No Lean Season 2017-2019 Pre-Analysis Plan

February 05, 2019

Contents

1 Overview

This document presents a description of the 2017-2019 No Lean Season (NLS) randomized impact evaluation experiment conducted in collaboration between Evidence Action and researchers from Yale University, the London School of Economics, and the University of California, Davis. The first round of the at-scale NLS research was done in 2017-2018, which was the study initially described in earlier versions of this pre-analysis plan. The current version of the pre-analysis plan extends the study to a second round in 2018-2019.

The three main goals of the study are:

1. Replicating previous findings showing positive treatment effects of incentivized migration on seasonal migration, caloric intake, food and non-food expenditure, income, and food security. Our aim is to estimate the impact of a scaled version of the NLS program: intensifying program implementation within branches and expanding the provision of loans to all eligible households.[1]
2. Investigating the program's spillover effects on workers at the migration destination who are not offered migration incentives. Given the scale of the No Lean Season program, we anticipate that there will be enough migration to noticeably affect destination labor markets. Destination workers include those who permanently reside at migration destinations as well as seasonal migrants from other areas. We aim to evaluate the effect of the program on these workers' income and location choice[2].
3. Comparing the differences in the program's effects between the 2008, 2014, 2017 and 2018 studies (the later two are the ones described in this document).

For the second round we make the following changes to the study:

1. Specifying some household, village, and program characteristics over which we plan to evaluate the program's effect heterogeneity.
2. Restricting the eligibility criteria for the program: now only covering a subset of the first round's eligible population.
3. Re-assigning branch level treatment assignment and expanding the study sample to new villages.

In this pre-analysis plan we refer to the scaled research studies as the "first" and "second" rounds. We should note that, *pre-scale*, the first study on migration incentivization predates the No Lean Season program (Bryan, Chowdhury, and Mobarak 2014). We will refer to this original study with the acroynm BCM for short.

In this document we describe the design of the intervention (the No Lean Season program as it will be scaled); lay out the research goals of the study; describe the design of the experiment to investigate our research

[1] In effect moving from making migration subsidies to a few thousand households to over 80,000 households.
[2] The pre-analysis for this section is presented in a separate document.

questions; describe the data we will collect; and finally how we will conduct our analysis.

Table 1: Members of the research team.

Role	
Principal Investigators	Gharad Bryan (London School of Economics)
	Mushfiq Mobarak (Yale University)
	Karim Naguib (Evidence Action)
	Maira Emy Reimao (Evidence Action/Yale University)
	Ashish Shenoy (University of California Davis)
Co-Investigator	Natalie Duarte (Evidence Action)

2 Intervention

No Lean Season is a program, implemented in collaboration between Evidence Action and RDRS, seeking to address seasonal poverty in rural Bangladesh by offering households small, interest-free loans, covering the costs of a round-trip bus fare to nearby areas that do not experience the same seasonal fluctuations. Loans are provided in the Northern Rangpur region of Bangladesh to households with limited land ownership and limited job opportunities in the lean season. The loan is generally offered right before and during the lean season, so participating households can take advantage of the opportunity when it is best for them. Upon return from their migration, households are asked to pay back the loan before the next program cycle begins.[3] Below is a description of each phase of the program.

2.1 Household Targeting

Targeting the appropriate households for the program is the first phase of activity. In this phase, the targeting (baseline) survey is administered to each household in sampled villages and branches, similar to that of a census. The targeting survey lists all households in each village and collects data on the targeting criteria. In BCM, before the No Lean Season program was scaled, researchers used land ownership (50 decimals or fewer) and food security in the last lean season (whether any member of the household skipped meals in the previous lean season) to define eligibility for the program. In the first round of the scaled study, the eligibility criteria were slightly modified to[4]:

1. Cultivable land ownership of 50 or fewer decimals; **or**,
2. Someone in the household did not have a sufficient amount of meals in the two weeks prior to the targeting survey.

In the second round of the program, the criteria were changed again to:

1. Cultivable land ownership of 50 or fewer decimals; **and**,
2. Someone in the household did not have a sufficient amount of meals during the last lean season (2017-2018).

This reversal in the second round to the BCM eligibility criteria is motivated by concerns stemming from the 2017-2018 round that the broader criteria, which increases the number of eligible households within a

[3]These are limited-liability loans, forgiven in the event of extreme adverse shocks or if the migrant fails to secure a job at destination. This feature is not disclosed up front, however, so as to not create negative incentives to recipient households."

[4]The decision in the first round to go with the second criteria on food security, instead of the criteria used in BCM, was made because recent food insecurity may be a more relevant condition for deciding to migrate. Previous research has revealed that decisions on migration were taken considering the household's current situation and negative shocks. Moreover, recall bias is reduced when referencing a more recent time period.

village, combined with the planned scale-up of the program (increasing the number of the villages) may have strained the capacity of the implementing partner, contributing to the low take-up in 2017-2018.[5]

Household targeting surveys are conducted by two partners - RDRS, the No Lean Season implementing partner, and Innovations for Poverty Action (IPA), the research implementing partner[6]. RDRS collected data for the survey in all program villages. In the first round of the study, IPA collected data in all non-treated sampled villages - this includes *spillover*, *branch-control*, and *pure-control* villages (more on these treatment arms below). In the second round of the study, we decided to have IPA collect eligibility data from all villages. So while eligibility in treated villages was still determined by RDRS's data, we could compare how the two data collection activities determine eligibility.

2.2 Information Outreach and Offers

In treated villages, after households have been deemed eligible for the program, they are invited to attend group meetings in their village to learn more about the program. Invited households who attend these meetings receive a standardized presentation on seasonality, the program, and migration for work. After the meeting concludes, households are asked whether they accept in the program offer. There are three possible responses: yes, no, or interested. If households respond yes, they will be added to a list for follow-up and can begin the loan application process. If households respond no or interested, they are given time to think about the offer and Migration Organizers (MOs) - RDRS' frontline of program implementation - will follow-up with the household up to three more times.

For those households unable to attend the offer meeting, they are assigned door-to-door offers and receive an offer at their household after the offer meeting has been completed. All following steps after the door-to-door offer are the same as those for offer meeting.

2.3 Disbursements

In the offer meeting, households are provided information about their respective branch offices and are instructed on the days in which they can pick up the loan. Disbursements are then provided to households that visit their RDRS branch office and successfully fill out the loan applications. The process for receiving a loan is short and only requires households to bring official identification.

In a select group of remote villages, village-based disbursements are carried out. This accounts for approximately 1% of the treated villages. During these village-based disbursements, households are only able to receive disbursements on that day and must travel to the branch office if they desire to take out the loan at a later date.

2.4 Follow-up (soft conditionality)

After households have taken out the loan, MOs follow-up at the household to remind loanees that they have taken out the loan with the intention of migrating for work. This visit is to remind the household that the loan is meant to encourage migration, and to ensure the loanee is aware that they have an obligation to pay back the loan at the end of the lean season. These visits were intended to initially occur one week after the household takes out the loan[7]. If a member from the household has migrated at the time of the first visit, the MO will only visit the household in one month increments to see if the migrant has returned. If a member from the household has not migrated by the first visit, the MO will visit a maximum of once a week until

[5]Included with this pre-analysis plan is a preliminary analysis of the 2017-2018 data showing weaker take-up of migration and a small effect on ecnomic welfare outcomes.

[6]IPA and RDRS implemented slightly different versions of the targeting survey. However, the majority of questions included in both surveys overlap, including eligibility questions. See attached survey instruments for details.

[7]These visits did not actually start until November, so a majority of loan recipients did not receive a visit within one week of taking the loan.

they have/he has confirmed a household member's migration. This phase is intended to monitor the soft conditionality of the loan, but it is important to note that household members are not forced to migrate if they take out the loan.

2.5 Loan repayment and debrief

The final phase of implementation occurs when the migrant has returned to the village. Once the MO learns of the migrant's return, he visits the household to collect repayment and conduct a short migration debrief survey. Households are not required to pay on the first visit after they have returned, and are given up to four chances to repay. The migration debrief survey conducted at the first visit after the migrant's return collects data on the migrant's destination, employment, wages, living expenses, and any experiences that may affect their ability to migrate or find a job. If a household is unable to pay back the loan in full, households are given the opportunity to describe their situation. All requests for loan exemption are then reviewed by RDRS management, and decisions on exemption are later communicated to households by the MO. At the end of the program year (March), all remaining unpaid balances are written off. A household's loan repayment in a given year will never affect its ability to participate in future years of the program.

3 Experiment Design

RDRS is organized administratively into branch offices. Each branch has a set of villages in its catchment area defined by the geographic (road) distance to the branch. Branch catchment areas are non-overlapping so each village in the experiment can be allocated to a single branch.[8]

Treatment, defined as the offer of a migration subsidy (incentivization), occurs at the village level. Every eligible household in a treated village is offered the migration subsidy. At each round of the study, our randomization strategy places villages in one of four categories:

1. *Incentivized*: Villages in which the migration subsidy offer is made.
2. *Spillover*: An untreated village geographically surrounded by treated villages.
3. *Branch-control*: An untreated village that belongs to a treated branch but is surrounded by other untreated villages from that branch.
4. *Pure-control*: An untreated village that belongs to a branch that has no treated villages.

To achieve this classification, we randomize at two levels. First, we randomly divide branches into treated and control. Branches assigned to be control contain only *pure-control* villages. Branches assigned to be treated contain the other three types of villages. Figure 1 shows the two-level branch and village treatment assignment. Transitioning from the first to the second round of the study we randomly assigned branches and villages to either remain at the same treatment assignment or to switch treatment assignment. Table 2 presents how all the branches' and villages' treatment status were changed, and how many new villages were added to the study[9].

[8]In general, a catchment area is defined as all villages that can be reached from the branch office within a 1-hour bike ride.

[9]One branch was dropped in the first round of the study due surveying problems but was re-introduced in the second round of the study; during the midline household survey of the first round of study, it was discovered that one of the *pure-control* village (Mandal Para in Berubari branch) was incorrectly assigned to receive incentivization. It was incorrectly placed in two different branches. In our analysis of the first round we will drop this *pure-control* village. In the second round of study, this branch was added back.

Table 2: Branch and village treatment assignment schedule over the two rounds of the No Lean Season study.

Branch Assignment 2017		Branch Assignment 2018	Number of Branches	Village Assignment 2017		Village Assignment 2018	Number of Villages: Original	New	Total
Control	→	Control	31	Pure-control	→	Pure-control	31	9	40
Treated	→	Treated	20	Incentivized	→	Incentivized	20	0	20
				Spillover	→	Spillover	20	0	20
				Branch-control	→	Branch-control	20	0	20
Control	→	Treated	39	Pure-control	→	Incentivized	39	39	78
				Pure-control	→	Spillover	0	20	20
				Pure-control	→	Branch-control	0	20	20
Treated	→	Control	20	Incentivized	→	Pure-control	20	0	20
				Spillover	→	Pure-control	20	0	20
				Branch-control	→	Pure-control	20	0	20
New	→	Treated	1	Incentivized	→	Incentivized	0	2	2
			111				190	90	280

Within treated branches, our randomization strategy generates a treated sector (designated as *incentivized*), a single untreated village within the treated sector (designated as *spillover*), and an untreated sector (designated as *branch-control*). In accordance with the RDRS workplan, the treated sector comprises approximately a quarter to a third of the villages in a treated branch. For assignment, we identify the centroid of the branch catchment area and then project each village onto a circle around the centroid. We randomly select one village on this circle and designate it as *spillover*. We then define the *incentivized* sector as the fraction of the circle surrounding the *spillover* village[10]. In effect, we create a "pie slice" (designated as the *incentivized* sector), with one village in the middle left untreated as *spillover*.

This strategy stems from the fact that incentivization may generate spillovers onto nearby villages. Spillovers come from three main sources. First, we find in previous work that migrants generally travel in groups and migrants from geographically close sources tend to go to geographically similar destinations. Therefore, inducing migration in one village may lower the returns to migration from nearby villages through the destination labor market. Second, labor markets may be locally integrated. Out-migration from an *incentivized* village may lower labor supply, raise wages, and induce in-migration from nearby villages. Third, household risk sharing networks may extend beyond village boundaries. An *incentivized* household may share the benefits of migration with others in nearby villages.

Our randomization strategy creates multiple types of *non-incentivized* villages to evaluate the geographic extent of these spillovers. The *spillover* village in a treated branch is on average closest to *incentivized* villages and therefore most exposed to treatment spillovers. At the other extreme, we believe *pure-control* villages are sufficiently far from treated regions that their workers are no more exposed to treatment spillovers than workers from anywhere else in the country. *Branch-control* villages fall between these extremes and allow us to estimate how quickly the spillovers dissipate with distance.

For evaluation, we plan to survey (record) households in only a subset of *incentivized*, *branch-control*, and *pure-control* villages. In the first round of th study, survey villages are selected as follows:

1. *Incentivized*: One randomly selected village in the *incentivized* sector per branch.
2. *Spillover*: The village in the middle of the *incentivized* sector, designated as *spillover* in each treated branch.
3. *Branch-control*: The village diametrically opposite the *spillover* village on the circle projection.
4. *Pure-control*: One randomly selected village in each untreated branch.

In the second round of the study, we generally followed the same protocol, however, with the changes in treatment status for branches and villages, as well as the addition of new villages to the study, some changes were made (refer to Table 3 for a break down of the treatment arm sizes, as of the second round of the study):

1. In branches that were control in the first round and that remained control in the second round, we added 9 new *pure-control* villages. Thus, in 9 of these 31 branches, there are two *pure-control* villages.
2. In branches that were treated in the first round and that were changed to control in the second round, all the villages that were previously in the study (*incentivized*, *spillover*, and *branch-control*) were converted to *pure-control*. Thus, is these branches we would have three *pure-control* villages per branch.
3. In branches that were control in the first round and that were changed to treated in the second round, all the *pure-control* villages were converted to *incentivized* with an additional 39 new *incentivized* villages. Thus, in these branches we have two *incentivized* villages per branch. In addition, 20 new *spillover* and 20 new *branch-control* villages were added to these branches. Therefore, 19 of these branches do not have any *spillover* or *branch-control* villages.

The randomization design generates the four experimental categories while ensuring that the status of a village is uncorrelated with other geographic characteristics. In particular, treatment status is orthogonal to the geographic density of villages and their proximity to a branch's boundary. The survey design preserves orthogonality between likelihood of being surveyed and geographic characteristic as well. Unfortunately, in maintaining this orthogonality, we cannot guarantee that *spillover* villages are closer to the *incentivized* sector than *branch-control* in every treated branch. We do not account for proximity to the centroid in randomization, meaning that a very central *branch-control* village may be closer to the treated region than a peripheral

[10] In the first round, this fraction was 1/3 to 1/4 of the villages, while in the second round it was around 1/8.

Table 3: Sizes of village treatment arms in the second (2018-2019) round of the No Lean Season study.

Village Assignment	Old Villages	New Villages	Total
Pure-control	91	9	100
Incentivized	59	41	100
Spillover	20	20	40
Branch-control	20	20	40
	190	90	280

spillover village. However, on average, *spillover* villages are closer to *incentivized* villages than *branch-control* villages. Similarly, an *incentivized* village in our sample is on average closer to the *incentivized* sector than a *branch-control* village, but slightly father on average than a *spillover* village.

The intervention and study sampling will be geographically clustered as follows:

1. *Households*, indexed by i.
2. *Villages*, defined as the set of households, indexed by j.
3. *Branches*, defined as the set of villages served by an RDRS branch office, indexed by k.
4. *Subdistricts*[11], defined as the subdistrincts within which study villages are located, index by m.

In addition, we model subgroups based on household and village covariates as its own level.[12] Thus we define the set of levels as

$$\mathcal{L} = \{\text{study-round, district, sub-district, branch, village, subgroup}\}.$$

Furthermore, we model time relevant outcomes as follows:

1. The study round is identified using the index $r \in \{1, 2\}$ for the first and second round.
2. In each round of the study, household surveys at origin villages be typically be conducted at three points in time: (a) a baseline/targeting survey, (b) a midline survey and (c) an endline survey. Hence, survey time will be indexed with a subscript $t \in \{0, 1, 2\}$ for the three survey rounds, respectively.

We define $\mathcal{V}_k = \{j : k[j] = k\}$, the set of all villages in branch k.

We define the following sizes:

- $N^b_r = N^p_r + N^c_r$ is the number of study branches: N^p is the number of program branches and N^c is the number of non-program branches.[13]
- $\mathbf{N}^v_r = \left(N^v_{kr}\right)_{kr}$ is an N^b-vector of the number of village in each branch.
- N^h_r is the total number of recorded (surveyed) households

For the study population we have the following treatment assignment indicators:

- $Z^p_{kr} \in \{0, 1\}$ is the branch level treatment assignment indicating whether branch k was a program branch.
- $Z^s_{jr} \in \{0, 1\}$ is the village level treatment assignment indicating whether village j is in a program branch and not receiving any migration loan incentives (spillover treatment).
- $Z^l_{jr} \in \{0, 1\}$ is the village level treatment assignment indicating the provision of migration loan incentives.

We also have the following recording assignment indicators:

- $W_{jr} \in \{0, 1\}$ is an indicator that village j has been selected to be surveyed.
- $W_{ir} \in \{0, 1\}$ is an indicator that household i has been selected to be surveyed (thus $W_{ir} \implies W_{j[i],r}$).

[11]Upazilas
[12]We will cover the *subgroup* level in the further detail in th Analysis section.
[13]Where 'program' means that subsidies were offered by that branch, within a portion of its catchment area.

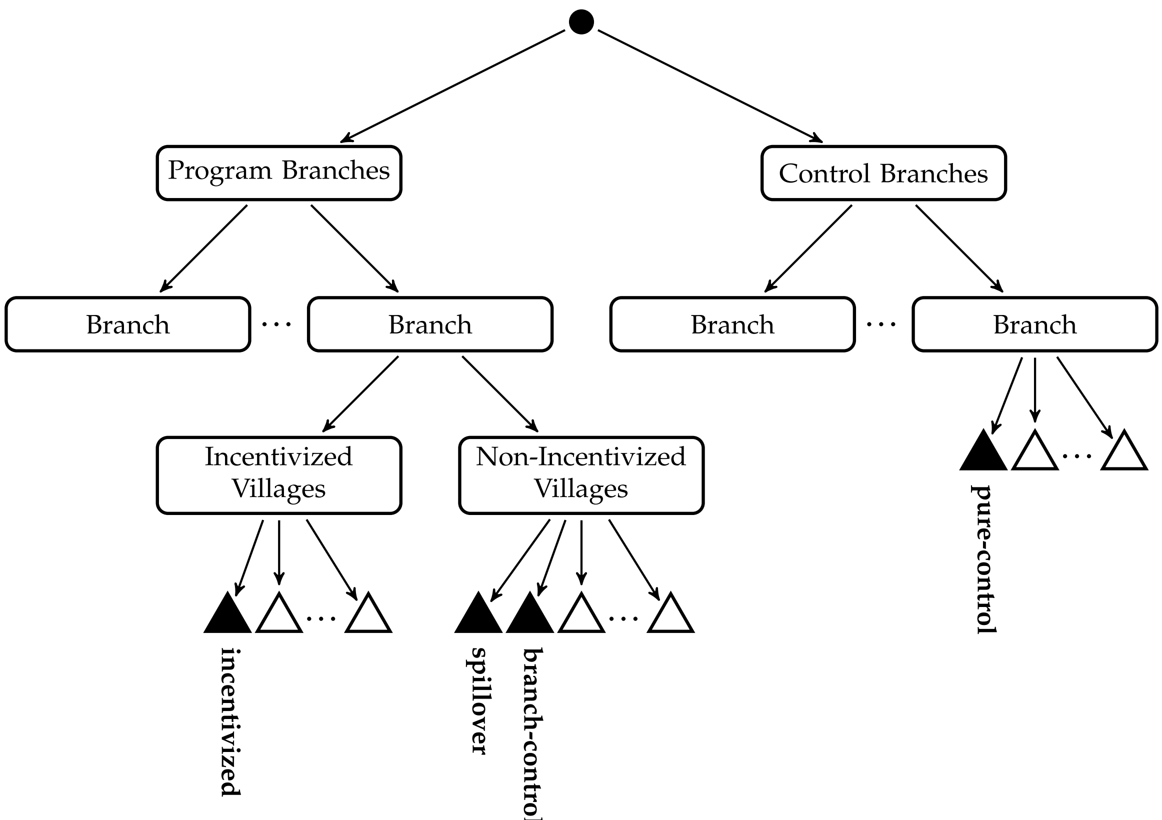


Figure 1: Tree diagram of assignment and recording mechanisms. Triangles represent villages and the solid triangles represent recorded villages.

We also define the categorical variable

$$\widetilde{Z}_{jr} \in \{\text{pure-control, branch-control, spillover, incentivized}\}$$

to identify the treatment group of village j in study round r.

Every village has the treatment assignment vector $Z_{jr} = (Z^p_{k[j],r}, Z^l_{jr}, Z^s_{jr})$. In addition, we define two within branch assignment $\|N^v_r\|$-vectors Z^s_r and Z^l_r, as well as Z^p_r as the program assignment N^b_r-vector. Likewise, W_r is an $\|N^v_r\|$-vector indicating village recording status.

Treatment assignment per study round will be split into the groups shown in Table 4. Since for some branches and villages treatment assignment changed between the first and second round of the study, we have the combined treatment arms shown in Table 5.

Table 4: Experiment Treatment Assignment Groups, at each study round .

$\widetilde{Z}_{jr}$	$Z^p_{k[j],r}$	Z^s_{jr}	Z^l_{jr}
pure-control	0	0	0
branch-control	1	0	0
spillover	1	1	0
incentivized	1	0	1

Table 5: Combined Treatment Assignment. *Original* villages were surveyed in the first round of the study while *New* villages were only surveyed in the second round.

$\widetilde{Z}_{j1}$	$\widetilde{Z}_{j2}$	Number of Villages		
		Original	New	Total
pure-control	pure-control	31	9	40
incentivized	incentivized	20	2	22
spillover	spillover	20	0	20
branch-control	branch-control	20	0	20
pure-control	incentivized	39	39	78
pure-control	spillover	0	20	20
pure-control	branch-control	0	20	20
incentivized	pure-control	20	0	20
spillover	pure-control	20	0	20
branch-control	pure-control	20	0	20

4 Data Collection

This section is to describe the data collection protocol (e.g., samples and random selection, stratification, clusters). Data collection at the origin can be divided into three phases:

1. Baseline/targeting
2. Midline
3. Endline

Household sampling differs between treated and untreated villages due to program implementation. In *incentivized* villages, RDRS carried out a census of all households and then was given a list of households whose responses determined they were eligible for the program. All of these eligible households were given a migration offer. Separately, in each of these treated villages, we chose a random subset of 40 eligible

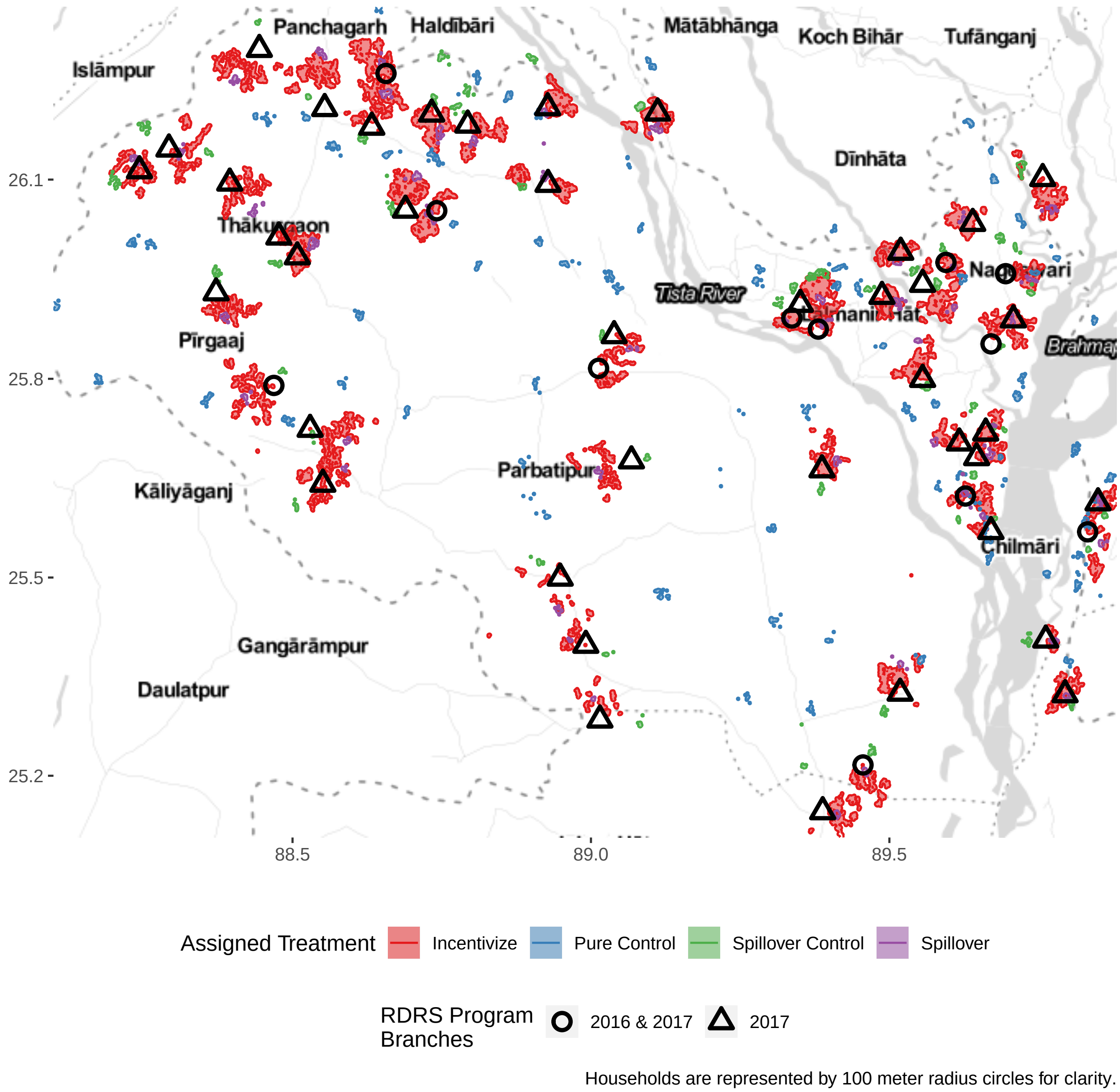


Figure 2: Household Sampling Frame Map, First Round (2017-2018).

households for our research and further surveying. In the first round of the study, for untreated villages (comprising *spillover*, *branch-control*, and *pure-control*), IPA selected households using a random walk. We then evaluated these households for eligibility applying the same criteria as used in treatment villages, and then randomly selected 20 households in each village for further surveying. Both sampling strategies are designed to generate a representative random sample of eligible households, but differ in their implementation. However, in the second round of the study, the baseline survey[14] was conducted in all villages. Eligibility in the *incentivized* villages was still based on RDRS's survey.

4.1 Baseline/Targeting

For program implementation, RDRS conducted a targeting survey in every village selected for incentivization (treatment), to determine eligibility for the loans. Within *incentivized* villages, all households were surveyed through a short questionnaire in June - September, through which basic information was collected on the composition of (potentially) working members in their household; ownership of land; recent migration patterns; and food security in the previous season as well as most recently.

To create a comparable group in non-incentivized villages for our study (*spillover*, *branch-control*, and *pure-control* villages):

- *In the first round of the study*, a questionnaire containing the same set of questions was deployed by IPA in this latter group of villages. Under this activity, carried out in August - September 2017, IPA enumerators applied the survey to 50 randomly selected households in each non-incentivized villages. This is in contrast to the RDRS activity, which covered every household in a village, as the purpose of the IPA survey was only to generate a list of 20 randomly selected households in each non-incentivized village that would have been eligible for the loan had the program been implemented in their village.
- *In the second round of the study*, IPA carried out an eligibility survey in both treated and control villages, again using a random walk. Besides asking the same questions as in the RDRS eligibility questionnaire, this survey included modules on the household roster and consumption as well. Eligibility for inclusion in the study was determined using the same eligiblity criteria as for the migration offers, but based on responses to the IPA survey for both treated and un-treated villages. Within each village, we randomly chose 20 eligible households for follow-up for the rest of the round.

4.2 Midline

A midline survey was deployed in January/February 2018, (first round of the study) and will be deployed in January 2019 (second round of the study) to all households in the study (*incentivized*, *spillover*, *branch-control*, and *pure-control*). The purpose of this survey is to gather data on migration, consumption, and wages/income during the migration period, and questions are largely based off previous survey instruments (e.g., 2008/2011 consumption modules; 2014 high-frequency employment survey). It will be deployed separately from the endline precisely to gather information on dimensions that cannot be reliably measured through recall (e.g., food consumption and wages in a given week) as well as to provide early information on migration patterns and decisions while these are taking place.

Definitions

Household. A household can be of a single person or a group of people living and sleeping in the same house and eating from the same pot.

Household member. A member would be counted as a household member if s/he usually lives and sleeps in the same house and eats from the same pot. If the household head or a student (household member)

[14] Which we refer to as the "round 1 survey" when deemed appropriate, since it was conducted shortly after program implementation had started. To avoid confusion we will continue to refer to it as the "baseline" survey in this document.

lives outside, however, sends to or receive money from the household regularly would be considered as household member. If an individual lives outside the household because they are a student, or due to temporary migration but regularly sends money to the household, they are considered a household member as well. *Respondent.* Ideally the household head should be the respondent. In absence of household head, the eldest son (if adult and knowledgeable) or HH's wife/husband should be selected as respondent. If the eldest son/wife/husband of head of household is not available, any other adult person from the household can be the respondent, if available and knowledgeable about the household.

Household Head. A household member who is recognized as the head of household by the all household members, and who takes decisions in the household.

Migration. We define migration as being outside one's upazila for work for at least 3 nights. In the midline surveys, we ask about migration between September and January, and the endline survey is designed to capture migration over the subsequent 3-4 months.

4.3 Endline

The endline survey was and will be deployed at the end of each migration period, in April/May (2018 and 2019). This survey is more extensive, and will again be applied to all households in our study. The primary purpose of this survey is to collect information on migration and income during the lean season, through we will also use it to investigate secondary outcomes, such as education investment and decisions, credit access and use, and intra-household effects. The instrument for this survey is again largely based off previous tools, particularly the 2011 and 2014 endline survey instruments.

4.3.1 Study Round 1 (2017-2018) RCT Modifications

After the midline household survey was conducted and the data was analyzed, it was decided to change the structure of the endline survey for two reasons: the length of the endline survey as originally proposed was deemed too long and we decided it would be useful to repeat the food and non-food consumption midline survey sections. Hence, endline households was randomly assigned to respond to one of three survey variants (A, B and C), such that a third of each village's respondents were assigned to each variant.

All variants shared a common core set of sections:

- Section 2 (Questions about household living conditions)
- Section 4 (Health)
- Section 5 (Economic activitives and wage employment), part C (Aggregated income from diverse sources)
- Section 9 (Food consumption)[15]
- Section 15 (Risk, coping and shocks)
- Section 16 (Household members' migration), part B (in-depth migration questions)
- Section 17 (Inter-village links)

In addition to the core sections, the three survey variant included:

- *Variant A*:
  - Section 5 (Economic activitives and wage employment), parts A and B
  - Section 6 (Non-agricultural activities)
  - Section 7 (Production)
  - Section 8 (Other assets and income)
- *Variant B*:
  - Section 3 (Land and agricultural wealth)
  - Section 11 (Household assets excluding agricultural assets)
  - Section 12 (Financial assistance)

[15] This is a much shorter version of the midline surey food consumption section.

  - Section 13 (Savings)
  - Section 14 (Gender and social norms)
  - Section 16 (Household members' migration), parts C, H, I and J
- *Variant C*:
  - Food consumption section from the midline survey
  - Non-food expenditure section from the midline survey

4.3.2 Study Round 1 (2017-2018) Survey Errors

In the endline survey of the first study round, two important errors were discovered in the data collected:

1. Respondents were supposed to be asked about their migration experience in the period between the midline and the endline, but due to an instrument coding error they were asked again about the same interval as in the midline survey. In order to remedy this problem, we added a migration module to the baseline survey of the second round of the study (September 2018) to collect information on migration in January-June 2018. Since households included in the first round of the study were tracked in the second round, we were able to recover this lost information.
2. In the food consumption module (added post the midline as described above), we were supposed to ask household respondents how many individuals were present in the household over the consumption interval asked about (*FC* variables). Due to an instrument coding error, this data was *not* collected, and thus we had to rely on the household roster to calculate per person consumption measures.

5 Surveys

This part of the document describes the contents of the datasets obtained from the research activities carried out/to be carried out as outlined in the Data Collection section. Survey instruments will be attached to the pre-analysis document for a more in-depth look at collected data. In the Analysis section of the document we will describe which variables (and how) will be used in our analysis. Table 6 contains details on when the various surveys were conducted and time frames for the main survey questions.

Table 6: Study Surveys Timing

	Survey Timing	Time Frames for Survey Questions			
		Migration	Food Consumption	Non-food Consumption	Income
2008-2011					
Baseline	July 2008	Month and year of first migration	Last 7-14 days	Last week-year	Last 12 months
Round 1 Follow-up	Oct-Nov 2008	Last 4 months	Last 7-14 days	Last week-year	Last 4 months
Round 2 Follow-up	May 2009	1 Sept - 13 Apr (7.5 months)			
Round 3 Follow-up	Nov 2009	Last 6 months (Apr 15 - date)	Last 7-14 days	Last week-year	
Round 4 Follow-up	July 2011	Last 4 months (Feb 15 - Jun 15)	Last 7-14 days	Last week-year	Last 30 days; Last 12 months
2014-2016					
High Frequency Origin Survey	Dec 2014 - Feb 2015	Last week	Last 7 days (short form)	Last 7 days (short form)	Last 7 days
Endline	Jun-Jul 2015	Sept 15 - Feb 15	Last 7 days (food security)		Sept 15 - Feb 15
Follow-up	Aug-Sep 2016	Sept 1 - May 31	Last 7 days (food security)		
2017					
Round 1	May-Sept 2017 (RDRS); Aug-Sept 2017 (IPA)	Last 3 years; Sept-Dec 2016			
Midline	Jan 2018	Last 3 years; Last week; Sept-Jan (detailed)	Last 7-14 days; Last 7 days (food security)	Last week-year	Last 7 days
Endline	Apr-May 2018	Last 6 months (migration income); Sept-Jan (detailed)	Last 7-14 days; Last 7 days (food security)	Last week-year	Last 30 days; Last 6 months (Oct 1 - Apr 1)
Follow-up (overlap with 2018 round 1)	Sept 2018	Last 3 years; Sept - Dec 2017; Feb - Apr 2018; Jan 16 - June 15 (detailed)	Last 7-14 days		
2018					
Round 1	Sept 2018	Last 3 years; Sept - Dec 2017; Feb - Apr 2018			
Midline	Jan 2019	Sept-Dec 2014-2018; Last week; Sept-Jan (detailed)	Last 7-14 days; Last 7 days (food security)		Last 7 days

5.1 Census/Baseline Household Surveys

5.1.1 Study Round 1 (2017-2018)

The purpose of the baseline household survey was to collect data which could be used to determine eligibility for the No Lean Season program. As such, the survey collected basic demographic data that could be used as a proxy to measure household vulnerability during the lean season. The survey was administered to an adult household member with knowledge about the working members of the household (or a working member themselves). The following pieces of data were collected in the baseline survey:

- A roster of (potential) working member names and their respective ages
- Household land ownership (on paper) in decimals - total land and cultivable land
- Amount of land cultivated by the household in the last aman season (includes land owned and rented)
- Household food security in the last 2 weeks preceding the survey (all members in household completed at least 2 meals with satisfaction each day)
- Household food security in the last lean season
- Previous work migration
- Location of previous work migration

Ultimately, cultivable land ownership and food security in the last 2 weeks were used to determine program eligibility.

The baseline survey was completed as a census in treated villages. In control villages, the baseline survey was conducted in random sample of 40 households per village. It was estimated that at least half of the households in treated and control villages would be eligible for the program, providing a sample of roughly 40 eligible households in treated villages and 20 eligible households in control villages, on average.

5.1.2 Study Round 2 (2018-2019)

In the second study round, the baseline survey was conducted after the program had started (but before loan disbursements). The purpose of the survey was similar to the first round survey. The following data was collected:

- A roster of working member names and their respective ages
- Household land ownership (on paper) in decimals - total land and cultivable land
- Amount of land cultivated by the household in the last aman season (includes land owned and rented)
- Household food security in the last 2 weeks preceding the survey (all members in household completed at least 2 meals with satisfaction each day)
- Household food security in the last lean season
- Previous work migration
  - Including **migration between the midline and endline of the first study round** to correct the data collection error in the 2018 endline survey.
- Location of previous work migration
- **Food consumption**

5.2 Midline Household Survey

The main goal of the midline household survey is to capture data that is likely to have a short recall period, for example, food consumption items within the last seven days. This is information critical to evaluating the impact of the program on consumption during the lean season, and it is not likely to be easy to recall at the endline survey 4-5 months after.

5.2.1 Study Round 1 (2017-2018)

5.2.1.1 Household Composition (section 1)

This section of the midline module is composed of two parts, both of which will be repeated in the endline survey.

1. A roster of all household members, their relationship to other household members and general demographic characteristics.
2. Three short questions on school participation within the last seven days.

5.2.1.2 Baseline Eligibility (section 2)

This is a short section of the module repeating some of the baseline survey questions determining eligibility to the NLS program. We will use this data, in addition to other household demographics, to check the balance between households that were surveyed at baseline by different organizations (RDRS and IPA). This and all following survey rounds will be conducted by IPA.

5.2.1.3 Food Consumption (section 3)

This section of the module is composed of two parts:

1. An itemized food consumption list, replicating the data collection used previous studies, and which will mainly be used to impute the caloric intake of household members.
2. A short survey on food security, replicating the data collection conducted in the last study (2014-2015). This part of the survey will be be repeated at the endline.

5.2.1.4 Non-food Expenditure (section 4)

This section replicates the data collection done in previous studies, measuring household expenditure on non-food items.

5.2.1.5 Employment (section 5)

This section asks about the employment over the last 7 days of household members who are 10 years old and above. This is mostly a new section with questions identifying possible inter-village employment. This section will be repeated in the endline survey.

5.2.1.6 Migration (section 6)

This section of the module will capture seasonal migration data for all household members. It is composed of two parts

1. A migration roster
2. In-depth questions on migration episodes leading up to the time of survey

This section will be repeated at the endline survey. The in-depth migration part of the section will be expanded with more detailed questions.

5.2.2 Study Round 2 (2018-2019)

In addtion to the changes described below, these changes were made to the baseline survey:

- Dropped the non-food expenditure section

5.2.2.1 Household Composition (section 1)

Same as above.

5.2.2.2 Baseline Eligibility (section 2)

Same as above, but with the addition of a more comprehensive **migration history** module.

5.2.2.3 Food Consumption (section 3)

Same as above, with the addition of a **consumption seasonality** module.

5.2.2.4 Health (section 4)

Self-reported illness or injuries in the past four weeks. This is a new section to the baseline.

5.2.2.5 Employment (section 5)

Same as above.

5.2.2.6 Migration (section 6)

Same as above with the following additions:

1. A disutility of migration/willingness-to-pay part.

5.2.2.7 Life Satisfaction and Well-being (section 7)

This is a new section, composed of two parts:

1. Standard self-reported measures of **happiness, trust, stress, and worries**.
2. Questions on **gender and empowerment**.

5.2.2.8 Middle-Upper Arm Circumference (MUAC) (section 8)

This is a new section and is measured for all present household members.

5.3 Endline Household Survey

The goal of the endline survey is to capture more comprehensive details about household characteristics, migration and economic activities that are unlikely to have short recall time. Some of the modules from the midline have been copied over to the endline and have been clearly identified in the documentation here and in the survey instruments themselves. We also attempted to keep the section ordering as it was in the 2011 endline survey module (we skip section numbers 10 and 15 since they are no longer applicable in this year's endline survey round).

5.3.1 Study Round 1 (2017-2018)

5.3.1.1 Household Composition (section 1)

This section contains the same questions as in the midline survey in addition to an employment part (copied from section 5 in the midline).

5.3.1.2 Questions about HH Roster (2)

This section asks more in-depth questions about the household such as household construction material and access to water.

5.3.1.3 Land and Agricultural Wealth (section 3)

This section's questions are aimed at measuring land and agricultural assets, similar to what was done in the baseline survey but more in-depth. This is repeat of the section used in the 2011 study.

5.3.1.4 Health (section 4)

This section's questions are aimed at health shocks within the last year. It is a repeat of the section used in the 2011 study.

5.3.1.5 Economic Activity (section 5)

These are detailed questions about economic activity within the last 12 months. There are similar employment questions in section 1 that focus on the last week.

5.3.1.6 Non-agricultural Enterprises (section 6)

These questions on non-agricultural enterprises were taken from the 2011 study instruments.

5.3.1.7 Production (section 7)

This section is split into three parts on

1. Agricultural production
2. Livestock and birds
3. Fishing
4. Forestry

It is a repeat of a module used in the 2011 study.

5.3.1.8 Other Assets and Income (section 8)

This section cover any assets or income not already recorded in other sections. It is a repeat of a module section used in the 2011 study.

5.3.1.9 Food Consumption (section 9)

This section in the endline survey focuses on food security during the past 12 months and the past 7 days. This is an expanded version of the food security part used in the midline survey.

5.3.1.10 Non-agricultural Household Assets (section 11)

This sections collects data on such household assets as appliances, furniture, televisions, radios, etc. This is a repeat of a module section used in the 2011 study.

5.3.1.11 Financial Assistance (section 12)

This section collects data on

1. Financial assistance received
2. Financial assistance given
3. Membership in any MFIs

This is a repeat of a module section used in the 2011 study.

5.3.1.12 Savings (section 13)

This is a short section on savings, copied from the 2011 study endline module.

5.3.1.13 Risk, Coping and Shocks (section 14)

This section focuses on natural and economic shocks that households confronted and how they were dealt with. This is a repeat of a module section used in the 2011 study.

5.3.1.14 Migration (section 16)

This is an expanded version of the migration section used in the midline survey (for example, with questions on remittances).

5.3.2 Study Round 2 (2018-2019)

In addition to the modules used in the first round of the study, we will add a module on treatment experience and compliance, to be deployed only in treatment villages. These include questions on their understanding of the use and conditions of the migration subsidy and contact with RDRS migration officers.

6 Analysis

6.1 Primary Outcomes

- $Y_{it}^{\text{mig}} \in \{0, 1\}$ is a binary indicator of whether any any member of the household migrated, as recorded at data collection round $t > 0$.
- Y_{it}^{con} is a vector of different consumption outcomes measured at round $t > 0$ (not all outcomes will be measured every round):
  - Food expenditure (Taka per person per month), measured at $t > 0$ during the first round of study and $t < 2$ in the second round.
  - Non-food expenditure (Taka per person per month), measured at $t = 1$.
  - Caloric intake (calories per person per day), measured during the midline and endline surveys of the first round of study and measured during the baseline and midline surveys of the second round.
  - Income, measured at $t > 0$.
  - Food security, measured at $t > 0$.
  - Food consumption seasonality, measured at $t = 1$ in the second round of the study.

We denote the potential outcomes in response to assigned treatment z as $Y_{it}^{mig}(z)$ and $Y_{it}^{con}(z)$. We index outcomes in the Y^{con} vector by d.

6.1.1 Food expenditure

- Based on data collected through the midline and endline surveys.
- Reports of items consumed in the last 7 days will be multiplied by 4 to convert to monthly expenditure; and those consumed in the last 14 days will be multiplied by 2 to convert to monthly expenditure.
- We will add Taka amount spent on each item consumed (`Q10`) as well as convert the amount consumed from own production (`Q12`) and other sources (`Q13` and `Q14`) into a Taka value using the mean price of the item.
- Denominator ("per person") will be calculated using information from `FC_2`, which give the number of household members present in the last 7 days.

6.1.2 Total expenditure

- Based on data collected during the midline survey, in the second round of the study, and during the endline endline in the second round.
- It is the sum of food and non-food expenditures over the previous month. Food expenditures will be calculated as described above.
- For non-food expenditures with 1 week recall period, we will multiply expenditures by 4 to convert to monthly; for those with 12 month recall periods, we will divide by 12 to convert to monthly. Those with a recall period of one month will be left as is.
- We will add the Taka amount spent on each item consumed (`Q3`) as well as convert the amount consumed from other sources (`Q4`) into a Taka value using the mean price of the item.
- Denominator ("per person") will be calculated using information from `FC_3`, which give the number of household members present in the last month.

6.1.3 Caloric intake

- Based on data collected through the midline survey, Section 3 Part A.
- Each item reported as consumed in the last 7 or 14 days will be converted to its caloric value using the amount (`Q5`) and unit (`Q6`) consumed, divided by 7 or 14, respectively. These calories will then be added up over all the food items in the survey. We will use the same caloric values for food items as was done in 2008 and 2011. The script used to calculate caloric intake in previous studies will be attached to this pre-analysis plan. As shown, we will calculate `tqtykglcal` and divide it by the number of household members present in the household over the last 7 days.
- Denominator ("per person") will be calculated using information from `FC_2`, which give the number of household members present in the last 7 days.

6.1.4 Income

In previous versions of the pre-analysis plan, the following was the way we described out income would be measured:

- Income will be measured at two points, and there are three versions of this outcome - one from the midline survey and three from the endline. **We will use the third measure from the endline survey as the primary income outcome (4 below)**.
    1. For the midline, the data comes from Section 5 (employment). In particular, we will add the Taka amount earned (`Q10`) from each member over the previous week.
    2. The same questions and data will be collected in the endline, generating a second measurement of labor income over the previous week.
    3. A third measurement of income will come from the data from the endline survey, based on income earned over the previous 30 days, and which includes income from household enterprises and

agriculture (Section 5). We will add to this any non-labor income received in the last 12 months (divided by 12; Section 8 Part B).

4. From the endline household survey section 5 part C, we will sum
   - The gross income from migration
   - As described in Akram, Chowdhury, and Mobarak (2017) "all income and profits earned at home (all income from household's enterprises, and both ag and non-ag wages minus the household's costs in the income-generating activities)" (Table 5, ITT effects on migration income, etc.).

This is the updated approach:

Income is measured, based on the aggregate income data, as:

$$\begin{aligned}\text{Net Income} = \text{Net Own Agricultural Income}\\ + \text{Net Off-farm Income}\\ + \text{Net Non-farm Income}\\ + \text{Net Migration Income},\end{aligned}$$

where

$$\begin{aligned}\text{Net Own Agricultural Income} &= \text{Gross Own Agricultural Income}\\ &\quad - \text{Own Agricultural Costs}\end{aligned}$$

$$\begin{aligned}\text{Net Off-farm Income} &= \text{Gross Off-farm Cash Income} + \text{Gross Off-farm Non-cash Income}\\ &\quad - \text{Off-farm Costs}\end{aligned}$$

$$\begin{aligned}\text{Net Non-farm Income} &= \text{Gross Non-farm Cash Income}\\ &\quad - \text{Non-farm Costs}\end{aligned}$$

$$\begin{aligned}\text{Net Migration Income} &= \text{Gross Migration Cash Income} + \text{Gross Migration Non-cash Income}\\ &\quad - \text{Migration Costs}\end{aligned}$$

- In the first study round:
  - Aggregate income is measured at the midline and endline.
- In the second study round:
  - Aggregate income is measured at the endline only.
  - We will estimate the effect of the program at the midline using reported monetary earnings in the previous 7 days (Section 5, question 10).

6.1.5 Food security

- Food security questions will be asked at the endline only in the first round of the study and at both the midline and endline of the second round. These will be based on a standardized set of questions, also employed in the 2016 follow-up survey (section 8).
- In particular, as is standard when dealing with this type of data, responses to these six questions will be summed to create a "food security index", where a higher score indicates a higher level of food insecurity[16].
- Any further analysis of each food security variable individually will be considered exploratory.

[16] See http://www.fao.org/3/a-i7835e.pdf

6.1.6 Food Consumption Seasonality

During the midline survey of the second round of the study, households will be asked about how often any of their members restricted the portion size of number of meals consumed in day. We will be using, as a primary outcomes, what is reported during Kartik (Oct-Nov 2018) and Agrahayan (Nov-Dec 2018). These are questions 10 and 11 of part C of section 3.

6.2 Primary Impact Estimands

For all the primary outcomes describe above, we are interested in estimating the impact (average treatment effect) between the following treatment groups:

1. **branch-control vs. pure-control**:

   This identifies the impact of program spillover on non-incentivized villages, excepting spillover from neighboring incentivized villages (i.e., impact on non-incentivized villages that are as distant from incentivized villages in their branches as possible). This captures any impact program implementation might have on villages while limiting inter-village spillover.

2. **incentivized vs. pure-control**:

   This identifies the total impact of the program as it would be scaled, combining the effects of incentivization, program, and inter-village spillovers, by comparing *incentivized* villages to the counterfactual of having no program at all. If we find estimand (1) to be lower than a specific threshold we will combine the *pure-control* and *branch-control* groups and estimate **incentivized vs. (pure-control and branch-control)**. This threshold will be a 5% increase in both total expenditure and migration, statistically significant at the 10% level using a Fisher exact test.

3. **spillover vs. branch-control**:

   This identifies the impact of inter-village spillover from neighboring *incentivized* villages on non-incentivized villages. This could be driven by a number of mechanisms:

   a) Spillover at destination labor markets, due to an increase in migration take-up from neighboring *incentivized* villages leading to an increase an in labor supply in destination markets.
   b) Spillover between villages as migration in neighboring incentivized villages leads to
      i. Greater employment opportunities in *incentivized* villages
      ii. Better networking and/or information sharing about seasonal migration
      iii. Changes in prices

4. **incentivized vs. spillover**:

   This identifies the impact of incentivization net spillover effects. This hinges on spillover effects being equal between *incentivized* and *spillover* villages which might not be true since *incentivized* villages were clustered by design around *spillover* villages and recorded *incentivized* villages were randomly selected from the pool of *incentivized* villages in each branch. Thus, recorded *incentivized* villages might not be exposed to the same intensity of spillover as the *spillover* villages.

In particular, for migration we are interested in (a) the average difference in the probability of seasonal migration and (b) the average number of seasonal migration episodes. For the other welfare outcomes (expenditure, caloric intake, income, etc.) we want to estimate the average scalar difference in measured outcomes.

6.3 Covariates

Our analysis will control for pre-treatment household characteristics:

- Household education
- Percentage expenditure on food[17]
- Number of adult males
- Number of children
- Borrowing

These are same endline covariates used in Bryan, Chowdhury, and Mobarak (2014). In addition, the following are added:

- Last migration experience (from the eligibility section)
  - (i) Never migrated or migrated >3 years ago
  - (ii) Migrated 2-3 years ago
  - (iii) Migrated a year ago
- Cultivable land (from the eligibility section). This is split into:
  - (i) No land
  - (ii) Below the median conditional on $0 < \text{land} \leq 50$ decimals
  - (iii) Above the median conditional on $0 < \text{land} \leq 50$ decimals
  - (iv) Below the median conditional on land > 50 decimals
  - (v) Above the median conditional on land > 50 decimals
- Consumed full meals in the previous year (from the eligibility section)
- Consumed full meals in the previous week (from the eligibility section)
- Village experienced flooding (this information was collected in the origin price survey of 2018)
- Loan disbursement[18] (from RDRS's administrative data)
- Baseline caloric intake (Second study round only). This continuous outcome will be discretized into quantiles.

6.4 Subgroups

6.4.1 Study Round 1 (2017-2018)

Criteria for program eligibility in this study have changed from those used in prior studies. In the current study, a household is eligible to receive an incentive loan if (i) ownership of cultivable land is less than or equal to 50 decimals *or* (ii) any household member has missed a meals in the previous week. In prior studies, a household was eligible if (i) ownership of cultivable land was less than or equal to 50 decimals *and* (ii) any household member had missed a meal in the previous year. Since prior findings were based on the second eligibility criteria, we plan to conduct our analysis on

1. Household eligible according to the current criteria
2. Household eligible according to the previous criteria

6.4.2 Study Round 2 (2018-2019)

As previously mentioned, in the second round of the study only the second subgroup (above) will be analyzed.

6.5 Outliers

For continuous outcomes

- i. Food expenditure,
- ii. Total expenditure,

[17] Only if available from the baseline survey.

[18] As this is an endogenous outcome observed only in the treated arm, we rely on multiple imputation to condition on loan-taking.

iii. Caloric intake, and
iv. Income

we will trim outliers, which we define as outside outside the interval

$$[\mathrm{Q}_1(\mathbf{y}) - 1.5 \times \mathrm{IQR}(\mathbf{y}), \mathrm{Q}_3(\mathbf{y}) + 1.5 \times \mathrm{IQR}(\mathbf{y})],$$

where $\mathrm{Q}_q(\mathbf{y})$ is the q^{th} quartile of observations $\mathbf{y}$ and $\mathrm{IQR}(\mathbf{y})$ is the inter-quartile range.

6.6 Empirical Strategy

We are committing to analyze the outcomes of this study using the three approaches described below, each with a slightly different motivation but altogether presenting a more robust inference.

6.6.1 Regression Analysis

For this analysis we will use the following linear model specifications to estimate and test the statistical significance of our primary estimands described above.

1. **branch-control vs. pure-control**: For all observations i such that $\widetilde{Z}_{j[i]} \in$ {pure-control, branch-control} we estimate the model
$$Y_{it}^q = \alpha_1^q + \beta_1^q \cdot Z_{k[i]}^p + \delta_{1,m[i]}^q + \varepsilon_{1,it}^q, \tag{1}$$
and test
$$\begin{aligned} \mathrm{H_o} &: \beta_1^q = 0 \\ \mathrm{H_a} &: \beta_1^q \neq 0 \end{aligned}$$

2. **incentivized vs. pure-control**: For all observations i such that $\widetilde{Z}_{j[i]} \in$ {pure-control, incentivized} we estimate the model
$$Y_{it}^q = \alpha_2^q + \beta_2^q \cdot Z_{j[i]}^l + \delta_{2,m[i]}^q + \varepsilon_{2,it}^q, \tag{2}$$
and test
$$\begin{aligned} \mathrm{H_o} &: \beta_2^q = 0 \\ \mathrm{H_a} &: \beta_2^q \neq 0 \end{aligned}$$

3. **spillover vs. branch-control**: For all observations i such that $\widetilde{Z}_{j[i]} \in$ {spillover, branch-control} we estimate the model
$$Y_{it}^q = \alpha_3^q + \beta_3^q \cdot Z_{j[i]}^s + \delta_{3,m[i]}^q + \varepsilon_{3,it}^q, \tag{3}$$
and test
$$\begin{aligned} \mathrm{H_o} &: \beta_3^q = 0 \\ \mathrm{H_a} &: \beta_3^q \neq 0 \end{aligned}$$

4. **incentivized vs. spillover**: For all observations i such that $\widetilde{Z}_{j[i]} \in$ {spillover, incentivized} we estimate the model
$$Y_{it}^q = \alpha_4^q + \beta_4^q \cdot Z_{j[i]}^l + \delta_{4,m[i]}^q + \varepsilon_{4,it}^q, \tag{4}$$
and test
$$\begin{aligned} \mathrm{H_o} &: \beta_4^q = 0 \\ \mathrm{H_a} &: \beta_4^q \neq 0 \end{aligned}$$

Where $q \in \{\text{mig}, \text{con}\}$, the primary outcomes studies, and δ^q is a subdistrict fixed effect.

We will conduct the following statistical tests:

1. For each individual hypothesis above and for each primary outcome, we will test the null hypothesis of *zero average treatment effect*:
   a) Using conventional cluster robust standard errors
   b) Multiple hypotheses corrected procedures (e.g. List, Shaikh, and Xu (2016), Westfall and Young (1993))
2. For each individual hypothesis above and for each primary outcome, we will test the null hypothesis of *zero treatment effect*, using a Fisher exact test (Young 2017).
3. For each primary outcome, we will test the null joint hypothesis of *zero treatment effect* for all experimental treatments using a cross equation joint test as proposed by Young (2017).
4. For all primary outcomes and all experimental treatments, we will conduct a omnibus test of no effect (Young 2017).

6.6.2 Model-Based Multilevel Analysis

We construct a multilevel Bayesian model incorporating all available and relevant regional/geographic information, as well as household and village characteristics. Thus we model heterogeneity in treatment effects over

(i) Study round
(ii) Districts (zila)
(iii) Sub-districts (upazila)
(iv) RDRS branches
(v) Villages
(vi) Covariate subgroups (more on these below).

Levels (iv) and (v) are the ones used in the study's two stage randomization.

First, we start with the process generating observed outcomes (the likelihood model), and then we delve deeper into the hierarchy of parameters.

- For **continuous outcomes**, such as caloric intake or income, we model observed outcomes as

  $$\widetilde{Y}_i \sim \texttt{Normal}(\mu_i, 1),$$

  where $\widetilde{Y}_i$ is scaled Y_i (we scale outcomes to easily define priors below) and μ_i is the linear predictor.

- For **binary discrete outcomes** we model observed outcomes as

  $$Y_i \sim \texttt{Bernoulli}(\text{logit}^{-1}(\mu_i)),$$

  where $Y_i \in \{0, 1\}$ and μ_i is the linear predictor (for the latent variable).

- For **ordered multiple discrete outcomes** we model observed outcomes as

  $$Y_i \sim \texttt{OrderedLogistic}(\mu_i, \mathbf{c}),$$

  where $Y_i \in \{1, L\}$, $\mathbf{c}$ is vector of *cutoffs*, of length $L - 1$, and μ_i is the linear predictor.

Table 7 presents all the outcomes modeled and their types.

We need to add a new indices to represent the different outcomes analyzed. We two sets of outcomes analyzed

$$\mathcal{Q}^{\text{endo}} = \text{All Endogenous Outcomes}$$
$$\mathcal{Q}^{\text{exo}} = \text{All Exogenous Outcomes.}$$

Table 7: Likelihood types for all modeled variables.

Outcome	Type	Likelihood Model
Endogenous Outcomes		
Migration	Binary Discrete	Logistic
Caloric Intake	Continuous	Normal
Expenditure	Continuous	Normal
Income	Continuous	Normal
Food Security	Ordered Multiple Discrete	Ordered Logistic
Exogenous Outcomes (Covariates)		
Last Migration Experience	Ordered Multiple Discrete	Ordered Logistic
Cultivable Land	Ordered Multiple Discrete	Ordered Logistic
Full Meal Last Year	Binary Discrete	Logistic
Full Meal Last Week	Binary Discrete	Logistic
Flooded Village	Binary Discrete	Logistic
Loan Taken	Binary Discrete	Logistic
Baseline Caloric Intake	Ordered Multiple Discrete	Ordered Logistic

The linear predictors mentioned above are modeled below.[19] However, we need to also define $e[i, l]$ as the *entity* of level l that observeration i belongs to. For example, if $l = \textit{village}$ then $e[i, l] = j[i]$, the village that i belongs to. For endogenous outcomes, $q \in \boldsymbol{Q}^{\text{endo}}$, we have

$$
\begin{aligned}
\mu_i = \alpha &+ \beta_1 \cdot \texttt{Incentivized}_{j[i]} + \beta_2 \cdot \texttt{Spillover}_{j[i]} + \beta_3 \cdot \texttt{BranchControl}_{j[i]} \\
&+ \sum_{l \in \mathcal{L}} \alpha_{l,e[i,l]} + \beta_{1,l,e[i,l]} \cdot \texttt{Incentivized}_{j[i]} + \beta_{2,l,e[i,l]} \cdot \texttt{Spillover}_{j[i]} + \beta_{3,l,e[i,l]} \cdot \texttt{BranchControl}_{j[i]} \\
&+ \eta_i^q,
\end{aligned} \tag{5}
$$

while for exogenous outcomes, $q \in \boldsymbol{Q}^{\text{exo}}$, we have

$$
\mu_i = \alpha + \sum_{l \in \mathcal{L}} \alpha_{l,e[i,l]} + \eta_i^q.
$$

The parameter η_i^q is an observation level random effect used to capture any potential correlation between the outcomes measures per observation i (modeling correlation will be covered when we discuss model priors).[20] This is equivalent to modeling the observation i (e.g., household) as another level of the model with multiple outcomes measured.

As mentioned above, we include a *subgroup* level to the model. This is a partition of the study population into cells defined by a fully saturated interaction model of the exogenous outcomes.[21] Thus, if there is a total of L^{total} possible outcomes of all the exogenous variables, there are $2^{L^{\text{total}}}$ subgroups (or level entities, using the level terminology used above).

[19] For readability, we use

$$
\begin{aligned}
\texttt{Incentivized}_j &\equiv \mathbb{1}\{\widetilde{Z}_j = \text{incentivized}\} \\
\texttt{Spillover}_j &\equiv \mathbb{1}\{\widetilde{Z}_j = \text{spillover}\} \\
\texttt{BranchControl}_j &\equiv \mathbb{1}\{\widetilde{Z}_j = \text{branch-control}\}.
\end{aligned}
$$

[20] We drop the q index from the rest of the parameters in the model for readability.

[21] Which have all been modeled as having finite discrete values.

While we are attempting here to lay out all the details of the model, we do expect the model to evolve in reponse to model checking and criticism. For example, as it is currently modeled, the subgroup level does not have any interaction with other levels. This is something that we might have to change to capture potential heterogeneity in the effects of covariates across geography or study rounds. The same is true for the priors that we present below.

6.6.2.1 Priors

While we do have some prior beliefs of the program's effects (from previous studies), we are not explicitly modeling them here. Instead, we use the model's weakly informative priors to regularize our analysis in order to mitigate overfitting. This is a part of the model that will also be checked after the data is analyzed.

Priors for the super-population parameters, $\alpha, \beta_1, \dots, \beta_3$, are

$$\begin{aligned} \alpha &\sim \texttt{Normal}(0, 5) \\ \beta_1, \dots, \beta_3 &\sim \texttt{Normal}(0, 1) \end{aligned}$$

Priors for all levels, $l \in \mathcal{L}$, and for all entities e in each level l, are

$$\begin{pmatrix} \alpha_{le} \\ \beta_{1le} \\ \vdots \\ \beta_{3le} \end{pmatrix} \sim \texttt{MultiNormal}(\mathbf{0}, \Sigma_l),$$

where Σ_l is the covariance matrix that we further define as

$$\Sigma_l \equiv \texttt{diag}(\tau_l)\Omega_l\texttt{diag}(\tau_l).$$

The parameters τ_l and Ω_l are the vector of treatment effect standard deviations and the treatment effect correlation matrix, respectively. For $l \in$ {study-round, district, sub-district, branch, subgroup}, we set the priors[22]

$$\begin{aligned} \tau_l &\sim \texttt{Normal}(0, 1) \\ \Omega_l &\sim \texttt{LKJCorr}(3), \end{aligned}$$

while for $l = \textit{village}$,

$$\begin{aligned} \tau_l &\sim \texttt{Normal}(0, 0.5) \\ \Omega_l &\sim \mathbf{I}_4. \end{aligned}$$

For the observation level random effects, η_i^q, we use the priors

$$\begin{aligned} \boldsymbol{\eta}_i^{\text{endo}} &\sim \texttt{Normal}(\mathbf{0}, \Sigma_\eta^{\text{endo}}) \\ \boldsymbol{\eta}_i^{\text{exo}} &\sim \texttt{Normal}(\mathbf{0}, \Sigma_\eta^{\text{exo}}), \end{aligned}$$

where $\boldsymbol{\eta}_i^{\text{endo}}$ and $\boldsymbol{\eta}_i^{\text{exo}}$ are the random effect vectors capturing correlation within units of observation between endogenous and exogenous outcomes, respectively. We define

$$\begin{aligned} \Sigma_\eta^{\text{endo}} &\equiv \texttt{diag}(\tau_\eta^{\text{endo}})\Omega_\eta^{\text{endo}}\texttt{diag}(\tau_\eta^{\text{endo}}) \\ \Sigma_\eta^{\text{exo}} &\equiv \texttt{diag}(\tau_\eta^{\text{exo}})\Omega_\eta^{\text{exo}}\texttt{diag}(\tau_\eta^{\text{eoo}}), \end{aligned}$$

[22] The second parameter for `Normal` distributions is the standard deviation.
For villages, we don't model any correlation between treatment effects.

where

$$
\begin{aligned}
\tau_\eta^{\text{endo}} &\sim \texttt{Normal}(0, 0.5) \\
\tau_\eta^{\text{exo}} &\sim \texttt{Normal}(0, 0.5) \\
\Omega_\eta^{\text{endo}} &\sim \texttt{LKJCorr}(2) \\
\Omega_\eta^{\text{exo}} &\sim \texttt{LKJCorr}(2).
\end{aligned}
$$

6.6.2.2 Estimation

The primary goal of the model-based analysis is generating posterior distributions for *finite sample* average/median treatment effects identifying the Primary Impact Estimands. This differs from the *super-population* estimation done in the Regression Analysis section. The purpose is to impute unobserved counterfactuals for all observations in the study and reporting their average/median differences (Imbens and Rubin 2015).

Let

$$\Delta_i(z, z')$$

represent the difference in i's counterfactual outcomes for treatments z and z', and thus the finite sample average treatment effect is defined as

$$\hat{E}[\Delta_i(z, z')]$$

and the median

$$\texttt{median}[\Delta_i(z, z')]$$

Additionally, we can use village size data to conduct *finite population* estimation, essentially by imputing the unobserved outcomes of the population in the villages under study.

6.6.2.3 Model-based Analysis Only Estimands

6.6.2.3.1 Within- and Post-Program Spillover

In the second round of the study, we test the effect of being in an incentivized, spillover, or branch-control village relative to pure control. Since this is the second year of the program, we can use the research design to test for temporal spillovers in these categories as well as main effects.

For each treatment assignment $\widetilde{Z} \in$ {incentivized, spillover, branch-control}, we define the within-program spillover and post-program spillover of $\widetilde{Z}$. We define the with-program spillover as the effect of having been assigned to $\widetilde{Z}$ in a prior year and continuing to be assigned to $\widetilde{Z}$. This value captures the interaction of treatment over multiple years of the program. In contrast, we define the post-program spillover as the effect of having been assigned to $\widetilde{Z}$ in a prior year and no longer being part of the program. This value captures the lasting impact of the program once subsidies are withdrawn.

To compute within-program spillovers, we make a comparison among villages that are currently assigned to treatment arm $\widetilde{Z}$. In this group, we compare villages that were previously assigned to $\widetilde{Z}$ against those that were previously assigned to pure-control. This comparison identifies the causal impact of having previously been in treatment arm $\widetilde{Z}$ on the treatment effect of $\widetilde{Z}$.

To compute post-program spillovers, we make a comparison among villages that are currently assigned to pure-control. In this group, we compare villages that were previously assigned to $\widetilde{Z}$ against those that were previously assigned to pure control. This comparison identifies the causal impact of having been treated after treatment is withdrawn relative to having never been treated.

Formally, let $Y(z_2, z_1)$ be the outcome for villages with current treatment status z_2 and prior treatment status z_1, where $z_1, z_2 \in$ {incentivized, spillover, branch-control, pure-control}. Then the two temporal spillover effects are:

- Within-program spillover: $\tau^{\text{within}}(z) = Y(z, z) - Y(z, \text{pure-control})$.
- Post-program spillover: $\tau^{\text{post}}(z) = Y(\text{pure-control}, z) - Y(\text{pure-control, pure-control})$.

To make estimands consistent between study rounds, counterfactual outcomes for all arms in first round of the study are modeled as $Y(z, \text{pure-control})$ where z is the treatment assigned; there is not prior treatment since the program was not yet scaled.

6.7 Comparing Pre- and Post-scale Studies and Interpreting their Differences

In this section we will posit a number of mechanisms for us to investigate in the event of finding a significant difference comparing the program's effect on migration and other primary outcomes over the two rounds of the study. Primarily, we want to try to understand to what extent implementation changes between the two post-scale studies rounds might have led to changes in the take-up of loans, migration, and other economic welfare outcomes.

If we observe a similar effect in both rounds, we would focus on investigating how the scaled program differed from the 2008 and 2014 RCTs that had a relatively stronger effect on migration and other outcomes.

In this section, we will lay out the possible mechanisms including what data we plan to use that would help us conduct non-experimental inference on these possible causes/explanations.

6.7.1 Household Mis-targeting

This mechanism dampens the effect of the program by effectively turning away households that would have migrated had they received a loan and not otherwise. This is the subpopulation that were induced to take up seasonal migration in the 2008 and 2014 studies — the compliers.

There are two main types of this mechanism that we will investigate. One is related to explicit quotas set by RDRS branches or other resource constraints. This could lead MOs to exclude households in the complier subpopulation, or perhaps compliers self-select themselves out by perhaps being slower to seek migration loans. The second type of mis-targeting could be due to MOs perceiving, without explicit instructions from their branches, households that are more experienced with season migration as more reliable or would benefit more from the program. This would also result in excluding compliers.

- Implementation constraints on household targeting.
  - Village-level quotas that if are set too low could cause MOs to stop seeking new clients once quotas are reached.
  - Time constraints/heavy MO workload.
  - Not enough cash to disburse at branch offices.
- Targeting households likely to have migrated without incentivization

Note also that while in 2017 we expanded the eligibility criteria to include households that had either skipped meals or less than 50 decimals of land, the eligibility criteria was reverted back in 2018 to require that households meet both criteria. This means that the 2018 eligibility criteria is the same as for 2008 and 2014.

6.7.1.1 Data

- Endline MO surveys
  - Explicit or implicit targeting criteria
  - Any reported targeting quota
- Heterogeneity based on eligibles per branch (or per MO)
- Household eligibility characteristics
- Migration history from household surveys
- Administrative data on
  - Cash for loan disbursement

  - MO allocation and workload
  - Loan disbursement

6.7.2 Breakdown in monitoring/conditionality enforcement

The NLS program relies on soft conditionality since it would impractical to closely monitor all loan recipients. This could lead households to increase their take-up of loans and not actually migrate. A more serious problem with conditionality is if MOs were actually more forceful in their monitoring which could discourage complier households.

6.7.2.1 Data

We will include in the endline a module on households' experience with RDRS, to be asked in treatment villages. This module will include their understanding of the terms of the offer, the extent to which migration was verified, and any expected repercussions if failing to follow the conditions.

6.7.3 2017 was an anomalous year

Environmental conditions could have resulted to 2017 being a strong outlier in terms of effectiveness. Some types of shocks that could have a strong negative effect on migration:

- Severe flooding
- Labor strikes
- Road closures
- Political unrest

6.7.3.1 Data

- Village level reports on flooding

6.7.4 Scale specific issues

This category captures negative effects on migration take-up that are caused by the program's scale. Studies prior to 2017 were significantly smaller and thus less likely to reveal these problems.

There are two main mechanisms we will look at. First, it is possible that size and concentration of the program could have caused negative spillovers between villages, possibly due to greater competition for employment at migration destination.

Second, it is possible that local heterogeneity results in some areas being more suitable for seasonal migration than others. This could be due to the distance needed to travel to get a loan from the closes RDRS branch, the distance to roads and destination cities, or perhaps sensitivity to pre-existing migration levels.

6.7.4.1 Data

- Spillover analysis from experimental data
- Treatment effect heterogeneity from experimental data
  - Multilevel analysis
  - Estimation of correlation between base rate and treatment effects within districts and sub-districts.
  - Consider effects when limited to the two districts in which the previous rounds were implemented.
- Household and village level characteristics
  - Distance to RDRS branch

- Distance to main road
- Distance to destination[23]

6.7.5 Economic growth eliminated the need for migration

Finally, the need for seasonal migration could have been eliminated due to changes in rural labor markets. Perhaps wages increased, there was an increase in local employment opportunities, or other solutions for smoothing consumption during the lean season became available.

6.7.5.1 Data

- Wage, employment and food security data in control villages

7 Document History

Date	
Jan 16, 2018	IPA notified of data embargo (restriction)
Jan 31, 2018	Describing the empirical strategies we plan to use
Feb 1, 2018	Clarifying how primary outcomes will be calculated
Feb 6, 2018	Specifying analysis subgroups
Feb 19, 2018	Plot showing village assignment mechanism, within program branches
Mar 15, 2018	Remove at destination sections (moved to separate document). Update on overlapped pure-control village. Study households map. Covariates. Assignment/recording mechanism implication for analysis.
May 21, 2018	Update household endline survey variants protocol
Jan 21, 2019	Renamed all *spillover-control* to *branch-control*. Document round 1 survey errors and workarounds. Document changes to the baseline and midline surveys in round 2. Document update to primary outcomes. Document new estimands to calculate in round 2. Update the list of covariates. Document trimming of outliers. Document multilevel model used in the 2017-2018 analysis.
Feb 4, 2019	Added survey timing and time frames table. Moved temporal spillover section to the model-based analysis.
Feb 5, 2019	Added a section on comparing study rounds.

[23] We will need to impute what destinations the compliers would have gone to in the control villages.

Appendices

Appendix A: Assignment Mechanism

First, we will describe the assignment mechanism in the first round of the study (we drop the r index for readability). Random assignment to treatment is designed in a manner concentrating *incentivized* villages ($Z^l_j = 1$) around the *spillover* village ($Z^s_j = 1$), in program villages ($Z^p_k[j] = 1$). This is done as follows:

1. Construct the circle-distance $N^v_k \times N^v_k$ matrices $\{\mathbf{D}_k\}_k$.

   For each branch k:

   a) Project all village onto the surface of a circle centered on the branch's centroid (based on the geographic location of all branch villages). The actual radius of the circle is irrelevant, but we set it to the smallest (geographic) distance between a branch village and the centroid. See Figure 3.

      i. For each village $j \in \mathcal{V}_k$: Calculate the point of intersection of a ray projecting from the centroiding and intersecting j geographic location.
      ii. Let $\mathcal{D}_k$ be the matrix of geographic distances between every pair of villages' projected locations.

2. For each program branch k ($Z^p_k = 1$):

   a) Randomly select one village to be the *spillover* village.
   b) Assign the $\text{round}\left(\frac{N^v_k}{3}\right)$ villages with the smallest circle-distance to the *spillover* village to be *incentivized*.
   c) Assign the village with the greatest circle-distance to the *spillover* village to be the *branch-control* village.

$$p(Z^p, Z^s, Z^l|\{\mathbf{D}_k\}_k) = p(Z^p) \cdot p(Z^s, Z^l|Z^p, \{\mathbf{D}_k\}_k)$$

where the branch level assignment mechanism is

$$p(Z^p) = \begin{cases} \binom{N^b}{N^p}^{-1} & \text{if } \|Z^p\| = N^p \\ 0 & \text{otherwise} \end{cases}$$

and the village level assignment mechanism is

$$p(Z^s, Z^l|Z^p, \{\mathbf{D_k}\}_k) = \prod_k p(Z^s_k, Z^l_k|Z^p_k, \mathbf{D}_k)$$

and[24]

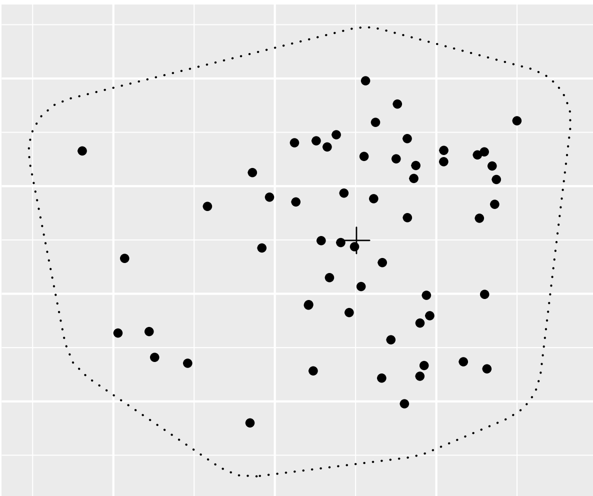
$$p(Z^s_k, Z^l_k|Z^p_k, N^v_k) = \begin{cases} 1 & \text{if } Z^p_k = 0 \wedge \|Z^s_k\| = 0 \wedge \|Z^l_k\| = 0 \\ 1/N^v_k & \text{if } Z^p_k = 1 \wedge Z^s_k \cdot Z^l_k = 0 \wedge \texttt{is\_in\_incentivize\_wedge}(Z^s_k, Z^l_k, \mathbf{D}_k) \\ 0 & \text{otherwise} \end{cases}$$

Z^s_k and Z^l_k are the N^v_k-vector for village treatment assignment within branch k and `is_in_incentivize_wedge`(s, l, d) is an indicator function for whether all incentivized villages (Z^l_k) are within a "incentivization wedge" around the *spillover* village (Z^s_k). Figure 4 shows how the (non-zero) probability of treatment assignment of villages in program branches varies.

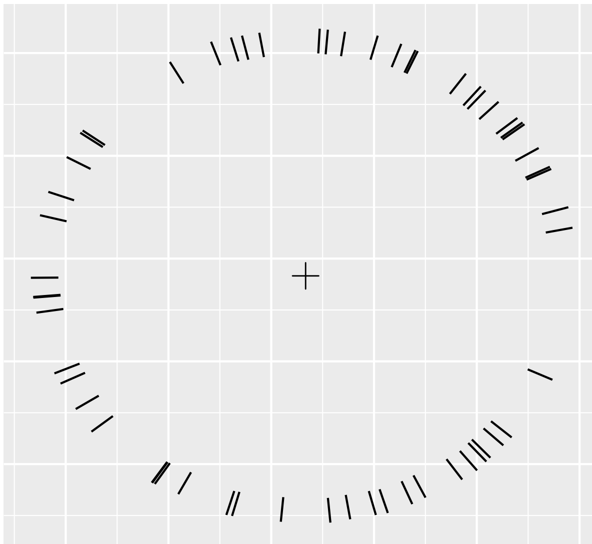
In the second round of the study, the only change in the treatment assignment mechanism is related to the "Control → Treated" branches. In those branches,

[24]"∧" is the logical *and* operator.

Geographic distribution of villages



Projection of villages to a circle



Final treatment assignment

Black filled points show the selection of villages to record (survey).

Figure 3: Example of within branch treatment assignment

1. We anchored the incentivization wedge on the village that was a *pure-control* village in the first study round.
2. The previous *pure-control* village was switched to an *incentivized* village.
3. The new *spillover* village was randomly selected such that the previously *pure-control* village is now within the incentivization wedge of the branch.
4. The determination of the incentivization wedge has changed a little in the second round of the study: instead of using the distances in $\{\mathbf{D}_k\}_k$, we used the rank or order of villages (in terms of distance). Thus instead of selecting the closest m villages to the *spillover*, we selected the $\frac{m}{2}$ villages that are closest to the *spillover* on either direction on the circle-distance (m being the number of villages to incentivize in the branch).

Appendix B: Recording Mechanism

$$p(W|Z^p, Z^s, Z^l, N^v, \{\mathbf{D}_k\}_k) = \prod_k p(W_k|Z^p_k, Z^s_k, Z^l_k, N^v_k, \mathbf{D}_k)$$

and

$$p(W_k|Z^p_k, Z^s_k, Z^l_k, N^v_k, \mathbf{D}_k) = \begin{cases} 1/N^v_k & \text{if } Z^p_k = 0 \wedge \|W_k\| = 1 \\ 1/\|Z^l_k\| & \text{if } Z^p_k = 1 \wedge W_k \cdot Z^s_k = 1 \wedge W_k \cdot Z^l_k = 1 \wedge W_k \cdot \texttt{spillover\_control}(Z^s_k, \mathbf{D}_k) = 1 \\ 0 & \text{otherwise} \end{cases}$$

W_k is the N^v_k-vector for village recording status within branch k and `spillover_control`(z, d) returns a vector mask indicating which villages are the "farthest" from the *spillover* village (Z^s_k).

The second case in the above recording mechanism above assigns the probability $1/\|Z^l_k\|$ when

i. k is a program branch
ii. The *spillover* village is recorded
iii. One random *incentivized* village is recorded
iv. The *branch-control* village is recorded

Details on how distance between villages is determined is in the experimental design section.

The first and second round of the study are similar in terms of recording, with the following exceptions:

a) Two *incentivized* villages in the "Control → Treated" branches will be surveyed.
b) Only 20 (out of 39) of the "Control → Treated" branches will have *spillover* and *branch-control* villages surveyed.

Appendix C: Implication of the Assignment and Recording Mechanisms (2017-2018)

Since the assignment and recording mechanisms rely on the circle-distance between villages in selecting *incentivized* and *branch-control* villages probabilities of assignment and recording vary within branches, as shown in Figure 4. Thus, while the experimental mechanisms are known, the study's design is only ignorable conditional on the circle-distance characteristics that affect assignment and recording probabilities. For that reason, we will condition our analysis on villages' propensity scores (probability of assignment and recording).

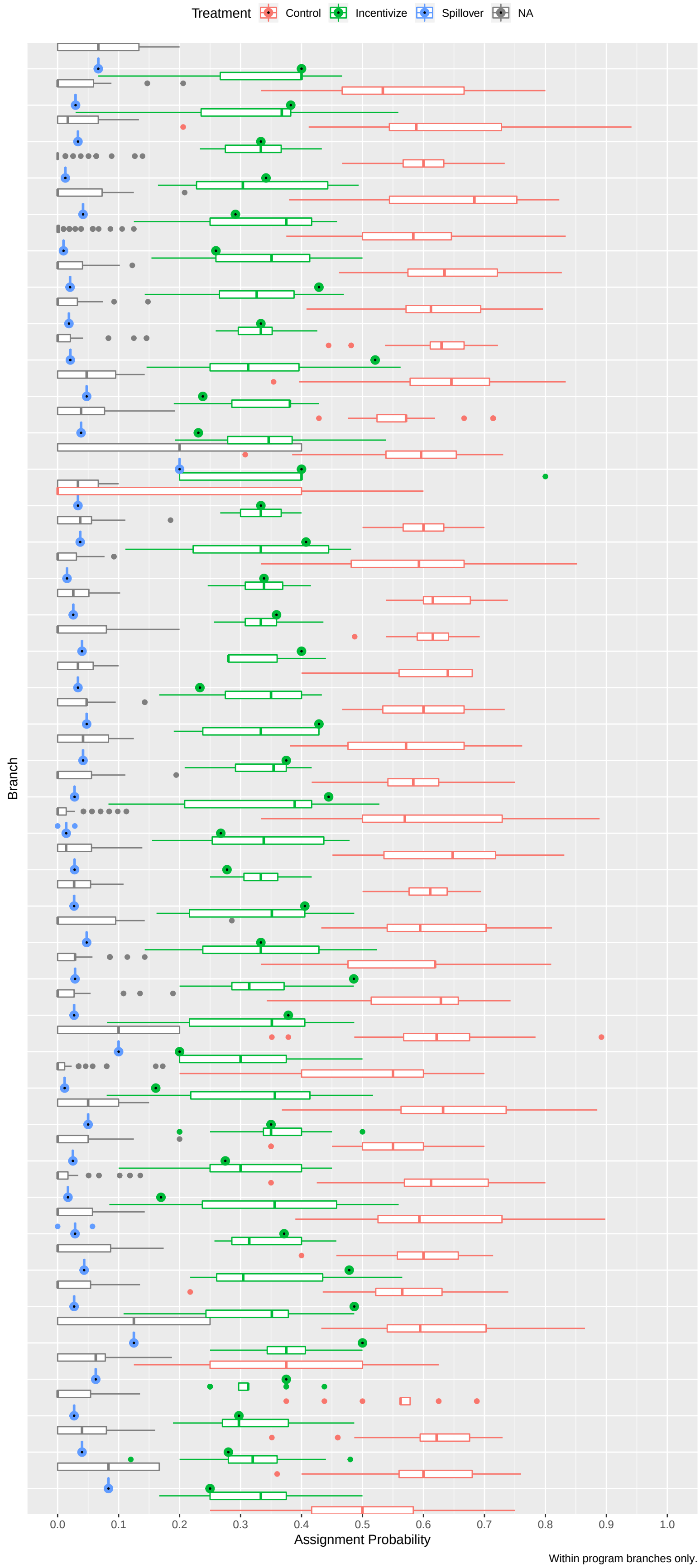


Figure 4: Treatment Assignment Probability

Appendix D: Bayesian Multilevel Causal Inference

We will build a multilevel/hierarchical Bayesian model (Imbens and Rubin 2015; Gelman et al. 2013) to estimate the posterior distribution of the estimands in the Primary Impact Estimands section. This allows us to study the heterogeneity in treatment effect between branches and villages (the experiment clusters). Such a model also enables us to address multiple hypothesis concerns (multiple treatments and outcomes) using regularizing priors, as well as use information more efficiently across model levels.

To calculate the finite sample posterior distribution of treatment effect, we will use Markov Chain Monte Carlo simulation (Carpenter et al. 2017) to estimate the prediction function

$$p(\mathbf{Y}^{mis}|\mathbf{Y}^{obs}, \widetilde{\mathbf{Z}}) = \int_\theta p(\mathbf{Y}^{mis}|\mathbf{Y}^{obs}, \widetilde{\mathbf{Z}}, \theta) \cdot p(\theta|\mathbf{Y}^{obs}, \widetilde{\mathbf{Z}})\, \mathrm{d}\theta \tag{6}$$

$$= \int_\theta \prod_j p(\mathbf{Y}_j^{mis}|\mathbf{Y}_j^{obs}, \widetilde{Z}_j, \theta_j) \cdot p(\theta|\mathbf{Y}^{obs}, \widetilde{\mathbf{Z}})\, \mathrm{d}\theta \tag{7}$$

where θ is a vector of all model hyper parameters and hierarchical parameters; $\mathbf{Y}^{mis}$ is a matrix of unobservable outcomes (columns) for all observations (rows); $\mathbf{Y}^{obs}$ is a matrix of observed outcomes (columns) for all observations (rows); and $\widetilde{\mathbf{Z}}$ is a vector of assigned treatments. This estimated distribution will allow us to make multiple imputations of unobservable counterfactuals. Thus we will be able to calculate the finite sample posterior distribution of treatment effects, $Y_{it}(z) - Y_{it}(z')$, for any treatments z, z'.

Posterior probability functions for all parameters will be estimated using observed data and prior probabilities. Model levels will be comprised of villages (indexed by j), branches (indexed by k), and subdistricts (indexed m).

The posterior probability of the model parameters is

$$p(\theta|\mathbf{Y}^{obs}, \widetilde{\mathbf{Z}}) = \prod_j p(\theta_j|\mathbf{Y}^{obs}, \widetilde{\mathbf{Z}}, \theta_{k[j]}) \prod_k p(\theta_k|\mathbf{Y}^{obs}, \widetilde{\mathbf{Z}}, \theta_{m[k]}) \prod_m p(\theta_m|\mathbf{Y}^{obs}, \widetilde{\mathbf{Z}}, \theta_o) \cdot p(\theta_o|\mathbf{Y}^{obs}, \widetilde{\mathbf{Z}}), \tag{8}$$

where model parameters are linear (and generalized linear) location, scale and correlation parameters for each level (parameters θ_o are the top-level model hyper parameters). Such model parameters will also allow us to calculate finite population and super-population estimands (the same as we do above with finite sample estimands).

References

Akram, Agha Ali, Shyama Chowdhury, and Ahmed Mushfiq Mobarak. 2017. "Effects of Emigration on Rural Labor Markets." https://doi.org/10.3386/w23929.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. 2014. "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh." *Econometrica* 82 (5): 1–43. https://doi.org/10.3982/ECTA10489.

Carpenter, Bob, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. "Stan : A Probabilistic Programming Language." *Journal of Statistical Software* 76 (1). https://doi.org/10.18637/jss.v076.i01.

Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2013. *Bayesian Data Analysis*. 3rd ed. Chapman; Hall/CRC.

Imbens, Guido W., and Donald B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. 1st ed. Cambridge: Cambridge University Press.

List, John A., Azeem M. Shaikh, and Yang Xu. 2016. "Multiple Hypothesis Testing in Experimental Economics." NBER. https://doi.org/10.3386/w21875.

Westfall, Peter H., and S. Stanley Young. 1993. *Resampling-Based Multiple-Testing: Examples and Methods for p-Value Adjustment*. New York: Wiley.

Young, Alwyn. 2017. "Channelling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results."