ANALYSIS PLAN FOR MIDLINE SURVEY

Graduating To Resilience in Kamwenge, Uganda
Pre-Analysis Plan
June 22, 2020

Principal researchers: Lasse Brune, Dean Karlan, Nathanael Goldberg, Doug Parkerson, Chris Udry

Implementing Partner: AVSI Foundation

Introduction
This document outlines the pre-analysis plan for analysis of the midline survey of the project "Graduating To Resilience in Kamwenge, Uganda." The purpose of the study is to evaluate the impacts of three versions of a comprehensive livelihood program among refugees in the Rwamwanja Refugee Settlement and among residents in neighboring "host" communities. A particular question of academic interest is the relative effectiveness of a group-coaching approach compared to an individual coaching approach.

Study villages were first randomly assigned to either treatment or control villages.[1] Within treatment villages, participants were randomly assigned at the household level to one of three intervention conditions or a control condition:

- T1, "Individual, asset": consumption support, cash asset transfer, coaching in individual household visits
- T2, "Group, asset": consumption support, cash asset transfer, coaching in group setting and all other aspects of program
- T3, "Individual, no asset": consumption support, *no* cash asset transfer, coaching in individual household visits and all other aspects of program
- C1, "Control in treatment villages": no intervention

Empirical Analysis

For our main empirical strategy, we use a Bayesian approach. While the basic setup, a randomized controlled trial, suggests a straightforward analysis, there are several important aspects of this evaluation that make the Bayesian framework useful in our setting. We can identify several comparisons ex-ante between groups of study participants that are potentially important but may turn out to be irrelevant ex-post. However, pre-specifying many sets of subgroups for analysis would leave us with limited power for hypothesis testing in the classical

[1] For randomization and program implementation, some small villages were combined into larger village combinations and some larger villages were divided up. We refer to the list of resulting village-derived administrative units as "villages" throughout this document.

framework. To the extent that treatment effects are not heterogeneous along an ex-ante identified dimension, we would like to pool observations across those dimensions to obtain more precise estimates. In the classical framework, however, we would need to pre-commit to a model that either estimates separately or one that fully pools observations, or pre-specify a decision-rule that will determine when to fully pool and when to estimate separately. In contrast, in a Bayesian Hierarchical model, the extent to which estimates are pooled is continuous and can be determined by the data. As such, a Bayesian model provides a unified framework for integrating several dimensions of heterogeneity with a transparent and ex-ante specified method for partially pooling estimates.

We include three sets of comparisons in the model that may turn out to be important but for which we have reasons to believe they might not be. This motivates a model structure that allows for partial pooling of information between groups if it turns out that outcomes in the groups are similar. First, members of the refugee and of the host communities may experience different treatment effects. That said, the two groups receive the same interventions and are located in immediate geographic proximity, with similar goods and services markets, which would mean we could expect similar treatment effects. Second, the individual coaching approach in the full program (T1) and the group coaching approach (T2) may produce different outcomes for participants. On the other hand, all other aspects of the program are the same and the modality of coaching could be considered a second-order program difference with respect to many downstream outcomes. Third, we first randomized villages into treatment (of any type) and control villages and then randomized treatments and control group status among households within villages. Comparing treated households in treatment villages with households in the control villages allows us to estimate the effect of the treatments including any spillover effects (among treated households). However, in the absence of spillovers, comparing treated households to control households in treatment villages also estimates treatment effects and we would like to pool information from the two comparisons for more precise estimation.

Lastly, the use of the Bayesian framework allows us to easily incorporate prior information on treatment effects. Since programs that are broadly similar the program being evaluated here have been evaluated with RCTs in a number of contexts, some of them with both village and household level randomizations, information on both treatment effects and spillovers exists and can be used to form priors for the treatment effects.

Modeling and Estimation

Because Bayesian analysis is less common in economics (and social science in general), and analytical methods and norms are still being developed (e.g., see Meager 2019 and 2020), we are not able to commit to the specific Bayesian setup for the analysis. We describe our current thinking below, and include an appendix which explains our process and how we chose the model we chose for final publication.

The starting point for our analysis is a linear regression of intent-to-treat effects. Within treatment villages, we specify the outcome Y of household i in village j located in community $c = c(j)$ as:

$$Y_{ij} = b_{1,c}\, T_{asset,\, ind,\, i} + b_{2,c}\, T_{asset,\, group,\, i} + b_{3,c}\, T_{no\ asset,\, i} + f_c X_{ij} + v_j + e_i \tag{1}$$

The vector X is a set of control variables that includes variables used for stratification and re-randomization and the baseline value of the outcome variable (or close proxies) where available. v_j and e_i are village-level and household-level error terms, respectively. We set baseline variables that are missing to zero and include missing indicators in the list of controls. $c(j)$ maps villages to refugee and host communities and so b_1, b_2, b_3 and f are separately specified for members of the two communities.

The omitted category is being in the control group in treatment villages ("C1"). Thus, parameters b_1 through b_3 capture the difference between households in treatment groups and the households in C1. The difference between b_1 and b_2 measures the relative difference of the individual and group coaching approaches; the difference between b_1 and b_3 is used to assess the marginal effect of the asset transfer.

Under the SUTVA assumptions (Imbens and Rubin 2005), we can interpret estimates of b_1 to b_3 as estimates of causal effects. However, since households interact with each other, no-interference might not hold and treatment effect estimators based on (1) may be biased. We use the first-stage randomization of villages into treatment and control villages to test for spillover effects and adjust estimates. We base our model of the *combined* spillover from households in the intervention arms T1-T3 to controls in treatment villages on the following regression with control households only:

$$Y_{ij} = a + d_c\, Tvil_j + u_{ij} \tag{2}$$

where $Tvil_j$ is a village-level treatment indicator. d captures the difference between control group households in treatment villages compared to households in control villages. We estimate d to test for the presence of spillovers. Under additional assumptions about spillovers between households assigned to T1 - T3 and about the relative contribution of each group to the estimate of the spillover on controls, we can adjust estimates of b_1 - b_3 using the estimate of d. For example, if we assume no spillovers between intervention households and equal spillovers from households across the three intervention groups to controls, $b_1 + d$ is the causal effect of intervention T1.

In addition to the above steps, we also have baseline data on social connections within villages. Define N_{ij} as the set of households in our sample in village j that household i claims to know based on a reading out of a random subsample of names and either a) that i selects as someone to consider getting advice from or b) that i has had a borrowing relationship with or c) that i would consider investing together with, plus any respondents who say the same about household i (even if they are not mentioned by that i). We will examine the more specific hypothesis that spillovers occur through social network connections by estimating

$$Y_{ij} = a + g_c \Sigma_{k \in N_{ij}} 1(k \in Treat) + h_c \#\{N_{ij}\} + u_{ij} \quad (3)$$

where $Treat$ is the set of households in one of the intervention groups T1 - T3. Equation (2) is a more general model of within-village spillovers than (3). If spillovers in fact occur through social connections as measured at baseline and there is variation across control households in the extent of their network connections, then (3) will have additional power to measure these spillover effects. Similar to the adjustment of coefficients in (1) using d from (2) described above, we can then also use the structure in (3) to adjust the estimates in (1).

To estimate the regression in the Bayesian framework we specify a likelihood for each outcome and prior distributions on the parameters of the model. For the most part, we will use weakly informative priors. Given the relatively lower power to estimate spillover effects, we will use literature-based priors based on the data from the experiments in Banerjee et al. (2015). Three of the six sites of this study had both village-level and household-level randomization and, thus, will be used to inform the prior distribution of coefficients in equations (2) and (3).

We plan to model the treatment effects in (1) using a non-nested grouping structure that allows for partial pooling of certain sets of parameters. We group households into refugees and hosts "communities". Similarly, we group together households that receive either the "group-coaching" or the "individual coaching" asset interventions (T1 and T2). Treatment effect parameters will share a prior distribution with a mean that varies with the groupings described above. We also model the grouping of control households into either treatment-village controls and control-village controls; to the extent that we find no spillovers, the "within" and "between" village comparisons should be pooled to increase precision by making use of the larger size of the combined control group.

Since we have few data points to estimate the hyper-parameters for some sets of grouped parameters, we plan to use group variance estimators with lower mean squared errors by using divisors that are larger than J-1 ---where J is the number of groups--- that are used for the standard, unbiased estimators.

Instead of multiple hypothesis test corrections, we will jointly model the three primary outcomes (see list of outcomes below) using a hierarchical structure, with parameters for each outcome sharing priors.

<u>Classical analysis for comparisons</u>

We will also analyze the data using classical (frequentist) regression tools for comparison. We will use randomization inference to compute standard errors and p-values where standard, closed-form variance estimators are not straightforward to apply. We will apply multiple hypothesis test (MHT) corrections across the three summary variables of the primary outcomes we examine (see list of outcomes) but not to the outcomes within groups since we have summary measures for each group. We will present results with separate coefficients as well as p-values from pooling across grouping as discussed further above.

Outcome variables

The outcome variables that we will consider for our final analysis are listed below. We classify them into groups of primary outcomes of interest and secondary outcomes. For each group of primary outcomes, we create a summary measure by summing, averaging or otherwise aggregating the different outcomes within an outcomes group. The summary measure is indicated in the list of variables below. We will also winsorize the data if the standard deviation of a variable is more than 50% larger than its 1%-winsorized version.

Indices

For the construction of outcome group summary measures, we will use the methodology detailed in by Kling, Liebman, and Katz (2007), unless the index is a specific index in the academic literature, in which case we will use the method employed in that literature to compute the index. For those concepts without a preconceived index formula, our methodology consists in first signing all variables consistently such that higher is telling a consistent story for the index. Then, we standardize the individual components of the index, by subtracting the comparison group mean and dividing by its standard deviation. Then, we take the average of the now-standardized components into a single measure, and then again finally standardize the average (again to the comparison group mean and standard deviation). We will use the endline control group for the standardization.

Primary outcomes of interest

Key indicators of livelihood engagement and well-being

1. Value of productive assets ($)
2. Income generating activity index
    - This outcome is an index with three components:
        - Non-agricultural business activity index
        - Livestock activity index
        - Farming activity index
    - The index will be computed as a weighted average of the three components, with weights equal to the proportion of declared intended asset use in each of the three

categories (shares are computed separately for host and refugee communities)

3. Food security index

The components of each summary measure are detailed below.

Time periods: for some variables we collected multiple nested time periods. To aggregate this data into single variables listed in the outcomes lists below or used for their construction, we take the difference of the longer recall period and the shorter recall period and create a weighted average of the shorter period and the difference with the longer period, with weights according to the relative length of the recall periods.

<u>Productive asset ownership</u>

- *Summary measure*: Value of productive assets = sum of outcomes below ($)
- Individual outcomes:
    - Value of non-fixed durable productive assets ($), based on estimated resale value
    - Value of fixed productive assets ($), based on respondent-estimated price of new replacement by contractor
    - Value of livestock ($), computed as sum across all animal types of quantity x price per animal, where price per animal for each type = ½*(median(value of sales/quantity sold) + median(value of purchases/quantity purchased))

<u>Business (non-agricultural)</u>

- *Summary measure*: index of outcomes below
- Individual outcomes:
    - Respondent currently engaged in non-ag business (=1)
    - Number of household members currently engaged in business (#) divided by baseline hh size (ratio)
    - Revenue ($)
    - Expenses ($)
    - Profits (direct question) ($)
    - Value of inventory (raw materials and unsold finished products) ($)

<u>Livestock activity</u>

- *Summary measure*: sum of outcomes below ($)
- Individual outcomes:
    - Household had any livestock activity (=1)

    - Value of sales ($)
    - Sum of variable costs (feed, rent, labor, maintenance, vaccines, medicine) ($)
    - Value of purchases ($)
    - Value of own consumption ($)
    - Value of loss (death, theft) ($)
    - Net sales ($) := sales + own consumption - variable costs - purchases - deaths - depreciated fixed assets

<u>Farming</u>

- Outcomes refer to the Sep-Dec 2019 agricultural season.
- *Summary measure*: index of outcomes below
- Individual outcomes:
    - Household was engaged in farming (=1)
    - Size of land under cultivation (acres)
    - Total spent on fertilizer ($)
    - Total spent on seeds ($)
    - Total spent on pesticides/herbicides/etc ($)

<u>Employment outside of household</u>

- Note we define "employment" defined as "activities that you or any member of your household may have done outside the household for pay or in-kind payment"
- *Summary measure*: index of outcomes below
- Individual outcomes:
    - Respondent was employed (=1)
    - Number of household members who were employed, divided by baseline household size
    - Total income from employment ($)

<u>Food security</u>

- *Summary measure*: index of FCS and HFIAS
- Individual outcomes:
    - Food Consumption Score (FCS) following INDDEX Project (2018), based on number of days in past 7 days household has consumed different major food categories
    - Household Food Insecurity Access Scale (HFIAS) following Coates et al. (2007), based on eighteen questions about incidence and frequency of food security problems in past four weeks

Secondary outcomes

All outcomes that are not listed as primary outcomes are considered secondary outcomes. These variables will be derived from the attached survey instrument and are not listed individually in this document. We will not do adjustments for multiple hypothesis testing when analyzing secondary outcomes, since these are exploratory analyses.

Data collection and timing of analysis

While data collection was completed at the time of submission of this analysis plan, the researchers have generally been blind to the treatment status. An exception was the analysis of survey attrition which was conducted during data collection and included comparisons of attrition rates by treatment status for the purpose of management of potential differential attrition.

References

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. 2015. “A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries.” Science 348 (6236).

Coates, Jennifer, Anne Swindale and Paula Bilinsky. 2007. *Household Food Insecurity Access Scale (HFIAS) for Measurement of Food Access: Indicator Guide*. Version 3. Washington, DC: Food and Nutrition Technical Assistance III Project (FANTA).

INDDEX Project. 2018. *Data4Diets: Building Blocks for Diet-related Food Security Analysis.* Tufts University, Boston, MA. https://inddex.nutrition.tufts.edu/data4diets. Accessed on 5 November 2019.

Kling, J.R., Liebman, J.B. and Katz, L.F. (2007), Experimental Analysis of Neighborhood Effects. Econometrica, 75: 83-119.

Meager, Rachael. 2019. "Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments." American Economic Journal: Applied Economics, 11 (1): 57-91.

Meager, Rachel. 2020. “Aggregating Distributional Treatment Effects: A Bayesian Hierarchical Analysis of the Microcredit Literature”, working paper.