year (the treatment villages) and 23 villages for implementation in the third year (the control villages). To improve the balance in observable and unobservable characteristics we grouped villages into strata based on certain characteristics believed to be correlated with central outcomes of interest (Bruhn and McKenzie, 2009).[15]

The randomization was carried out under our supervision by field officers from the NGO, who drew village names from seven hats containing the villages in each stratum. Fig. 1 below shows the physical location of the village centers, with the shape of the inner symbol indicating the stratum to which the village belongs, and the outer symbol indicating allocation to the treatment or control group, respectively circle and square.

As part of the VSLA intervention, the implementing partner carried out awareness meetings in *all* 46 villages prior to randomization. A more detailed description of the project introduction activities is contained in Appendix A. We used these awareness meetings to have field officers from the implementing partner collect lists with the names of the villagers who expressed an interest in joining the VSLA groups to allow for stratification on initial interest. Fig. 2 provides an overview of the timing of the randomization, the implementation and the data collection.

The unusual element of asking respondents both in the treatment and control group whether they were interested was done to increase power by oversampling respondents more likely to participate since take-up is often found to be low in microfinance projects.[16] Appendix Table A1 shows that the being on the initial list of interested households does increase the participation probability as expected. However, households not on this list also had substantial participation rates. In the control villages, the probability of participation is 15 percentage points higher for those households on the initial list of interested household, while in treatment villages being on the list raises participation by 25 percentage points relative to households not on the list of initially interested households.

There are two major downsides to this design: First, since oversampling necessitates weighted regressions, a mechanical consequence is that the gains of over-sampling are not as large as one might think, in particular since participation rates among those *not* on the list of interested households turned out to be quite high, between 35 and 40%. Second, asking the control group whether they are interested can alter the behavior of the control group and can lead to ethical issues unless combined with a roll-out design (i.e. the control group will eventually also be treated).

As a consequence, the treatment effect could be biased if the control group alters its behavior in anticipation of later participation in the treatment.[17] In our case, households in control villages may wait to invest in businesses, hoping to get better interest rates on credit once members of a VSLA; disinvest from their buffer stock savings in the hope that the insurance mechanism from the group they might join in the future would cover them; or the initial awareness meetings in all villages may have inspired households from control villages to seek participation in VSLA groups in treatment villages. In fact, more than 20% of households in control villages have become members of a VSLA before the roll-out into control villages started in 2011, see Table 4. This further reduces the power that was to be gained from oversampling interested households. In hindsight, the range of problems the announcement of the program might bring, in particular to such a complex topic as financial markets, suggests that it should be used with care.

[15] The non-overlapping strata were defined as follows: large fishing villages, small fishing villages, particularly eager villages (identified by the field officers based on the reaction from villagers at the awareness meetings), large non-fishing villages, villages with a rice irrigation scheme, villages with another NGO-led intervention, and a final group of the remaining villages.

[16] See for example Banerjee et al. (2015a) for an evaluation of a microfinance project that led to an increase in microcredit of 8.5%.

[17] This is slightly different from the John Henry effect which refers to the control group trying to prove themselves by over-performing. Here they could also be thought to "under-perform", by not investing for example.

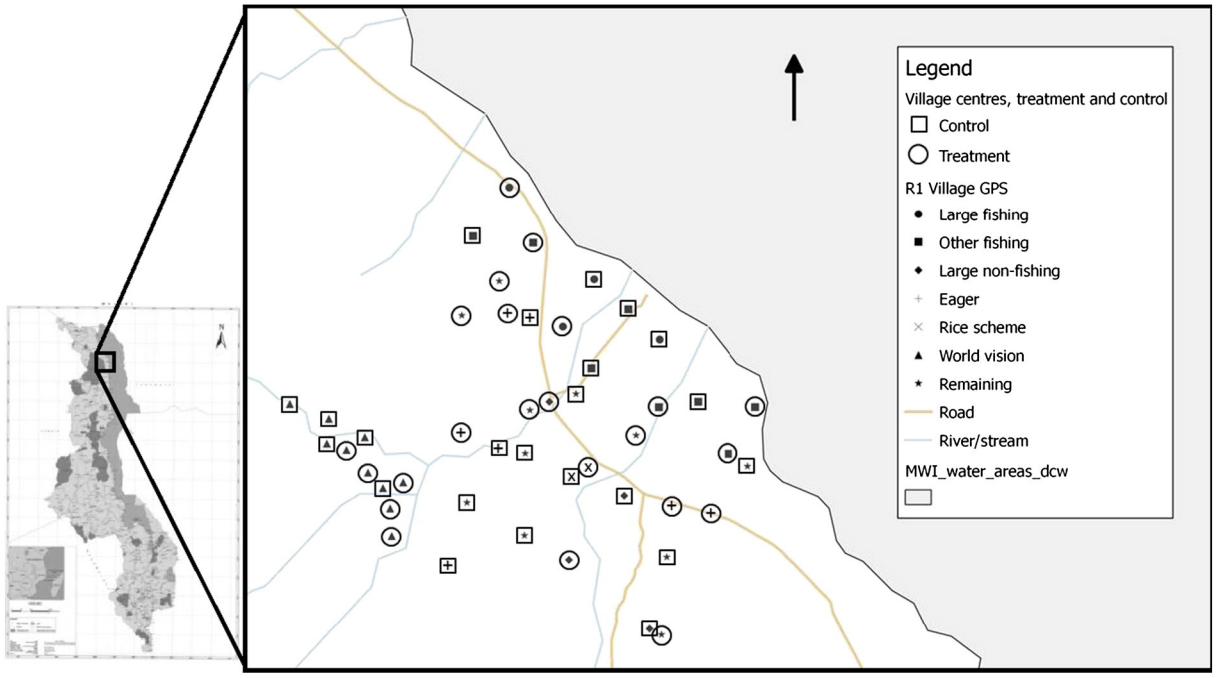


Fig. 1. Map of treatment and control villages in randomization blocks.

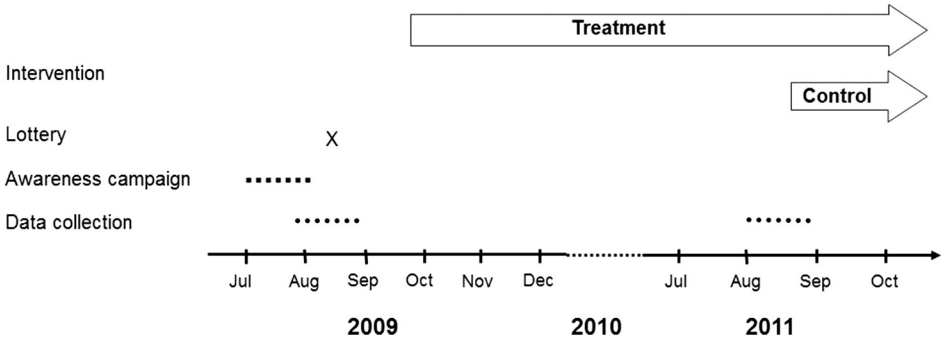


Fig. 2. Timeline of intervention and data collection.

2.3. Outcomes of interest

The final outcome variables for the primary impact analysis were closely linked to the development outcomes originally stated by the implementing NGO in the project's logical framework analysis (LFA) matrix, in which the NGO described the desired and anticipated outcomes and associated assumptions.[18] The matrix shows indicators on which the intervention is expected to have an impact, but it does not specify exactly how each of these indicators is to be measured. We select final outcome variables which come close to the original specification (see Table 2).[19] Although this is not equivalent to having laid out a pre-analysis plan as suggested by Casey et al. (2012), we follow this strategy to effectively tie the analysis – and our hands – to these outcome measures.

We do, however, also report extensively on non-predefined outcomes in order to investigate possible channels of the observed impact. Below we explicitly distinguish between the predefined primary analyses following the LFA indicators and the non-predefined secondary analyses investigating mechanisms.

3. Data and estimation strategy

3.1. Data

To enable oversampling of households which had expressed an interest in participating in the groups to be formed, we stratified the sample within each village by whether the household had declared an interest in participation, using the information gathered by the NGO during the awareness meetings. We sampled roughly 38 households in total in each village.[20]

[18] The logical framework matrix is a part of the logical framework analysis (LFA), by far the most common project management tool used in international development (Dale, 2003). The matrix is made by the implementing partner prior to initiating a project and specifies the expected impacts of the intervention. Since the matrix is made by practitioners rather than researchers, it is not very detailed when it comes to measurement and analysis.

[19] For a variety of reasons we did not collect information necessary to compute a total yearly household income measure.

[20] Due to differences in village size, our procedure led to considerable variation in sampling probability between observations, and we used weighted regressions throughout. There is no agreement on whether to use sampling weights or not (see Angrist and Pischke, 2009). If the population regression function is interpreted as causal in all its parameters, as would be the case in structural equation modeling, then weighting is in some cases not justified, regardless of the underlying method of sampling (Cameron and Trivedi, 2005). If, however, there is only one parameter of interest, for example because the causal interpretation stems from exogenous variation rather than from a correct structural specification, weights should be applied. This is the case within a counterfactual framework, or more specifically the Rubin causal model (Angrist and Pischke, 2009; Rubin, 1974).

Table 2
Predefined outcome measures and corresponding outcomes for the analysis.

Logical framework indicator (Soldev's choice)	Corresponding final outcome variable (authors' choice)
Hungry period is reduced	1) Number of months in past year with less than three meals per day.
Increase in the consumption of food	2) Number of meals previous day. 3) Food consumption per week per adult equivalent (log)
The average number of IGAs carried out by the VSLA participants has increased	4) Number of income-generating activities in each household.
Increase in the volume of savings by the VSLA groups from project-related activities by 2012	5) Total savings (log)
The share of the targeted population living below USD 1.25/day has decreased as measured using USAID's PAT	6) USAID PAT's prediction of per capita expenditure (log)
HHs have improved their housing standards	7) Number of rooms in dwelling 8) House has cement floor
Increase in household asset ownership	9) Asset count

Data were collected in 2009 and 2011 by the Invest in Knowledge Initiative under the supervision of the authors. In 2009, data collection took place from July 26 to August 30. The endline data was collected between July 8 and August 14, 2011. In both the 2009 and 2011 data collection, 24 interviewers completed interviews with up to 1775 households. A sub-sample of 834 households was also interviewed using a longer questionnaire that went into greater depth, in particular with respect to more detailed questions on loans and business activities. Fig. 1 shows the geographical locations of the villages in which the project was implemented surveyed. As can be seen from the figure, the treatment villages (marked by circles) and control villages (marked by squares) are contiguous, which made it possible for some households from control villages to participate in VSLA groups in the treatment villages, an issue which we address below.

To limit attrition, we implemented a tracking survey in 2011, with the result that a total of 1715 households were surveyed in that year. The attrition between the 2009 and the 2011 survey rounds was only 60 households, or less than 4% of the initial sample, which compares favorably with other panel surveys (Glewwe and Jacoby, 2000).

Since we also tracked households where one of the designated respondents had moved out due to some reason such as divorce, we have a number of split households, i.e. instances where one household from the 2009 survey had become two households in 2011. In the estimations below we drop one of these new households at random to ensure a balanced panel. We assess the implications of this for the results in the robustness section by limiting the sample to intact households.

3.2. Baseline balance

We test whether the randomization of villages into treatment and control groups did in fact lead to two groups with the same observable characteristics.

Table 3 below provides descriptive statistics for our final outcome variables and some characteristics of the households for the entire sample of households in 2009. Men headed 85% of households in the sample. Household heads on average had just below seven years of education. Households had an average of just below six household members each who, at the time of the interview, had a level of expenditures of approximately USD 1.17 per person per day (2005 values) assessed using the USAID PAT measurement,[21] and ate fewer than three meals a day for an average of more than four months of the year.

Columns four and five report the mean values of the variables for the households in the treatment villages and control villages respectively. Column six shows t-values for the differences between treatment and control villages, indicating whether the differences are statistically significant. These t-values were calculated by implementing an OLS regression which controls for randomization strata while using sampling weights and clustering at the village level. None of our predefined outcomes or observable household characteristics are significantly different at a conventional 5% level and the treatment and control samples appear reasonably balanced. Only the number of income generating activities is slightly lower in the treatment than in the control villages, which is significant at the 10% level of significance. It should be noted, that total savings were only collected for the subset of households, which were administered the longer in-depth questionnaire mentioned above.

3.3. Empirical strategy

The randomization allows for simple estimation strategies to be used to estimate the treatment effects of introducing VSLAs to villages, as it ensures that without treatment, the expected outcomes are the same in the treatment villages and in the control villages. Several different estimators can be employed to estimate the intention-to-treat effect.

We first estimate the intention-to-treat (ITT) effect, δ^{DM}, i.e. the average effect of introducing the VSLA intervention for all the households in the treatment villages compared to all the households in the control villages, irrespective of whether the households actually participated in VSLAs as follows:[22]

$$y_{ij} = \alpha + \delta^{DM} VSLA_j + \theta Block_j + \varepsilon_{ij}$$

where y_{ij} is the outcome for household i in village j, $VSLA_j$ is a dummy indicating whether village j was assigned to participate in the VSLA program, and δ^{DM} is the difference-in-means ITT estimate.

We include the randomization strata, termed blocks, in all our linear regressions, as recommended by Duflo et al. (2007). We implement weighted regressions using the inverse sampling probability as the weight for each household, and cluster standard errors at the village level. These sampling weights account for the fact that we over-sampled interested households.

Our second specification includes the lagged value of the dependent variable.

$$y_{ij} = \alpha + \delta^{DMlag} VSLA_j + \tau y_{ij-1} + \theta Block_j + \varepsilon_{ij}$$

Third, we use the pooled difference-in-differences estimator (Angrist and Pischke, 2009):

$$y_{ijt} = \alpha + \beta VSLA_j + \gamma post_t + \delta^{DiD}(VSLA_j \cdot post_t)_{jt} + \theta Block_j + \varepsilon_{ijt}$$

where $post_t$ indicates whether the observation is from the 2011 post-treatment survey, and $(VSLA_j \cdot post_t)_{jt}$ is the interaction of the two — i.e. whether the observation is from a treatment village in the 2011 survey. δ^{DiD} is our parameter of interest — the ITT using pooled difference-in-differences.

Finally, we take out any time-invariant unobserved heterogeneity at the household level through first-differencing.

$$\Delta y_{ij} = \alpha + \delta^{FD} VSLA_j + \Delta\varepsilon_{ijt}$$

δ^{FD} gives us the ITT estimate using first differences, which is our preferred estimate for uncensored outcomes.

[21] Throughout the analysis, we used the 2005 poverty-adjusted purchase power parity exchange rates suggested by Deaton and Dupriez (2011).

[22] Note that because the compliance in control villages was imperfect, this will underestimate the impact of VSLAs.

Table 3
Baseline characteristics and balance between treatment and control groups.

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	N	Mean	SD	Treatment average	Control average	Difference (t-value)
Project outcomes						
Number of months with fewer than three meals a day	1737	4.10	4.03	4.26	3.92	1.18
Number of meals yesterday	1737	2.65	0.56	2.61	2.70	1.45
17-Food consumption per week per adult equivalent (MK, log)	1737	6.10	0.58	6.07	6.14	1.52
Number of income-generating activities (including agriculture and livestock)	1737	1.99	1.10	1.94	2.04	1.73*
Total savings (log)	576	7.87	1.66	7.77	8.00	0.59
Per capita expenditure predicted by USAID PAT (log)	1737	0.16	0.42	0.16	0.16	0.39
Size of house (number of rooms)	1737	2.75	1.25	2.73	2.77	0.37
House has cement floor	1737	0.10	0.30	0.11	0.09	1.47
Asset count	1737	12.89	7.87	12.84	12.94	0.46
Other household characteristics						
Age of household head	1726	38.94	15.34	39.00	38.88	0.15
Household size	1737	5.77	2.46	5.71	5.83	0.68
Household is female-headed	1737	0.15	0.36	0.17	0.14	0.87
Years of education in household	1734	6.87	3.26	7.07	6.65	1.61
Household owns land	1732	0.96	0.19	0.95	0.97	1.48
Household is member of VSLA	1734	0.06	0.23	0.06	0.05	0.66

Notes: Table presents the mean and standard deviation of baseline characteristics in the sample as a whole (columns 2 and 3), as well as the mean separately by treatment status in columns 4 and 5. Displayed results are on the full sample of 1737 households for which we have information for all pre-defined outcomes. Total savings information was only collected from the subset of 834 households, which were administered the longer questionnaire. Log (savings) only presented for observations with non-zero savings. Observation numbers are reduced when there is missing information in a central variable. Column 6 reports the absolute value of the t-statistic from a weighted regression of the dependent variable on an indicator variable for VSLA and stratification fixed effects, testing for baseline difference in means between treatment and control villages. Huber–White standard errors clustered at the village level presented in parentheses. * $p < 0.1$.

3.4. Non-linear models

We implement tobit regressions for censored variables on a logarithmic scale (Cameron and Trivedi, 2005). As it is not obvious how to provide the correct estimate and inference in difference-in-differences and difference-in-means with lagged variable tobit models, we focus on difference-in-means models for variables on a logarithmic scale.[23]

When we investigate impacts on the profitability of businesses, we implement median regressions for business income, as we have a large proportion of zeros (suggesting a censored model). However, tobit regressions are inappropriate as business incomes can be negative. The OLS estimates are sensitive to outliers, resulting in large standard errors. The quantile regressions are robust to the large positive and negative values we observe, and are useful at establishing whether there is any impact of VSLAs on business income.

3.5. Multiple hypothesis testing

When multiple hypotheses are tested, the probability of rejecting at least one null hypothesis is higher than the significance level, even when a program has no effect on any of the outcomes (see, for example, Anderson, 2008). To deal with this issue, a number of techniques have been developed, the most famous of which is the so-called Bonferroni correction. Bonferroni's correction is however excessively conservative, leading to a substantial loss of power. Anderson (2008) provides an overview and implements a range of methods to correct standard errors for multiple hypothesis testing, including controlling for the false discovery rate (Benjamini and Hochberg, 1995), a procedure recently used by Banerjee et al. (2015b). We present the sharpened q-values of Benjamini et al. (2006) and Anderson (2008) for our main results, see Appendix B for more detail.

[23] See Ai and Norton (2003) and Puhani (2012). The Stata program for the fixed effects tobit estimator based on Honoré (1992) did not converge in our case.

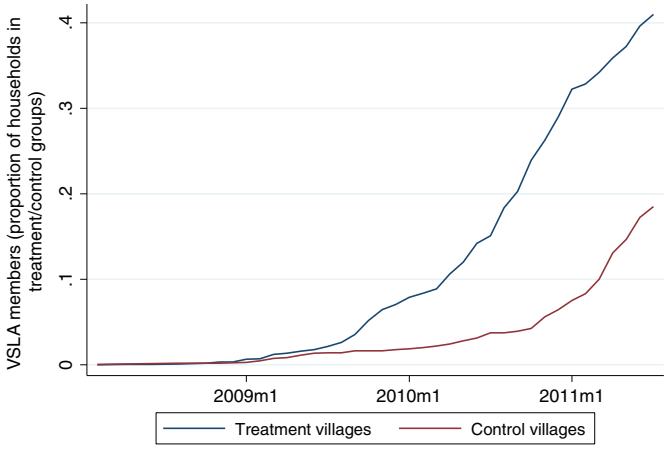


Fig. 3. Cumulative VSLA membership by village category. Notes: This figure presents the timing of membership in VSLA in treatment and control villages. Sample observations are reweighted so that they represent (estimates of) the proportion of households. Information on the timing of membership is based on survey data.

4. Results for main outcomes

4.1. Take-up

Before assessing the impact of the intervention, we describe how successful it was in attracting participants to form VSLAs. Fig. 3 gives

Table 4
VSLA membership.

	Control villages	Treatment villages	Differences
Baseline (2009)	0.052	0.061	0.010
Endline (2011)	0.207	0.451	0.259***
Difference	0.155***	0.390***	0.235**

Notes: Standard errors used in calculating significance are clustered at the village level. ** $p < 0.05$, *** $p < 0.01$.

Table 5
Effects on predefined outcomes.

Outcome	(1) Difference in means	(2) Difference in means with lag	(3) Difference-in-difference	(4) First-difference
Number of months with fewer than three meals a day	−0.203	−0.272	−0.493	−0.555
	(0.38)	(0.36)	(0.45)	(0.38)
	[0.703]	[0.35]	[0.303]	[0.181]
Number of meals yesterday	0.055	0.069**	0.145**	0.126**
	(0.03)	(0.03)	(0.07)	(0.05)
	[0.378]	[0.121]	[0.252]	[0.091]
17-Food consumption per week per adult equivalent (MK, log)	0.026	0.046	0.080	0.097
	(0.05)	(0.05)	(0.07)	(0.06)
	[0.703]	[0.326]	[0.303]	[0.181]
Number of income-generating activities (including agriculture and livestock)	−0.179***	−0.144***	−0.104	−0.060
	(0.06)	(0.05)	(0.11)	(0.07)
	[0.034]	[0.063]	[0.338]	[0.324]
Total savings (log)	1.034**			
	(0.43)	*n.a.*	*n.a.*	*n.a.*
	[0.07]			
Per capita expenditure predicted by USAID PAT (log)	0.030	0.037*	0.041*	0.042**
	(0.03)	(0.02)	(0.02)	(0.02)
	[0.53]	[0.121]	[0.252]	[0.107]
Size of house (number of rooms)	0.104	0.133*	0.135*	0.158**
	(0.10)	(0.07)	(0.07)	(0.06)
	[0.53]	[0.121]	[0.252]	[0.091]
House has cement floor	0.019	0.002	0.001	−0.010
	(0.02)	(0.02)	(0.02)	(0.02)
	[0.53]	[0.61]	[0.593]	[0.324]
Asset count	−0.643	−0.455	−0.484	−0.374
	(0.62)	(0.38)	(0.42)	(0.35)
	[0.53]	[0.234]	[0.303]	[0.324]

Notes: Table presents results for the ITT impact of VSLA on predefined outcomes. With the exception of row 5 (total savings) all regressions are weighted least squares regressions. For total savings, weighted tobit regressions are reported, which is why we do not report results in columns (2)–(4). All columns report the estimates from a regression of the respective predefined outcome on a VSLA treatment assignment dummy and stratification fixed effects, accounting for sampling weights. Huber–White standard errors clustered at the village level are presented in round parentheses. *p < 0.1, **p < 0.05, ***p < 0.01. Multiple hypothesis corrected sharpened q-values (as in Anderson, 2008) in square brackets. Column 1 (simple means) reports estimates where no additional controls are included. Column 2 includes the lagged baseline covariate as dependent variable. Column 3 presents results from difference-in-differences specifications, and column 4 includes household-level fixed effects in the specification. Sample size as in Table 3.

an overview. Total membership increased during the 2009–11 period, and in the 2011 survey the levels of membership found were substantially higher in the treatment villages than in the control villages.

Table 4 shows that, at the time of the baseline survey in 2009, 6% of the population in both the control and the treatment villages reported that they were members of a VSLA or a similar savings group. Two years later, the figure was 20.7% for the control group and 45% for the treatment group.

This 23.6 percentage point difference in the take-up of the VSLA intervention between the treatment and control villages is significant at the 1% level, suggesting that the randomization was effective in inducing more treatment villagers to participate in VSLAs. However, compliance with the treatment is not perfect, and the randomization suffers from two-sided non-compliance (Gerber and Green, 2012). Not all households living in villages assigned to treatment participated, just as some households from the control villages found their way into VSLA groups. But as Fig. 3 shows, the timing between treatment and control villages differs. Until mid-2009 there was a general, but very small, increase in membership in the area regardless of the random assignment into treatment and control groups. In 2009, when the intervention commenced, membership took off in the treatment villages. In the control villages, however, membership seems to have followed the general pre-project trend until late 2010. This indicates that control group contamination happened relatively late. However, we cannot rule out that the estimated results presented below suffer from some degree of downward bias due to this contamination.

Lønborg and Rasmussen (2014) describe the profile of households who participate in the VSLAs studied here. They suggest that VSLAs do reach very poor households (our study area is one of the poorest areas in one of the poorest countries in the world and so most households in our study area are very poor), but that in the area studied here, the very poorest do not participate.

4.2. Impact on predefined outcomes

The four columns of Table 5 below present the estimated intention-to-treat effects of the VSLA intervention using, respectively, difference-in-means, difference-in-means with lagged value of the dependent variable, pooled difference-in-differences, and first-differenced regression methods on the predefined outcomes. We first discuss the results using the significance levels from individual outcome regressions as the starting point. We then present the interpretation based on the multiple hypothesis q-values.

Overall, we find that the introduction of VSLAs in the treatment villages had a significant impact on a number of outcome variables over the two-year intervention period. Food security improved, particularly in terms of the number of meals consumed the day before the interview. Although the number of income-generating activities decreased, the total volume of savings increased. Finally, there is some evidence of increased total expenditures as predicted using PAT as well as increases in the size of the dwelling, while we do not find any significant impact on the length of the hunger period, on food consumption (measured in terms of the 17 most common food items), on floor quality or a gross asset count.

The increase in total savings arises from an increase in the proportion of people who have liquid savings[24], but mostly from an increase in savings amounts among those who already had savings at baseline. The impact on savings is sizeable and corresponds to an increase in savings of 153% in addition to the increase in saving experienced in the control group. Appendix Table A2 shows that there is no evidence

[24] In raw numbers 71.4% of households hold liquid savings at baseline, at endline this has increased to 79.3 and 87.5% in control and treatment villages, respectively.

that savings went down in the control group. On the contrary, there is evidence that there was a positive and significant impact on savings among respondents on the list of interested households in control villages, an increase which is entirely driven by VSLA members. There is also no evidence of a decrease in savings among the non-VSLA members. As such, Appendix Table A2 provides no evidence to support worries that control villages households postponed savings while waiting for the VSLA program. Furthermore, when we look into membership into alternative organizations or groups with microfinance components, we find that there is very little crowding out. The introduction of VSLAs in treatment villages reduces membership in village-based groups (some of which may also have a savings and credit function) by 3–4 percentage points.[25]

The difference-in-means estimates of the impact on the number of income-generating activities are negative across the first two specifications, indicating that households in the treatment villages decreased the number of such activities. *A priori* it is unclear which impact to expect here. The implementing NGO had an aim of increasing the number of income-generating activities of the households involved in the program, but as mentioned in the introduction the fact that a household gains access to savings, credit and insurance against major unforeseen problems may lead to more efficient economic choices, in particular to specialize in a few more profitable income-generating activities rather than to have a diversified set of such activities. A decrease in the number of income-generating activities is consistent with such an explanation, although we cannot identify this to be the exact channel that accounts for the change.

In columns (3) and (4) we control for baseline differences between treatment and control groups by estimating difference-in-differences regressions taking village-level time-invariant characteristics into account as well as household-level time-invariant characteristics using first-differencing. The results across these two specifications are very stable. First of all, the number of meals consumed the day prior to the survey significantly increased: the households from the treatment villages consumed on average 0.13–0.15 more meals per day (s.e. = 0.05 and s.e. = 0.07 respectively), which is equivalent to one in seven households consuming an extra meal. Given that the baseline average is 2.65 meals per day, this is a sizable impact.[26]

The effect on the number of meals per day is mirrored in a positive effect on overall expenditure as predicted by PAT (marginally significant at the 10% level in the difference-in-differences specification, and at the 5% level for first differences), of about 4.1–4.2% (s.e. = 0.02).

The number of rooms increased by 0.14–0.16 (s.e. = 0.06 and s.e. = 0.07, respectively), which is statistically significant at the 10% level in the difference-in-differences specification, and at the 5% level for first differences. In rural areas in Malawi, where households usually live in huts built of mud plastered onto a wooden frame, it is quite common to build additional rooms onto the existing structure or in the compound when money allows, or – if possible – to replace the mud hut with burnt or un-burnt brick structures.[27] However, we found no significant effects on the type of floor in the dwelling. We also found no significant effects on total food consumption measured in terms of the 17 most common food items or on the length of the "hungry" period where households live on less than three meals per day, though the signs of both of these are consistent with an increase in food security.

[25] Results are available upon request from the authors.

[26] The impacts on number of meals do not seem to be driven by those households who share-out close to the date of the survey (Fig. 4). When we drop those households, the impacts on number of meals remain very similar. These results are available upon request from the authors.

[27] Appendix Fig. A1 shows that the changes in the number of rooms are primarily due to households in the treatment group adding rooms, as opposed to the housing stock in the control group depreciating.

4.3. Robustness

Our results are weaker in terms of statistical significance and not robust across specifications when we use the multiple hypotheses corrected standard errors. In the simple means specification, the impact on savings and on the number of income generating activities (remains) significant, though at lower levels of significance (5% and 10% levels respectively). In the difference-in-differences specification, not a single variable is significant, while in the first difference specification, our preferred specification for continuous outcomes, the impact on the number of meals consumed yesterday and the size of the house are significant, though only at the 10% level of significance. The impact on expenditure levels as predicted by PAT is not significant at conventional levels of significance, with a q-value of 0.11.

Our conclusion from the multiple hypothesis corrected tests of significance is that there is still evidence of VSLAs having a positive effect on food security, savings and the number of rooms. However, the evidence is much weakened in its statistical significance and in the lack of robustness across specifications.

In Table 6, we assess the robustness of the results to dropping split households, focusing on the estimated intention-to-treat effects. In the 2011 survey, we tracked any households that had moved or split into multiple households during the survey period. Thus, for a number of households, defined from the 2009 survey, we had two observations in the 2011 survey. In the predefined analysis above, we drop one of each of these split households at random from the 2011 sample. In Table 6 we drop all split households, i.e. restricting the sample to intact households. Columns (1) through (4) replicate the specifications of Table 5. The point estimates do not change much, and the significance level of all estimates is nearly identical to the previous specifications.

4.4. Results for intermediate outcomes

In order to investigate possible causal mechanisms for the effects found, we begin by studying whether the intervention actually improved access to and take-up of the financial services offered, namely savings and credit. We then investigate households' agricultural production as well as their business related activities.

4.5. Savings volumes and share-outs

The core component of the VSLA intervention is savings, these being a prerequisite for subsequent credit opportunities and insurance. We showed above the estimated ITT effects on total savings volume. In Table 7 we extend this analysis by looking at the effects on different types of highly liquid savings. We find that the change in overall savings is driven by changes in savings in VSLAs. We do not observe any significant changes in the other savings options (non-VSLA savings as a whole) nor in any of the sub-categories of non-VSLA savings.[28]

A natural question is where the funds saved in VSLAs come from, since savings in other categories of savings do not decline. The results for consumption suggest that the extra savings do also not come from reduced consumption in terms of meals per day or measured by PAT, our asset-based expenditure estimate.[29] The increased VSLA savings may be a result of extra income, just as credit may also play a role. The next sections investigate this further by estimating the impact on credit and on income from business and agriculture.

Before exploring the use of VSLAs, it is important to note that saving in the VSLA groups is not identical to saving in a regular liquid savings account: a core feature of the intervention is the annual share-out of all savings along with any interest earned from loans made during the

[28] We do not find evidence that VSLAs crowded out ROSCAs. Only 20 households had been a member of a ROSCA at baseline.

[29] Note that these consumption measures are unlikely to capture the more hidden consumption used for instance on alcohol or gambling.

Table 6
Regressions without split households.

Outcome	(1) Difference in means no split household	(2) Difference in means with lag and no split household	(3) Difference in difference with no split household	(4) First-differences without split households
Number of months with fewer than three meals a day	−0.204	−0.266	−0.478	−0.541
	(0.38)	(0.36)	(0.45)	(0.38)
Number of meals yesterday	0.058	0.071**	0.150**	0.129**
	(0.04)	(0.03)	(0.07)	(0.05)
17-Food consumption per week per adult equivalent (MK, log)	0.023	0.043	0.079	0.094
	(0.05)	(0.05)	(0.07)	(0.06)
Number of income-generating activities (including agriculture and livestock)	−0.185***	−0.151***	−0.113	−0.068
	(0.06)	(0.05)	(0.11)	(0.07)
Total savings (log)	1.043**	*n.a.*	*n.a.*	*n.a.*
	(0.43)			
Per capita expenditure predicted by USAID PAT (log)	0.030	0.035*	0.041*	0.040*
	(0.03)	(0.02)	(0.02)	(0.02)
Size of house (number of rooms)	0.102	0.133*	0.137*	0.159**
	(0.10)	(0.07)	(0.07)	(0.06)
House has cement floor	0.017	0.001	0.002	−0.010
	(0.02)	(0.02)	(0.02)	(0.02)
Asset count	−0.652	−0.475	−0.504	−0.414
	(0.63)	(0.39)	(0.43)	(0.36)

Notes: The table presents weighted regressions of predefined outcomes on a treatment dummy, following the same specifications as in Table 5. The sample excludes all households that split between baseline and endline. For total savings, weighted tobit regressions are reported, which is why we do not report difference-in-differences and fixed effects results. Huber–White standard errors clustered at the village level are presented in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7
ITT effects on savings outcomes.

Outcome	(1) Difference in means
Total savings (log)	1.034**
	(0.4304)
VSLA savings (log)	3.480***
	(0.902)
Non-VSLA savings (log)	0.048
	(0.479)
Savings with friend/relative (log)	−0.763
	(1.681)
Savings at home (log)	0.192
	(0.569)
Savings with bank (log)	0.910
	(1.761)

Notes: Table presents results for tobit regressions estimating the impact of VSLA on the log of the value of savings. We implement the suggestion of Cameron and Trivedi (2005) of setting the cutoff at a very small value below the lowest observed value for zero values. Columns 1 presents the results from a tobit regression on the VSLA dummy and block fixed effects. ** $p < 0.05$, *** $p < 0.01$.

cycle. In other words, savings in VSLAs are tied to the VSLA cash box until the date of the share-out, and thus they also become a commitment device. Among the households participating in VSLAs, more than 30% had received a share-out at the time of the survey in 2011.[30] Based on households' self-reports, the funds from the share-outs were primarily used for agricultural inputs or investments. This corresponds to the timing of the share-out observed, which was typically during the planting season when there is a need for seeds and fertilizer (see Fig. 4).

[30] From the information recorded by the implementing NGO, only 3 of the 102 groups that had been initiated by September 2011 had shared out twice, while another 40 groups had shared out once.

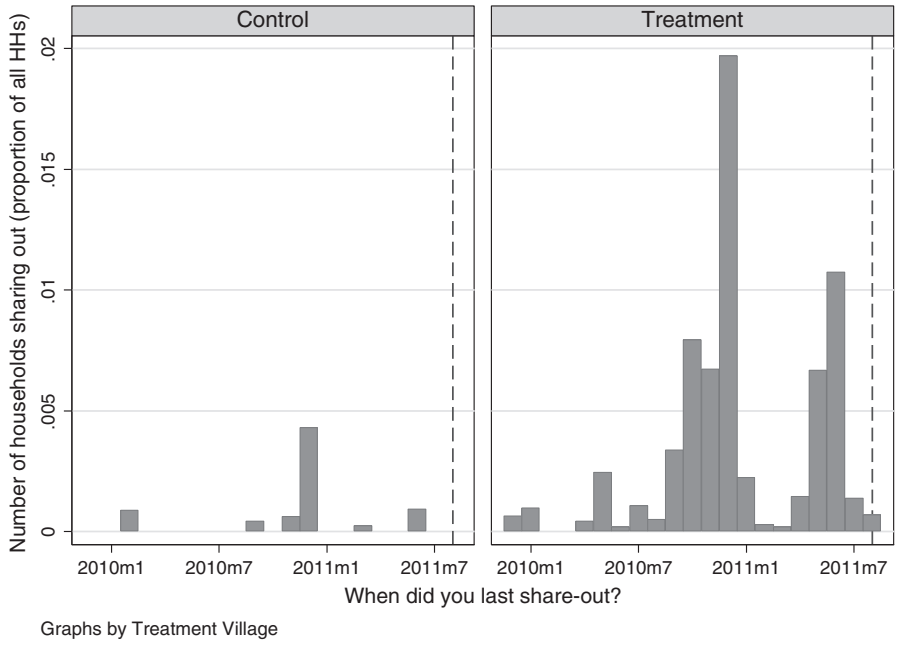
4.6. Credit volume

The other key component of the intervention is the use of pooled savings as credit for the VSLA members.

Table 8 below shows the estimated ITT effects on a range of credit-related outcomes using the same estimation strategies as in Table 5. In addition to increasing the likelihood of taking out at least one loan, living in a treatment village increased the number as well as the value of loans that had been active within the previous twelve months (total loan amount). The intervention also increased the take-up of loans for investment purposes: the probability of having taken a loan for investment purposes and the number of loans taken for investment purposes increased, just as the value of loans taken for investment in agriculture increased significantly. The increases are sizeable and generally significant. Total loan amounts increased significantly and so did the amount borrowed for agriculture; however there was no significant increase in the amount borrowed for business purposes. If we include all types of loan, the raw share of households with loans in treatment villages more than quadrupled from 6% to 26%,[31] suggesting that access to and use of credit increased substantially.

Although the total loan value for agricultural investments increased considerably more than that for business purposes, more VSLA members report to have taken out loans for trading and other business purposes than for agricultural input or investments. Self-reported use of credit might very well be wrong (Karlan et al., 2015) and thus, we investigate these two potential channels for the effects of VSLAs in the following. Specifically, we look at whether VSLAs seem to have influenced agricultural inputs or outputs and whether VSLAs have influenced other types of business activities.

[31] Note that this increase is larger than the ITT estimate by itself, as the proportion of control households with loans also doubled, possibly because of spill-over of the intervention into control villages. The proportion of treatment households with any loan at baseline was 5.9% (8.1% for control households). Between baseline and endline, the control villages increased the proportion of households with a loan by 6% and the treatment villages by 19.6%. In total at endline, 25.5% of treatment households had a loan. A little more than 50% of the VSLA members took out loans.



Notes: Dashed vertical line shows median date of endline survey.

Fig. 4. Timing of share-outs (by treatment status).

Table 8
ITT effects on credit outcomes.

Outcome	(1) Baseline mean	(2) Difference in means	(3) Difference in means with lag	(4) Difference-in-difference	(5) First-difference
Household had any loan in past 12 months	0.070	0.122***	0.122***	0.129**	0.145***
	[0.26]	(0.04)	(0.04)	(0.06)	(0.05)
Number of loans active within past 12 months	0.074	0.138***	0.138***	0.144**	0.160***
	[0.28]	(0.04)	(0.04)	(0.06)	(0.05)
Total loan amount (log)[a]	8.72	2.714***	*n.a.*	*n.a.*	*n.a.*
	[1.20]	(0.76)	*n.a.*	*n.a.*	*n.a.*
Household took out loan for investment purposes in past 12 months	0.044	0.096**	0.095**	0.076*	0.082*
	[0.21]	(0.04)	(0.04)	(0.04)	(0.04)
Number of investment loans	0.046	0.107***	0.107***	0.086*	0.092**
	[0.22]	(0.04)	(0.04)	(0.04)	(0.04)
Total amount borrowed for agricultural investments (log)[a]	9.38	3.507***	*n.a.*	*n.a.*	*n.a.*
	[0.85]	(1.01)	*n.a.*	*n.a.*	*n.a.*
Total amount borrowed for business purposes (log)[a]	8.53	1.183	*n.a.*	*n.a.*	*n.a.*
	[1.14]	(0.76)	*n.a.*	*n.a.*	*n.a.*

Notes: Table reports estimates of the impact of VSLA on credit outcomes. Column (1) presents the mean of the outcome at baseline, with standard deviations in square brackets. Column (2) presents the results from the difference in means specification, column (3) from the difference in means with lagged value specification. Columns (4) and (5) present results from difference-in-differences and first differences specifications, respectively. Rows without "a" superscript contain Least squares regressions. Regressions with "a" superscript are tobit regressions, following Cameron and Trivedi (2005) suggestion of replacing zero with the lowest observed value minus a small amount. For these regressions, column 1 presents the estimated mean for observations in the population with non-zero values, whereas the standard deviation represents the sample standard deviation (as some strata do not contain observations at baseline STATA is not able to compute population standard deviations). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.7. Agricultural activities

Since most share-outs are at the end of the year, when purchases of seed and fertilizer are made, Table 9 investigates whether inputs into and outputs of maize have changed. The first column shows the weighted baseline means, whereas columns 2 through 5 provide results for the difference in means, difference-in-means with lag, the pooled difference-in-differences and the first difference regressions, respectively.

On the input side, we find evidence that households in treatment villages were more likely to use fertilizer for growing maize. On the output side, we find a positive effect on the likelihood that households in treatment villages sold part of their maize production and that the total value of maize sales increased. We do, however, not find evidence that production increased, nor of an increase in maize output. This evidence does point to a change in agricultural practices. However, because this has not led to significantly higher yields, this is unlikely to be the primary channel for the observed increase in food security and expenditures.[32]

[32] Appendix Table A3 shows the ITT effects on a range of other agricultural inputs and outputs, though we do not find any consistent effect.

Table 9
ITT effects on maize planting practice and outcomes.

	(1)	(2)	(3)	(4)	(5)
Outcome	Baseline mean	Difference in means	Difference in means with lag	Difference-in-difference	First-difference
Household uses any fertilizer on maize	0.452	0.135**	0.127**	0.106*	0.093*
	[.498]	(0.05)	(0.05)	(0.06)	(0.05)
Area with maize (acres)	1.375	−0.306*	−0.241	−0.120	−0.154
	[.897]	(0.17)	(0.15)	(0.15)	(0.16)
Quantity of maize harvested (kg, log)[a]	5.759	−0.103	*n.a.*	*n.a.*	*n.a.*
	[.925]	(0.13)	*n.a.*	*n.a.*	*n.a.*
Household sold any maize	0.202	0.095*	0.116**	0.140*	0.198***
	[.402]	(0.05)	(0.03)	(0.07)	(0.06)
Value of agricultural sale (MK, log)[a]	9.375	−0.582	*n.a.*	*n.a.*	*n.a.*
	[1.55]	(0.56)	*n.a.*	*n.a.*	*n.a.*
Value of maize sale (MK, log)[a]	7.936	1.489**	*n.a.*	*n.a.*	*n.a.*
	[.958]	(0.73)	*n.a.*	*n.a.*	*n.a.*

Notes: Table reports estimates of the impact of VSLA on agricultural inputs. Column (1) presents the mean of the outcome at baseline, with standard deviations in square brackets. Column (2) presents the results from the difference in means specification, column (3) from the difference in means with lagged value specification. Columns (4) and (5) present results from difference-in-differences and first differences specifications, respectively. Rows without "a" superscript contain Least squares regressions. Regressions with "a" superscript are tobit regressions, following Cameron and Trivedi (2005) suggestion of replacing zero with the lowest observed value minus a small amount. For these regressions, column 1 presents the estimated mean for observations in the population with non-zero values. * $p < 0.01$, ** $p < 0.05$, *** $p < 0.01$.

4.8. Business activities

Table 10 reports the estimated ITT effects on income derived from business. This is defined as the money earned (or lost) in the last month of operation, i.e. it is a measure of profit. Based on OLS regressions, the first panel shows that there is some weak evidence that the average number of businesses in the treatment villages increased. Turning to the mean outcome, we see that business income increased by about 6590 MWK (about the size of the share-out) *on average* (s.e. = 8926). However, due to the very large variation in business income, the estimated standard errors are very large and the result is not statistically significant.

The results in the second panel are therefore based on quantile regressions, giving us results robust to outliers and allowing us to investigate whether there are heterogeneous effects at different place in the business income distribution. We find that median income from business has increased by between 1295 (s.e. = 758) and 1500 MWK (s.e. = 904) more for the median treatment household than for the median control household, an effect that is statistically significant at the 10% level. The treatment effect is higher than the median income from business at baseline. We then investigate whether there are impacts at the 25th and 75th percentiles and find that we cannot reject that there are no effects on average. We then investigate whether business incomes rose for the 66% of households who had already been

Table 10
Effects on small business outcomes.

Outcome		Baseline mean	Difference in means	Difference in means with lag
OLS regressions				
Number of businesses (excluding agriculture and livestock)		0.93	0.192	0.241*
		[0.778]	(0.129)	(0.125)
Total income from all businesses (MWK)		10523.2	6589.8	2453.4
		[49834.1]	(8926.315)	(5609.115)
Quantile regressions				
			600	386.5
	25th	0	(463.3)	(435.8)
			1,500*	1295*
	50th (median)	700	(904.4)	(758.3)
			1150	1652
Business income all respondents	75th	4000	(1747)	(1559)
			900	615.0
	25th	723	(730.2)	(694.2)
			2500*	2811**
	50th (median)	2500	(1330)	(1208)
			4600*	3911*
Business income respondents *with* business at baseline	75th	6100	(2407)	(2069)
			0	*n.a.*
	25th	0	(123.1)	
			0	*n.a.*
	50th (median)	0	(1108)	
			−2400*	*n.a.*
Business income respondents *without* business at baseline	75th	0	(1421)	

Notes: Table reports estimates of the impact of VSLA on income generating activities. Column 1 presents the mean of the outcome at baseline, with standard deviations in square brackets. Columns 3 and 4 present results from difference in means and difference in means with lag, respectively. OLS regressions follow specifications described in Table 5. Quantile regressions include stratification dummies, as well as a dummy for baseline business activity. Quantile standard errors are cluster-bootstrapped and include the randomization strata. Sample: long questionnaire respondents.* $p < 0.1$, ** $p < 0.05$.

involved in a small-scale business at baseline. Here we find more precisely estimated treatment effects, suggesting significant positive impacts on both the 50th and 75th percentiles. Conversely, for households who had not yet been engaged in some small-scale business, there is no evidence that their business income increased for the median household.

Overall, the evidence suggests that the program seems to have led to significant increases in investments in agriculture, although we cannot detect any impact on agricultural outcomes. On the other hand, the VSLAs were more successful in increasing the income from existing small-scale businesses.

4.9. Cost effectiveness

In order to assess whether VSLAs are worthwhile, this section provides some simple cost-effectiveness calculations. We estimate the cost per year of the implementation in the project studied as total costs (USD 201,000) divided by three. The specific intervention investigated was quite costly compared to the average VSLA intervention due to its relatively small scale and the fact that VSLAs often use village agents to a greater extent than the randomization allowed in the project setup.[33]

The implementing partner reported having a total number of 1783 VSLA members by September 2011 out of the 3800 households in the treatment villages. The cost per member was thus USD 75,[34] while the cost per household in a treatment village was USD 35 per year of implementation.

The estimated effect on total household expenditures of 4.2 per cent corresponds to an increase in the average household expenditures of 0.124 USD per household per day, which is equivalent to USD 0.31 at PPP adjusted priced (PPP 2005). On this basis the project broke even after 280 days or a bit over nine months of impact (based on the two-year impact estimates).[35] These estimates do not include impacts beyond the duration of the two years or impacts on other measures beyond the impact on household expenditures.

5. Conclusion

This paper provides evidence that improvements in local financial market intermediation can have a significant impact on household welfare and economic activities in remote rural areas of developing countries. Village Savings and Loan Associations have been heavily promoted in the hope of remedying what has been termed the "last mile problem of microfinance": operating a microfinance institution and (even more so a bank) is rarely profitable in remote rural areas in developing countries. VSLAs, on the other hand, do not rely on any outside injection of funds or professional management once set up. They harness savings in the local community which are re-invested locally, and are entirely self-managed. A set of carefully thought-out safeguards and governance features seem to be able to sustain the trust necessary for an improved local financial market intermediation. As household wealth increases and VSLAs mature, it is possible that these semi-formal institutions might even be integrated into the formal financial sector.

This paper provides evidence of the effects on household welfare of introducing VSLAs in rural Malawi. We find that this introduction raises welfare along a number of dimensions, in particular for improved food security and strengthened household income indicators. The results provide evidence that households in the poorest, most rural areas are constrained by lack of access to financial institutions providing savings and credit opportunities. Having analyzed the impact only two years after the onset of the project, we find these indications of the positive effects of simply organizing savings groups which also create credit opportunities for their members rather encouraging.

In studying how the money from share-outs and credit is used, we find that the use of fertilizer rises, and that loans are used primarily in agriculture and small-scale businesses. Agricultural output increases in response to the increased investments; however we cannot detect any impact on agricultural outcomes. On the other hand, we do see an increase in total business income among households, which already had a business at baseline, whereas there was no significant increase in business income among households with newly established businesses. Judging by these results, the provision of opportunities for the village to save and take out loans for agricultural and business investments may play a key role in travelling "the last mile".[36] The mechanism through which VSLAs might have impacts in different contexts might well differ, given the different roles participants can play in these savings groups. The mechanism we identify suggests the presence of financial market imperfections, whereas Beaman et al. (2014) suggest that women in their treatment villages become more forward looking.

The above conclusions have limitations, however, which should be kept in mind. Out of the 80,000 VSLAs across the world, we have studied only a few. The results may be specific to the context, cultural and economic, in which they appear, and the particular constraints faced by households in these communities. Future assessments from other places, including the ones we know are under way, will shed light on the extent to which these findings can be generalized.

Appendix A. Protocol of intervention of NGO in the villages

Both the villages in the treatment and control group received the same information, according to the field reports from the implementing agency. Two types of meetings were held, following the VSLA manual from 2009.

1. One orientation of community leaders and administration officials.
2. Introducing VSL to the community.

The first meeting included Senior Chiefs, Group Village Headmen, Village Headmen, religious leaders, various development committees currently present in the project area as well as government workers. The objective of this meeting was to obtain the support of key stakeholders. The agenda was the following:

1. Introduction of the implementing organisation.
2. Project goals and objectives.
3. Target group to be served.
4. Services offered.
5. Role of local leaders and administrators.
6. The stakeholders were told about similar programs in nearby regions (that had been successful).

The second set of meetings was carried out in each of the 46 villages. The goal of these meetings was to create general awareness in the community of the programme's purpose, methodology and process and to offer the opportunity to register for training. The purposes were described to be to enable villagers to manage their own savings and to administrate a loan facility through which they could earn a profit. It was emphasized that there would be no injection of funds into the saving groups other

[33] Normally, VSLAs are introduced to a single village. Some VSLA members then become village agents, who subsequently start VSLAs in neighboring villages. While this limits the costs of the intervention, we do not know what the effects of relying on village agents rather than trained field officers would be.

[34] This is higher than the typical cost of implementing VSLAs as reported by Allen and Panetta (2010), which was USD 18 to 48 per member.

[35] The estimates are very close, when instead we use the treatment effect on meals per day of 0.126 meals per day as a starting point.

[36] We do not know to what extent participation and savings in VSLAs are driven by the commitment aspect of the associations—the obligation to contribute a minimum amount each week. Duflo et al. (2011) suggest that in the Kenyan context, the inability to save for investments is due to behavioral biases.

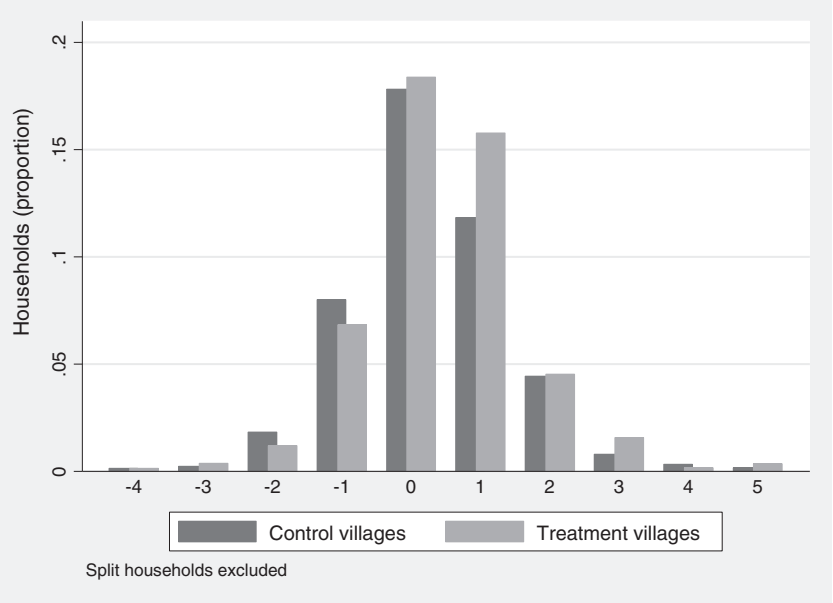


Fig. A1. Change in number of rooms over time

than a cash kit, which would be issued out as a loan. All villages were informed that 50 groups would be formed each year for three years and not all villages could receive training in the first year of implementation. Rather a lottery draw would be conducted to find which villages would be included in the project between 2009 and 2011 and which would be included after 2011. It was in these meetings that villagers could voice their interest and put their names on the initial list of interested households. Following meetings in all 46 villages, the randomization was carried out and an introductory meeting was held with groups in the 23 treatment villages.

Appendix B. Multiple hypothesis testing procedure

The usual p-values are underestimated when multiple hypotheses are tested, and need to be adjusted so that the rejection rate is correct under the null. To gain an intuition of how this works, take the case where you are testing two variables, X and Y with p-values of 0.045 and 0.04. The standard Bonferroni correction would say that since you are testing two variables, you have (almost) twice the probability of finding at least one error at the 5% level of significance, so you have to "adjust" the p-values by a factor of 2. In this example, such an adjustment would mean that the variables are not significant at the 5% level of significance, even though *both* are individually significant at that level.

An alternative adjustment procedure is suggested by Benjamini and Hochberg (1995). Their method proceeds as follows: it orders the p-values from the lowest to highest. Define p_i as the ith p-value in this ordering. Then analyse whether $p_i < i/m^*\alpha$, where i is the ordering parameter, m are the number of total hypotheses tested, and α the level of significance. If the inequality is met for any variable, then all variables with a lower p-value are considered significant for significance level α. In the example of X and Y the inequality is not met for the first variable, but is true for the second variable, so both variables are significant at the 5% level.

Benjamini et al. (2006), as explained in Anderson (2008), show how to improve power further. They implement a two-stage procedure, where the idea is as follows: the cut-off from the Benjamini and Hochberg method mentioned above (whether $p_i < i/m^*\alpha$) assumes that all null hypotheses are true. If say c of the first variables is significant at very, very high levels of significance, we should take that into account. This can be done as follows (see Anderson, 2008): Implement the Benjamini and Hochberg method with a significance level of $\alpha/(1+\alpha)$. Define c as the number of rejected hypotheses. Compute $m_0 = m - c$, which is an estimate of the number of incorrect hypotheses. Redo the Benjamini and Hochberg method with the following cutoff: $p_i < i/m_0 * \alpha/(1+\alpha)$. Anderson (2008) then shows how to go from rejection at specified levels of significance to the lowest level of significance, at which the hypotheses can be rejected, these adjusted p-values are denoted as sharpened q-values.

Table A1
VSLA interest and participation.

	(1)
Dep var: VSLA membership	Linear Probability Model
Panel 1	
HHs not on initial list of interested in control villages	0.113
Panel 2: Increase in participation over base category (HHs not on initial list of interested in control villages)	
On interested list in control villages	0.152***
	(0.0334)
Not on interested list in treatment villages	0.241***
	(0.0398)
On interested list in treatment villages	0.491***
	(0.0439)
Panel 3: Interaction between Interest and treatment (coefficient and significance)	
On interested list * treatment village	0.098**
	(0.0455)
Observations	1680
R-squared	0.107

Notes: Table reports estimates of probability of participation in VSLA using a Linear Probability Model. Panel 1 presents the proportion of households not on the initial list of interested households in the control villages who participate in VSLAs. This is the base category. Panel 2 presents estimates of the increase in the participation probability for the three categories: On interested list in control villages, not on interested list in treatment villages, and on interested list in treatment villages. The estimated coefficients are also from a Linear Probability Model that includes block fixed effects. Panel 3 presents the estimate of whether the effect of being both interested and in a treatment village (0.491) at the same time is different from the sum of the coefficients of being interested (0.152) and being treated (0.241). Robust standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Table A2
Time trend in savings in control villages.

Dep var: Log(savings + 1)	(1) Entire control group	(2) Control villages: on list of interested HH	(3) Control villages: not on list of interested HH	(4) Control group VSLA members	(5) Control group non-VSLA members
Endline	0.795 (0.798)	1.317*** (0.409)	0.687 (0.945)	2.906*** (0.577)	-0.038 (0.942)
Observations	811	509	302	267	544
R-squared	0.018	0.045	0.018	0.234	0.016

Notes: Table reports estimates of a regression of savings on a dummy for endline, as well as block fixed effects. Column 1 presents results for the entire sample of control group members. Column 2 shows the estimate from a regression of savings on a dummy for endline for HHs on the interested list. Column 3 present equivalent results for HHs *not* on the interested list. Column 4 presents results for respondents who in either of the two survey rounds were members of a VSLA. Column 5 presents results for respondents who were never VSLA members. The dependent variable is different than in Table 5, as implementing a tobit regression in the specification of Table 5 results in an error, due to one strata only containing a single village in the control group. Robust standard errors in parentheses. ***p < 0.01, **p < 0.05, *p < 0.1.

Table A3
Effects on other agricultural outputs.

Outcome	(1) Baseline mean	(3) Difference in means	(2) Difference in means with lag	(4) Difference-in-difference	(5) First-difference
Household uses any fertilizer	0.702 [0.458]	0.088 (0.06)	0.092 (0.05)	0.108 (0.08)	0.1 (0.07)
Household cultivates vegetable plot	0.247 [0.431]	−0.047 (0.03)	−0.023 (0.03)	0.037 (0.04)	0.037 (0.03)
Household has any irrigated plots	0.28 [0.449]	0.01 (0.05)	0.047 (0.05)	0.125 (0.08)	0.166** (0.07)
Household used any purchased seeds	0.803 [0.398]	0.011 (0.05)	0.017 (0.05)	0.054 (0.07)	0.073 (0.08)
Total area cultivated (acres)	2.876 [1.716]	−0.251 (0.19)	−0.119 (0.16)	−0.013 (0.21)	0.01 (0.21)
Area with local maize (acres)	0.53 [0.784]	−0.309** (0.15)	−0.286* (0.15)	−0.183 (0.12)	−0.227* (0.13)
Area with composite maize (acres)	0.079 [0.36]	−0.009 (0.01)	−0.009 (0.01)	−0.022 (0.03)	−0.02 (0.02)
Area with hybrid maize (acres)	0.766 [0.86]	0.012 (0.08)	0.034 (0.08)	0.084 (0.08)	0.094 (0.07)
Area with tobacco (acres)	0.184 [0.439]	0.023 (0.06)	0.043 (0.05)	0.06 (0.06)	0.071 (0.06)
Area with cotton (acres)	0.315 [0.64]	−0.022 (0.06)	−0.021 (0.04)	0.002 (0.07)	−0.022 (0.06)
Area with rice (acres)	0.316 [0.593]	−0.18 (0.13)	−0.098 (0.09)	−0.038 (0.11)	−0.015 (0.11)
Local maize harvested (kg, log)[a]	5.397 [0.913]	−0.920* (0.53)	*n.a.* *n.a.*	*n.a.* *n.a.*	*n.a.* *n.a.*
Composite maize harvested (kg, log)[a]	5.911 [1.008]	0.292 (0.61)	*n.a.* *n.a.*	*n.a.* *n.a.*	*n.a.* *n.a.*
Hybrid maize harvested (kg, log)[a]	5.668 [0.957]	0.33 (0.28)	*n.a.* *n.a.*	*n.a.* *n.a.*	*n.a.* *n.a.*
Local maize harvested per acre (kg, log)[a]	5.35 [0.8]	−0.816 (0.50)	*n.a.* *n.a.*	*n.a.* *n.a.*	*n.a.* *n.a.*
Composite maize harvested per acre (kg, log)[a]	5.741 [0.897]	0.417 (0.72)	*n.a.* *n.a.*	*n.a.* *n.a.*	*n.a.* *n.a.*
Hybrid maize harvested per acre (kg, log)[a]	5.639 [0.837]	0.374 (0.28)	*n.a.* *n.a.*	*n.a.* *n.a.*	*n.a.* *n.a.*
Household sold any crops	0.592 [0.492]	−0.041 (0.06)	−0.021 (0.06)	−0.016 (0.08)	0.027 (0.07)

Notes: Table reports estimates of the impact of VSLA on agricultural inputs. Column (1) presents the mean of the outcome at baseline, with standard deviations in square brackets. Column (2) presents the results from the difference in means specification, column (3) from the difference in means with lagged value specification. Columns (4) and (5) present results from difference-in-differences and first difference specifications, respectively. Rows without "a" superscript contain Least squares regressions. Regressions with "a" superscript are tobit regressions, following Cameron and Trivedi (2005) suggestion of replacing zero with the lowest observed value minus a small amount. For these regressions, column 1 presents the estimated mean for observations in the population with non-zero values. *p < 0.10, **p < 0.05, ***p < 0.01.

Appendix C. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jdeveco.2015.12.003.

References

Ai, C., Norton, E.C., 2003. Interaction terms in logit and probit models. Econ. Lett. 80 (1), 123–129. http://dx.doi.org/10.1016/S0165-1765(03)00032-6.

Allen, H., Panetta, D., 2010. Savings Groups: What Are They? The SEEP Network, Washington DC

Allen, H., Staehle, M., 2007. Programme Guide — Field Operations Manual Version 3.1. VSL Associates, Stuttgarten (available from http://vsla.net [Accessed October 2012]).

Anderson, M.L., 2008. Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. J. Am. Stat. Assoc. 103 (484), 1481–1495. http://dx.doi.org/10.1198/016214508000000841.

Anderson, S., Baland, J.-M., 2002. The economics of ROSCAs and intrahousehold resource allocation. Q. J. Econ. 117, 963–995.

Angrist, J.D., Pischke, J.-S., 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, Princeton, NJ.

Annan, J., Bundervoet, T., Seban, J., Costigan, J., 2013. A Randomized Impact Evaluation of Village Savings and Loans Associations and Family-based Interventions in Burundi. USAID (Retrieved from http://www.rescue.org/sites/default/files/resource-file/New_Generation_Final_Report_05312013.pdf).

Ashe, J., 2002. A symposium on savings-led microfinance and the rural poor. J. Microfinance 4, 127–135.

Ashraf, N., Karlan, D., Yin, W., 2006. Tying Odysseus to the mast: evidence from a commitment savings product in the Philippines. Q. J. Econ. 121, 635–672.

Banerjee, A.V., 2013. Microcredit under the microscope: what have we learned in the past two decades, and what do we need to know? Ann. Rev. Econ. 5 (1), 487–519. http://dx.doi.org/10.1146/annurev-economics-082912-110220.

Banerjee, A.V., Duflo, E., 2009. The experimental approach to development economics. Ann. Rev. Econ. 1, 151–178.

Banerjee, A., Duflo, E., Glennerster, R., Kinnan, C., 2015a. The miracle of microfinance? Evidence from a randomized evaluation. Am. Econ. J. Appl. Econ. 7 (1), 22–53. http://dx.doi.org/10.1257/app.20130533.

Banerjee, A., Duflo, E., Goldberg, N., Karlan, D., Osei, R., Parienté, W., Udry, C., 2015b. A multifaceted program causes lasting progress for the very poor: evidence from six countries. Science 348 (6236), 1260799. http://dx.doi.org/10.1126/science.1260799.

Banerjee, A., Karlan, D., Zinman, J., 2015c. Six randomized evaluations of microcredit: introduction and further steps. Am. Econ. J. Appl. Econ. 7 (1), 1–21. http://dx.doi.org/10.1257/app.20140287.

Beaman, L., Karlan, D., Thuysbaert, B., 2014. Saving for a (not so) rainy day: a randomized evaluation of savings groups in Mali. Working Paper No. 20600. National Bureau of Economic Research (Retrieved from http://www.nber.org/papers/w20600).

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B (Methodol.) 57 (1), 289–300.

Benjamini, Y., Krieger, A.M., Yekutieli, D., 2006. Adaptive linear step-up procedures that control the false discovery rate. Biometrika 93 (3), 491–507. http://dx.doi.org/10.1093/biomet/93.3.491.

Besley, T., Coate, S., Loury, G., 1994. Rotating savings and credit associations, credit markets and efficiency. Rev. Econ. Stud. 61 (4), 701–719. http://dx.doi.org/10.2307/2297915.

Bouman, F.J.A., 1995. Rotating and accumulating savings and credit associations: a development perspective. World Dev. 23, 371–384.

Bruhn, M., McKenzie, D., 2009. In pursuit of balance: randomization in practice in development field experiments. Am. Econ. J. Appl. Econ. 1, 200–232.

Cameron, A.C., Trivedi, P.K., 2005. Microeconometrics: Methods and Applications. Cambridge University Press.

Carter, M.R., Zimmerman, F.J., 2003. Asset smoothing, consumption smoothing and the reproduction of inequality under risk and subsistence constraints. J. Dev. Econ. 71, 233–260.

Casey, K., Glennerster, R., Miguel, E., 2012. Reshaping institutions: evidence on aid impacts using a preanalysis plan. Q. J. Econ. 127, 1755–1812.

Collins, D., Morduch, J., Rutherford, S., Ruthven, O., 2009. Portfolios of the Poor: How the World's Poor Live on $2 a Day. Princeton University Press.

Conning, J., Udry, C., 2007. Rural financial markets in developing countries. In: Pingali, P., Evenson, R. (Eds.), Handbook of Agricultural Economics vol. 3. Elsevier, The Netherlands, pp. 2857–2908.

Dale, R., 2003. The logical framework: an easy escape, a straitjacket, or a useful planning tool? Dev. Pract. 13, 57–70.

Daley-Harris, S., 2009. State of the Microcredit Summit Campaign Report 2009. Microcredit Summit Campaign, Washington DC.
Deaton, A., Dupriez, O., 2011. Purchasing power parity exchange rates for the global poor. Am. Econ. J. Appl. Econ. 3, 137–166.
Demirgüç-Kunt, A., Klapper, L., 2012. Measuring financial inclusion: the Global Findex Database. World Bank Policy Research Working Paper No. 6025. World Bank, Washington, D.C. (Available at SSRN 2043012).
Dercon, S., 1996. Risk, crop choice, and savings: evidence from Tanzania. Econ. Dev. Cult. Chang. 44 (3), 485–513.
Diamond, D.W., 1984. Financial intermediation and delegated monitoring. Rev. Econ. Stud. 51, 393–414.
Duflo, E., Glennerster, R., Kremer, M., 2007. Using randomization in development economics research: a toolkit. In: Schultz, T.P., Strauss, J.A. (Eds.), Handbook of development economics vol. 4. Elsevier, pp. 3895–3962.
Duflo, E., Kremer, M., Robinson, J., 2011. Nudging farmers to use fertilizer: theory and experimental evidence from Kenya. Am. Econ. Rev. 101, 2350–2390.
Fulford, S.L., 2013. The effects of financial development in the short and long run: theory and evidence from India. J. Dev. Econ. 104, 56–72. http://dx.doi.org/10.1016/j.jdeveco.2013.04.005.
Gerber, A.S., Green, D.P., 2012. Field Experiments : Design, Analysis, and Interpretation. 1 ed. W. W. Norton, New York.
Glewwe, P., Jacoby, H.G., 2000. Recommendations for collecting panel data. In: Grosh, M.E., Glewwe, P. (Eds.), Designing Household Survey Questionnaires for Developing Countries : Lessons From 15 years of the Living Standards Measurement Study vol. 2. World Bank, Washington DC, pp. 275–314.
Greaney, B., Kaboski, J.P., Leemput, E.V., 2013. Can self-help groups really be. Working Paper No. 18970. National Bureau of Economic Research (Retrieved from http://www.nber.org/papers/w18970.
Honoré, B.E., 1992. Trimmed lad and least squares estimation of truncated and censored regression models with fixed effects. Econometrica 60 (3), 533–565. http://dx.doi.org/10.2307/2951583.
Jackson, M.O., Yariv, L., 2014. Present bias and collective dynamic choice in the lab. Am. Econ. Rev. 104 (12), 4184–4204.
Karlan, D., Osman, A., Zinman, J., 2015. Follow the money: methods for identifying consumption and investment responses to a liquidity shock. J. Dev. Econ. http://dx.doi.org/10.1016/j.jdeveco.2015.10.009.
Klonner, S., 2008. Private information and altruism in bidding ROSCAs. Econ. J. 118 (528), 775–800. http://dx.doi.org/10.1111/j.1468-0297.2008.02142.x.
Lønborg, J.H., Rasmussen, O.D., 2014. Can microfinance reach the poorest: evidence from a community-managed microfinance intervention. World Dev. 64, 460–472. http://dx.doi.org/10.1016/j.worlddev.2014.06.021.
O'Donoghue, T., Rabin, M., 1999. Doing it now or later. Am. Econ. Rev. 89, 103–124.
Puhani, P.A., 2012. The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. Econ. Lett. 115 (1), 85–87. http://dx.doi.org/10.1016/j.econlet.2011.11.025.
Rasmussen, O.D., 2012. Small groups, large profits: calculating interest rates in community-managed microfinance. Enterp. Dev. Microfinance 23, 298–318.
Rubin, D.B., 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. J. Educ. Psychol. 66 (5), 688–701. http://dx.doi.org/10.1037/h0037350.
Rutherford, S., 2001. The Poor and Their Money. Oxford University Press, Oxford.
Thaler, R.H., 1990. Anomalies: saving, fungibility, and mental accounts. J. Econ. Perspect. 4, 193–205.
USAID, 2009. Poverty Assessment Tool Accuracy Submission, USAID/IRIS Tool for Malawi. USAID (Retrieved from http://www.povertytools.org/countries/Malawi/USAID_PAT_Malawi.pdf).