Pre-Analysis Plan:
Relaxing Borrowing Constraints in Savings Groups: Evidence from Uganda

Alfredo Burlando, Luciana Etcheverry, and Jessica Goldberg *

April 30, 2019

Contents

---

*Email: burlando@uoregon.edu, lucianae@uoregon.edu, goldberg@umd.edu

1 Introduction

1.1 Abstract

The study randomly assigns VSLAs to two treatments: one providing a group-level savings account from a local bank, and the second providing the same savings account plus the possibility of getting a group level loan from the bank. The study is intended to provide groups with access to a safe place to store money, in addition to relaxing possible constraints that limit the ability of groups to lend to members. At the time of writing of this PAP, researchers had implemented the intervention and carried out one round of data collection (midline). An endline will be concluded in 2019. The data have yet to be analyzed. *Edits to this pre-analysis plan were incorporated – and are indicated in italics – in April 2019, after the PIs reviewed the midline data and before viewing endline data.*

1.2 Motivation

Savings groups (SGs), including VSLAs, generally self-fund all their activities, including the provision of loans to one another through the accumulation of group savings. Successful groups are capable of meeting all internal capital needs, allowing members to satisfy the credit needs of their families and their income generating activities. In practice, however, many VSLAs are formed by rural poor with limited access to capital, and they often face significant financial barriers, with members unable to accumulate enough savings to meet the financial needs of their members (Burlando et al, 2018). In addition, groups also face the possibility of theft from their savings, which itself could create distortionary behavior; for instance, it could induce people to reduce savings, or to over-lend at the end of the cycle to reduce the amount of cash in their savings box. In this context, linkage programs with external financial providers that provide loans to VSLAs could alleviate these financial barriers.

In this study, we propose to carry out an evaluation of a randomized control trial implemented in 2015-2016 to measure the impact of introducing formal banking products to members of savings groups in Uganda. The main question we are interested in addressing is whether savings groups participants benefit from an enhanced access to external (bank) credit. The study partnered a research team with a financial services provider (FSP), Opportunity Bank Uganda (OBUL), two local implementing partners with ties with savings group, READ Uganda and Project SCORE, and the Strmme Foundation. OBUL developed two financial products specifically targeted to savings groups: a group saving account, and a group loan. In our study, we randomly

assigned savings groups to three interventions: a loan intervention, in which groups receive access to loans and group savings; a savings intervention, in which groups receive access to group savings only; and a control. This PAP covers the analysis of midline data; this analysis will inform our plans for an endline, and we plan to create a new PAP that covers the endline.

Access to external credit also affects incentives to join and remain in savings groups. The groups may become more attractive to households seeking larger loans, or less attractive to more risk averse members. We will monitor the composition of groups across cycles.

The inclusion of a savings intervention serves two purposes. First, it allows us to understand the importance of providing savings groups with a safe place to store excess funds. If concerns over the safety of savings leads to over-lending at the end of each cycle (which is something some group members have hinted at to us in the past), or if money saved in the bank is not easily accessible to members, we expect bank lending to be lower in this treatment. Second, the presence of the savings intervention allows us to cleanly separate the effect of credit from the effect of bank savings. That is because OBUL loan products require groups to maintain a savings account, which makes it impossible to evaluate a lending program without a savings component.

The group saving account allows savings groups to keep part or all of their balance deposited in a mobile-money linked OBUL saving account. The loan product involves a financial transfer to the group saving account. Loans carry a 2.5% monthly interest rate, and repayments occur monthly according to a payment schedule. Repayment is initially set at 3 months. The bank also charges a "search fee" and application fee to each loan; these fees are rolled into the payment schedule. Finally, a one-time "financial card application fee is applied to groups" for the purpose of obtaining Bank of Uganda permission to receive credit.

1.3 Research Questions

- **Financial outcomes**

  - *Effects of linkage on takeup:* What is the likelihood of the group adopting a savings account? The group getting a loan from the bank? The amount of money the group saved in the bank?

  - *Effects of linkage on borrowing:* What is the effect of access to loans on the number and size of household loans to members? On net, we expect access to bank loans to increase the overall supply of loans to members.

- *Effects of linkage on saving:* What is the effect of access to loans on the amount saved by members? Savings might increase for those hoping to gain access to larger loans, which are now available thanks to the bank linkage. Savings may decrease for those who typically do not borrow, as the returns to savings may fall.

- *Effects of savings intervention:* what is the effect of bank savings accounts on group savings and the number and size of household loans to members? We do not have a clear hypothesis for this outcome. Bank savings could increase the level of trust in the security of the group, increasing overall savings and credit. On the other hand, funds in a bank account are not readily accessible to the group, making on-lending more difficult and potentially reducing the pressure or opportunity to disperse funds ahead of schedule.

- *Welfare impacts:* What are the effects of the loan and savings intervention on investments in agriculture or micro enterprise? Other household investments? Profits?

- **Effects on group composition**

  - *Group cohesion and trust:* what are the effects of the intervention on trust in leadership and overall group cohesion?

  - *Group membership turnover:* What are the effects of the intervention on entry to and exit from the group?

2 Research Components

2.1 Timeline

The intervention took place in several steps.

- **Sampling of VSLA** In the fall of 2014, a research team representative visited approximately 200 VSLAs in three districts served by READ Uganda. Baseline information on group development and performance was collected to assess eligibility into the study and implement randomization. In 2015, VSLAs supported by project SCORE in two additional districts was added to the sampling frame. In total, 156 groups were selected for the study.

- **Baseline data collection** In February-April 2015, baseline interviews were carried out in all study groups. For each group, 15 respondents were selected for the baseline.

- **Intervention** The interventions were slated for early 2015. It gained full steam in January 2016, and concluded in December 2016.

- **Midline** The midline took place in February-April of 2018. The sample was composed of respondents who participated in the 2015 baseline.

- **Endline** Scheduled to take place in February-April of 2019. The sampling frame includes all current participants of the VSLA, and a random sample of baseline participants who have dropped out of the study groups.

2.2 Instruments

Data for the study comes from (1) screening questionnaires for study groups at baseline, (2) three rounds of household surveys, (3) one round of endline SG leader interviews, (4) administrative records from our partner commercial bank and other group lenders in the area, (5) administrative records from each savings group, and (6) internal records of visits from the implementing team.

- Screening scores of groups: collected in the fall of 2014, this is a questionnaire developed by CARE to identify groups that could benefit from a formal bank linkage. The questionnaire includes a scoring mechanism to determine group eligibility in the study. It was also used to determine treatment arm balance.

- Household surveys: includes a baseline (carried out in 2015), a midline (carried out in 2018) and endline (to be carried out in 2019). Baseline and midline are completed. Data from the surveys includes primary outcomes of interest: self reported amount of savings and loans, participation status with the group, satisfaction with the group; household assets, earnings, investments.

  - An error in a loop in the CAPI interface in the midline survey means that while detailed data about the amount borrowed from and saved with various institutions (savings groups, SACCOs, formal banks, etc.) were collected as intended, these data were not captured for respondents' borrowing from and saving with the study savings group. Therefore, we do not have midline data for the most obvious outcomes of interest: the amount borrowed from or saved with the study savings groups. Instead,

we'll rely on information collected directly from the savings groups records to fill in this gap.

- SG leader interviews: this will be carried in 2019. We will use this survey to obtain detailed information about group characteristics such as longevity, history of group linkages with financial providers and other businesses, group rules and regulations. These variables will be used for additional first stage analysis (determinants of takeup as a function of group characteristics).

- Administrative records from financial service providers: these records have already been collected. Takeup of group savings and loans, including loan amounts, repayment history, and dates of account usages.

- Administrative records of VSLAs: Collection of all cashbooks from 2014 to the end of the study. Cashbooks identify, for each weekly meeting, income from savings, repayments and fines, and outlays from lending. Identify groups that are cash constrained for takeup analysis. Identify treatment effects for savings, lending, and repayment.

- Internal records of visits: Data included frequency of visits to study groups by implementing team; and some first stage outcomes such as whether the group requested a savings account, whether they received it, whether they requested a loan and whether they received it.

3 Empirical Analysis

3.1 Variables

- **Takeup** Takeup is measured at the VSLA level (since the outcome is a group level outcome: a savings account or loan for the group).

  - **Primary take up measure**: whether the VSLA opened a financial account with any financial institution. We call this variable "linked". This will include both savings accounts and loan accounts. We will use this variable to estimate TOT.

  The following are other variables associated with takeup that we will study.

  - Takeup of Opportunity Bank:

6

* *Savings account takeup-Opportunity bank* indicator for whether the VSLA opened a group savings account with Opportunity bank.
* *Loan requested-Opportunity Bank* Indicator for whether the VSLA requested a loan through Opportunity Bank.
* *Loan obtained-Opportunity Bank* Indicator for whether the VSLA obtained a loan from Opportunity Bank.
* *Number of loans-Opportunity Bank* A variable indicating the number of loans obtained from Opportunity Bank.
* *Value of first loan-Opportunity Bank*
* *Total value of loans-Opportunity Bank*

– Takeup for any formal bank (including Opportunity Bank and any other commercial bank or MFI)

* *Savings account takeup-Any bank* indicator for whether the VSLA opened a group savings account with any formal bank.
* *Loan requested-Any Bank* Indicator for whether the VSLA requested a loan through any bank.
* *Loan obtained-Any Bank* Indicator for whether the VSLA obtained a loan from any Bank. *Number of loans-Any Bank* A variable indicating the number of loans obtained from any bank.
* *Value of first loan-Any Bank* Amount obtained (in shillings)from first loan ever received from any bank
* *Total value of loans-Any Bank* Total value (in shillings) of all loans obtained from any bank(s).

• **Primary outcome variables of interest: VSLA level variables**

– *Savings deposit*: total amount (in shillings) of savings collected from members each week.

– *Loan disbursements:* total amount (in shillings) of loans provided to household members each week.

– *Shareout amount:* Total value of money available for shareout.

– *April 2019 edit: VSLA turnover: share of baseline members who were still members at endline; share of endline members who were members at baseline.*

– *April 2019 edit: Loan repayment: percent of total group loans that were not repaid by their due dates.*

- **Primary outcome variables of interest: household level survey outcomes**

  - *Outcome family 1: financial inclusion.* Three primary measures and an index of financial inclusion, constructed with the same three measures:

    * Total savings, from all sources
    * Total formal borrowing, from all sources: number of loans and total value of loans acquired from formal credit sources
    * Credit rationing: an indicator that is equal to 1 if the household requested a loan from some source and was denied; and the value of loan not given (which is zero if hhld did not have rejected loans.) *April 2019 edit: these measures will be moved to secondary analysis at endline if there is insufficient variation at midline.*

    Additional variables under this family:

    * *savings* total amount saved by member in (1) the study VSLA,[1] (2) Other VSLAs (3) formal banks and MFI (4) total savings (all formal and informal sources included in the questionnaire).
    * *borrowing* total amounts borrowed in (1) the study VSLA,[2] (2) other VSLAs (3) formal banks and MFIs (4) total loans, (5) indicator for borrowing from study VSLA (6) value of informal borrowing (7) percent of all borrowing that was from formal sources.

  - *Outcome family 2: household investments* Index of total amounts spent in investments in housing, agriculture, and enterprise. *April 2019 edit: at endline, look at asset stocks in each category, not flows, because of long interval from treatment.*

    Additional variables under this family:

    * Whether household made housing investments, and amount spent on housing investments
    * Whether household bought agricultural inputs, and amount spent on agricultural inputs (paid labor, materials),
    * size of land used for farming (in acres).
    * Amount spent on equipment
    * Amount spent on enterprise (sum total over 10 categories)

---

[1]Not available at midline survey; see above.
[2]Not available at midline survey; see above.

– *Outcome family 3: Household income* Income variable constructed from survey asking household all sources of income. For each, earning for the typical week scaled by the total number of months household is employed. *April 2019 edit: This text is unclear. Our intention is to calculate annual income by adding the products of monthly income from each activity times the number of months worked in that activity.*

- **Secondary outcome variables of interest: household level variables** In addition to the primary variables of interest, we are also interest in a number of variables which either explain indirect effects of financial linkages, or shed light on the transmission mechanism between the interventions and the outcomes of interest. These outcomes are excluded from multiple hypothesis testing and should be considered exploratory.

    – Trust in VSLA: Index of opinions about trust in the VSLA. All variables included in the questionnaire will be used to construct the index. *April 2019 edit: see notes below. This variable will be analyzed at the individual-group level and then aggregated for analysis at the individual level. The variables will be used to construct three separate indices: trust in other members, agreement/disagreement in group, and overall satisfaction with the group. Note that the questions at midline and endline are slightly different than those at baseline, and omit some baseline items. Field teams reported that respondents found the baseline questions confusing and frustrating, so revisions were made at midline and endline. This means that the baseline controls included for these variables will be measured differently than the outcome variables.*

    – Index of household assets (constructed using principal component analysis)

    – livestock index (using principal component analysis)

    – Food security index, constructed from 5 variables

    – Dropped out of the study VSLA. (indicator) *April 2019 edit: see notes below. This variable will be analyzed at the individual-group level and then aggregated for analysis at the individual level.*

Because of the potential high variance in savings across groups, we will also measure some of the financial variables of interest (savings, loans) using logs, which better controls for dispersion in the outcome variable. *April 2019 edit:*

we will use inverse hyperbolic sine transformation instead of logs, because the outcomes of interest are zero for some observations.

3.2 Treatment Effects

We will employ two methods for data analysis of the midline. An ITT and TOT methodology will be used on the full sample of VSLA participants at endline. We will also take advantage of the panel feature of our midline data to use an ANCOVA methodology with the baseline. Finally, we will use the baseline information to carry out exploratory ITT analysis, develop hypothesis for the second endline, and then resubmit a PAP centering on hypotheses that appear to be promising (Anderson and Magruder 2017).

3.2.1 Intent to Treat

ITT regression on primary outcomes and exploratory outcomes for individuals i in group g include:

$$y_{ig} = \alpha_0 + \alpha_1 T_{1g} + \alpha_2 T_{2g} + X_{ig}\beta + \epsilon_{ig}$$

Here T_1 is the indicator for the savings only intervention, T_2 is the intervention for the savings + loan intervention, and X is a set of group controls and time invariant household controls. Group controls will include all the variables used for randomization (from screening scores). Household controls include sex, age, education level of household head, employment sector indicators. In addition, the vector of controls will include location controls (district f.e., and parish f.e.). The regressions will be estimated with and without controls.

April 2019 edit: the notation above indicates clustering at the group level. This may have been appropriate if each individual was a member of exactly one study group. However, many individuals in the sample are members of multiple groups – as a portfolio diversification strategy or to increase the total volume of their saving and potential borrowing. Therefore, we will cluster at the village level, as the vast majority of individuals with multiple group memberships were members of groups within a single village. The village was also the unit of randomization, to account for potential spillovers across groups.

For outcomes in the financial inclusion, household investment, and household income variable families described above, plus the secondary measures of household assets and food security, the correct unit of analysis is the individual. For other outcomes, including trust in the VSLA and dropout (or other measures of VSLA churn), we will consider several alternatives: regressions at the individual-group level, with

individual fixed effects and frequency weights to account for cross sectional variation in the number of group memberships at baseline, and regressions at the individual level, aggregating outcome variables. For trust in the VSLA, we will consider the average level of trust across all of the individual's groups, and for drop out (and other churn variables), we will consider three different measures: dropped out of at least one group; dropped out of all groups; and the share of groups dropped.

We will also consider a richer set of individual level covariates than originally described. We will use two approaches. First, we will control for common observables that include the original list of covariates and other baseline measures: housing quality, land owned, asset index, livestock index. Second, we will use LASSO to choose covariates from those collected in the full baseline questionnaire.

We will estimate this model for five subsamples in order to explore mechanisms and heterogenity. These subsamples are :

- All baseline survey respondents (representative of group membership in 2015, and re-interviewed at endline even if they have left the group; this is the primary sample and is not affected by endogenous group formation in response to the intervention)

- All endline group members respondents (representative of group membership in 2018)

- Original members who remained in the group

- Original members who left the group

- New members who were interviewed at endline but were not members at baseline

April 2019: we will also estimate the ITT and TOT models limiting the sample to study groups facilitated by READ, and considering each of the subsamples among the READ-only groups. Facilitation for SCORE groups ended during our study period, and take up of treatments among SCORE groups was therefore very low.

The second ITT model is as follows:

$$y_{ig} = \alpha_0 + \alpha_1 Banked_g + X_{ig}\beta + \epsilon_{ig}$$

Here, the indicator *Banked* identifies groups assigned to any of the treatments. This indicator will capture the effects of exposure to any formal banking (savings or credit). Controls are as above. *April 2019 edit: the notes on clustering and unit of analysis in the previous section apply to this analysis, as well.*

3.2.2   ANCOVA

The ANCOVA regression can be employed with the panel data (baseline and midline):

$$y_{ig,m} = \quad \alpha_0 + \alpha_1 T_{1g} + \alpha_2 T_{2g} + X_{ig,b}\beta + y_{ig,b} + \epsilon_{ig,m} \tag{1}$$

$$y_{ig,m} = \quad \alpha_0 + \alpha_1 Banked_g + X_{ig,b}\beta + y_{ig,b} + \epsilon_{ig,m} \tag{2}$$

$$\tag{3}$$

the regression controls for baseline characteristics X and baseline measurement of the outcome variable. *April 2019 edit: the notes on clustering and unit of analysis in the previous section apply to this analysis, as well.*

3.2.3   Treatment on the Treated

Finally, we will estimate the TOT effect of linkage to formal banking as well as the TOT effect of linkage to formal bank loans. The TOT regression is

$$y_{ig} = \alpha_0 + \gamma \widehat{Linked}_g + X_{ig}\beta + \epsilon_{ig}$$

where $Linked$ is the indicator for whether the group was ever able to access any banking services (Savings or loans), from any bank, instrumented by the assignment to both savings and savings + loan treatment. *April 2019 edit: the notes on clustering and unit of analysis in the previous section apply to this analysis, as well.*

The second TOT is going to be:

$$y_{ig} = \alpha_0 + \gamma \widehat{Bank\_Loan}_g + X_{ig}\beta + \epsilon_{ig}$$

where $Bank\_Loan$ is the indicator for whether the group was ever able to access a group bank loan, from any bank, instrumented by the assignment to the savings+Loan treatment. The covariate set is the same as described above. *April 2019 edit: the notes on clustering and unit of analysis in the previous section apply to this analysis, as well.*

3.2.4   VSLA level outcomes

We will conduct ITT and TOT for VSLA level outcomes. The ITT regression will be:

$$y_g = \alpha_0 + \alpha_1 T_{1g} + \alpha_2 T_{2g} + X_g \beta + \epsilon_g$$
$$y_g = \alpha_0 + \alpha_1 Banked_g + X_g \beta + \epsilon_g$$

The set of controls X_g includes baseline variables from the screening tool, plus location controls (district and parish fixed effects). *April 2019 edit: most parishes include only a single study village, which makes parish fixed effects colinear with treatment status. We will include district fixed effects, but not parish fixed effects.* TOT regressions are:

$$y_g = \alpha_0 + \alpha_1 \widehat{Linked}_g + X_g \beta + \epsilon_g$$
$$y_g = \alpha_0 + \alpha_1 \widehat{Bank\_Loan}_g + X_g \beta + \epsilon_g$$

3.3 Heterogeneous Effects

One hypothesis we have is that the impact of financial linkages is heterogeneous across members. Specifically, we hypothesize that bank loans are beneficial to members who borrow, as they gain access to credit. We expect (weakly) higher savings amounts among this group, as savings are used to leverage loans. Among savers who do not borrow, we expect the savings rate to decrease, as loans should reduce returns on savings. One caveat to the latter expectation is that it is theoretically possible that savers do not borrow because they are credit rationed. If this is the case, then we should see switching from savings only to borrowing in this group.

To test these hypotheses, we'll use the baseline to separate members who requested loans from the group from members who didn't, and measure treatment effects separately for these two groups.

In addition to the above, we are interested in measuring heterogeneous impacts across groups. We will study whether groups with above median group cohesion (as indicated by baseline measures of group cohesion and satisfaction) differ from groups with below median group cohesion. We will also use a methodology similar to Burlando, Canidio and Selby (2018) to identify credit constrained vs. unconstrained groups at baseline.

3.4 Standard Error Adjustments

We will adjust the p-values of the coefficients of interest for the three primary outcome variables of interest to control for the false discovery rate (FDR) (Benjamini et al. 2006). We will also do the same for the three primary outcome variables of interest for the group level variables.

References