

# Delegation Risk and Implementation at Scale: Evidence from a Migration Loan Program in Bangladesh

Harrison Mitchell<sup>1</sup>, A. Mushfiq Mobarak<sup>2</sup>, Karim Naguib , Maira Emy Reimão<sup>3</sup>, and  
Ashish Shenoy<sup>4\*</sup>

<sup>1</sup>University of California, San Diego

<sup>2</sup>Yale University

<sup>3</sup>Villanova University

<sup>4</sup>University of California, Davis

\*Corresponding author: shenoy@ucdavis.edu

This version: December 2025

## Abstract

Many economic policies show promising pilot results but fail to replicate at scale. We evaluate a large-scale migration loan program reaching 150,000 households in northern Bangladesh following multiple successful pilots. Administrative data on loan take-up and subsequent migration closely match the pilots, but experimental evaluation reveals the large-scale implementation only raised migration by 12p.p. above control, far below the 25–40p.p. achieved in piloting. To reconcile this difference, we introduce a new theory of participant response to economic policy that distinguishes implementation compliance (loan acceptance) from impact (migration). Policy loses effectiveness when pilot benefits accrue to program compliers (those induced to migrate with a loan) but capacity constraints lead implementation to prioritize other loan recipients at scale, a possibility we call "delegation risk". We present evidence consistent with this theory that loans to induced migrants were crowded out by those that accepted a loan and did not migrate. Mis-targeting lowers program effectiveness by 15–18% over capacity constraints alone and can account for the diminished impact. We also investigate scaling concerns raised by expansion to new geographic areas, changes in the beneficiary population over time, and crowd-out of migration opportunities. Delegation risk has the potential to undermine many development policies, and may be exacerbated by management practices common at development organizations.

**JEL Codes:** O18, C93, J43, R23

## 1 Introduction

Evidence-based innovation in the development sector follows a common pattern: new ideas are piloted at small scale, and both governments and NGOs fundraise to scale up successful trials. This development strategy creates rising demand for rigorous evidence of effectiveness from governments and policymakers, institutional funders such as the Gates Foundation and Open Philanthropy, and small-scale donors coordinated through rating agencies such as GiveWell. However, recent evidence has uncovered many cases of promising pilots that have failed to replicate at scale (see [List, 2022](#)). Failure to scale has been documented in education ([Kraft et al., 2018](#); [Bold et al., 2018](#); [Andersen and Hvidman, 2020](#); [Evans and Yuan, 2020](#); [Ganimian, 2020](#); [Kerwin and Thornton, 2021](#); [Bellés-Obrero and Lombardi, 2022](#)), public health ([Global Innovations Fund, 2018](#); [Cameron et al., 2019](#)), early childhood intervention ([Araujo et al., 2021](#); [Bloem and Wydick, 2023](#)), microfinance ([Giné et al., 2021](#)), and behavioral nudges ([DellaVigna and Linos, 2022](#); [Rabb et al., 2022](#)). In fact, [Vivaldi \(2020\)](#) reports a systematic negative relationship between a program's scale and the size of its impacts across a broad range of development policies.

In this paper, we identify a new threat to scaling that arises when implementation decisions are delegated to field staff within the implementing organization. We study a migration loan program, an intervention motivated by the prevalence of seasonal poverty in rural northern Bangladesh, where a combination of low labor demand and low food supply in September–December preceding the main rice harvest lead to sharp rises in food insecurity ([Khandker, 2012](#)). Many households in the region supplement income by sending a member to work elsewhere in the country during this period. The No Lean Season (NLS) loan program was designed to facilitate this option for those that lack the liquidity to finance such migration. The program offered low-interest, short-term loans to landless rural households ahead of the lean season to cover the cost of transportation and a few nights' lodging for one household member.

NLS loans generated large, positive returns in two independent rounds of pilot evaluation. In randomized trials conducted in 2008 and 2014 with 1,292 and 5,764 households, respectively, loan offers raised the fraction of households who sent a migrant during the lean season by 25–40 percentage points from a baseline rate of around 35 percent. Local average treatment effect (LATE) estimates of the return to migration indicate households enabled to migrate earn nearly 50% more over the following months, and this income translates to consuming roughly 600 more calories per person per day during the lean season ([Bryan et al., 2014](#); [Akram et al., 2017](#)).

Motivated by these promising results, the NLS loan program was implemented at large scale in 2017 and 2018. A local microfinance institution (MFI) offered loans to over 140,000 households—reaching about 5% of the total population in the region—in each of these two years. This effort was independently evaluated by our research team. Researchers only intervened in implementation to select the regions of operation to allow randomized evaluation, and otherwise maintained minimal contact with the MFI to accurately measure impacts of the program as it would be administered

by local implementers.

The first contribution of this paper is to report evidence that the transformative impacts achieved in multiple rounds of piloting did not replicate at scale. Loss of effectiveness can be attributed to the large-scale program enabling far less migration—only a 12p.p. increase in the pilot districts in 2018, and far less in other districts and in 2017. We can easily reject the 25–40p.p. increases that had been observed during the piloting phase. Notably, the loss of program effectiveness in enabling migration is not evident in the administrative data from the MFI tracking loan acceptance and subsequent migration among the treated population. These implementation metrics suggest that the ‘at-scale’ program performance is almost exactly the same as the pilots. The program would have looked successful if not for experimental comparison to an untreated control population during the 2017 and 2018 rounds of intervention.

The second contribution of this paper is to explain the discrepancy between implementation metrics within the treated population and program impact in comparison to a control group. We introduce a model that draws a distinction between policy implementation—in this case, disbursement of migration loans—and induced behavior—in this case, migration. Gains in earnings and consumption primarily accrue to those who cannot finance migration on their own but are enabled to migrate when they receive a loan. For them, policy implementation induces the target behavior; in the language of program evaluation these are policy compliers. In contrast, those who do not change their migration behavior in response to the policy, both never-takers and always-takers, derive little benefit from the loan offer. However, among both of these groups are households that accept a migration loan when offered, and thus can appear to be implementation successes. The inability to distinguish induced migrants from these other types that accept a loan lead implementation metrics to diverge from program impact.

Mismatch between implementation and impact can weaken effectiveness as a program expands when discretion is delegated to implementing agents. In the case of NLS, scaling up necessitated hiring a large number of new loan officers, so there was less oversight per officer relative to pilot, and each officer was assigned a much larger caseload of target households. As a result, loan officers had greater discretion regarding where to focus recruitment and marketing, and managers used number of loans awarded as a metric to evaluate officer effort. These conditions together create incentives for program implementers to direct effort toward those most likely to be implementation successes. If households who accept a loan without changing their migration behavior are the easiest to reach—for instance because they can signal their demand faster than induced migrants who have to first make migration plans, or because loan officers can easily identify households already intending to migrate—then implementer effort may systematically deprioritize compliers who benefit from the program most. We refer to the potential misallocation of implementer outreach as “delegation risk”, and it poses a threat to scaling independent of specific policy design features or implementer quality.

Many categories of development policy contain the necessary conditions for the risk that delega-

tion of authority undermines effectiveness as a program expands in scale. In general, implementation metrics may not reflect program impact whenever benefits derive from an induced behavior, there are high fractions of always- and never-takers in the population, and compliance status is not readily observed. Delegation risk is present when resource or capacity constraints necessitate selectivity in treatment or outreach intensity as a program grows, creating conditions where implementers may systematically mis-target their efforts. In addition to directed lending such as NLS, these factors describe microfinance more generally<sup>1</sup>, conditional cash transfers, occupational training, agricultural extension, and a number of other common policies designed to enable or encourage a specific economic activity. It is especially concerning because implementation metrics (e.g. number of loans disbursed or number of workers/farmers trained) are commonly reported as evidence of impact in these sectors.

The third contribution of this paper is to present quantitative evidence that delegation risk lowered NLS program effectiveness at scale. We first establish that loan officers lacked the capacity to consistently reach all households assigned to treatment. A debrief survey among households assigned to treatment reveals that only just over half remember being offered a loan. Recall propensity is correlated with factors related to the cost of outreach such as distance to MFI branch and village caseload, suggesting that this measure carries signal about loan officer outreach intensity.

Next, we show that the decline in treatment effect is consistent with selective focus on households who take a loan but do not change their migration behavior, to the exclusion of induced migrants who would have been enabled to migrate with a loan. To do so, we hold the population distribution of compliance status—the fractions of loan acceptance and migration conditional on loan availability—constant at their pilot levels. Given the evaluation data on loan acceptance and migration, loan officer outreach to induced migrants in 2018 would have to have been 80% as strong as outreach to other households that accept a loan in order to generate the observed treatment effect.

In a separate exercise, we predict loan recall using household baseline characteristics, and then downscale the 2008 estimated treatment effect by households' predicted recall frequency. We find that, had households in 2008 been reached with the same heterogeneous intensity as in 2018, we would have expected the pilot to generate only a 6.3p.p. increase in migration. Selectivity in outreach intensity contributes 15–18% to this treatment effect attenuation, and the overall magnitude suggests delegation risk alone is sufficient to explain the full loss in effectiveness from pilot to at-scale implementation.

The final contribution of this paper is to explore other commonly recognized threats to external validity. Implementation at scale incorporated both the two districts where the pilot evaluations took place as well as six adjacent districts with similar patterns of lean season deprivation and food insecurity. We find that in addition to delegation risk, geographic heterogeneity further

---

<sup>1</sup>We present several prominent examples of major microfinance institutions that report service delivery as a measure of impact in Appendix A.

limits the external validity of the NLS pilot. The 12p.p. increase in migration at scale was only observed in the two districts of the original pilot; NLS had no effect on migration in the six expansion districts. While site-specific heterogeneity is commonly observed in development policy (e.g. Pritchett and Sandefur, 2015; Meager, 2019), we find this difference surprising in our setting due to the fact that expansion districts are geographically adjacent and the treated population is observably similar in baseline characteristics. There was no indication from (lack of) differences in income, education, household size, or migration history among eligible households that treatment effects may be diminished in new districts. However, site-based attenuation is consistent with other literature suggesting initial policy implementation frequently targets populations with the greatest propensity for success (Allcott, 2015; DellaVigna and Linos, 2022; List, 2022).

We rule out two alternative explanations for why NLS treatment effects may be diminished in this evaluation round. First, our study design includes a novel clustered randomization strategy to enable measurement of spillovers across villages in treated areas. In principle, large-scale policies may shift market prices or other macroeconomic conditions in ways that alter the policy impact (e.g. Cunha et al., 2018; Sraer and Thesmar, 2018; Egger et al., 2022; Khanna, 2023). In our study, this concern manifests as the possibility that mass migration lowers wages at the destination, limits the number of available migration opportunities, or otherwise crowds out other would-be migrants. To the contrary, we find that adjacency to treated villages has a positive effect on migration. The positive spillover effect only appears in the original pilot districts in 2018 where treatment effects are the strongest, and its strength increases with proximity to treated villages. These facts indicate local migration opportunities are not exhausted, destination labor demand remains sufficiently elastic, and complementarities generated by traveling together overwhelm any general equilibrium crowd-out at this scale of operation.

Second, we find no evidence that the effect of the NLS program diminished due to changes in population characteristics over time. Rosenzweig and Udry (2019) demonstrate that the effects of economic policy can be sensitive to changes in the macroeconomic environment. We evaluate the potential effect of differential rainfall, changes in population wealth, and other changes by using machine learning to estimate treatment effect heterogeneity by baseline characteristics in the pilot experiments. We then simulate the treatment effect we would have observed at scale had conditional treatment effects remained constant and only the distribution of population characteristics changed over time. This exercise generates counterfactual results very similar to the actual effect measured in pilot rounds, indicating that changes in observable population characteristics over time cannot account for the decline in program effectiveness.

This paper contributes to the small but growing number of randomized evaluations of large-scale development programs. Other recent evaluations include conditional cash transfers (Schultz, 2004; Rivera et al., 2004), public assistance (Muralidharan et al., 2016, 2017; Cunha et al., 2018; Egger et al., 2022), education policy (de Ree et al., 2017; Bold et al., 2018; Khanna, 2023), police reform (Banerjee et al., 2021), and celebrity messaging (Alatas et al., 2019), though most of this literature

does not have the benefit of data progressing from a pilot RCT to large-scale evaluation as we do here. We demonstrate the value of such large-scale evaluation by quantifying how treatment effects differ across evaluations of the same program implemented at different scales. Linking these various evaluations allows us to quantify the importance of sources of change from pilot to scale. Economic explanations for the relationship between effectiveness and scale typically center around changes in program goals, changes in the nature of the intervention as it is implemented, changes in participant population or context, or general equilibrium spillovers (see Banerjee et al., 2017; Al-Ubaydli et al., 2017, 2019; Brooks et al., 2020).

Delegation risk can be seen as a counterpart to endogenous participant response. Chassang et al. (2012); Bulte et al. (2014) establish a relationship between program effectiveness and (often unobserved) effort taken by program participants. Such effort may be motivated by knowledge of the program and beliefs about its effectiveness. In this study, we highlight the role of implementers in disseminating such knowledge and facilitating take-up, and present evidence of one channel through which their efforts may miss the intended beneficiaries.

Mistargeting in our setting arises from the inability to base the incentives for implementing agents directly on treatment of intended program beneficiaries (i.e., compliers). This concern arises in a number of other settings where implementer incentives have been shown to affect policy outcomes. Empirical studies show the importance of incentive design in banking (Hertzberg et al., 2010), health worker attendance (Dhaliwal and Hanna, 2017), environmental audits (Duflo et al., 2013), and tax collection (Khan et al., 2015; Balán et al., 2022). In each of these cases, the implementing agent mediates an adversarial relationship between the policymaker and the program participant. For instance, a tax collector balances the government's interest in raising revenue against the taxpayer's interest in minimizing payment. Our study extends this investigation to a setting in which policymakers' and participants' interests are perfectly aligned: both want to enable the migration of poor households during the lean season. We show that even with this alignment of interests, delegation of authority to constrained implementing agents can undermine policy goals.

Finally, this paper sheds light on the household response to seasonal liquidity constraints. In rural areas around the world, seasonality in agriculture creates regular periods of economic distress that lead households to take costly actions to satisfy immediate consumption needs. Seasonal constraints have been shown to limit agricultural investment (e.g. Duflo et al., 2011), food storage and sales (Stephens and Barrett, 2011; Basu and Wong, 2015; Burke et al., 2018), and labor market activity (Bryan et al., 2014; Akram et al., 2017; Fink et al., 2020). We show that the mere availability of credit for migration as a coping strategy is not sufficient to alleviate these constraints as lack of concerted outreach can substantially limit uptake.

## 2 Background and Data

### 2.1 Setting

This study takes place in rural parts of the Rangpur Division in northern Bangladesh. This is a poor and largely agrarian part of the country, with an urbanization rate under 15%. Among the rural population of Rangpur, 48% of households were classified as moderately or extremely poor in 2016, compared to 26% for the country overall ([Bangladesh Bureau of Statistics, 2017](#)). A map of the region of study is provided in Appendix Figure S2. Pilot evaluation rounds included villages randomly selected from the Kurigram and Lalmonirhat districts, and the at-scale implementation expanded to include these as well as the six other districts in the division. We report results separately for pilot and for expansion districts.

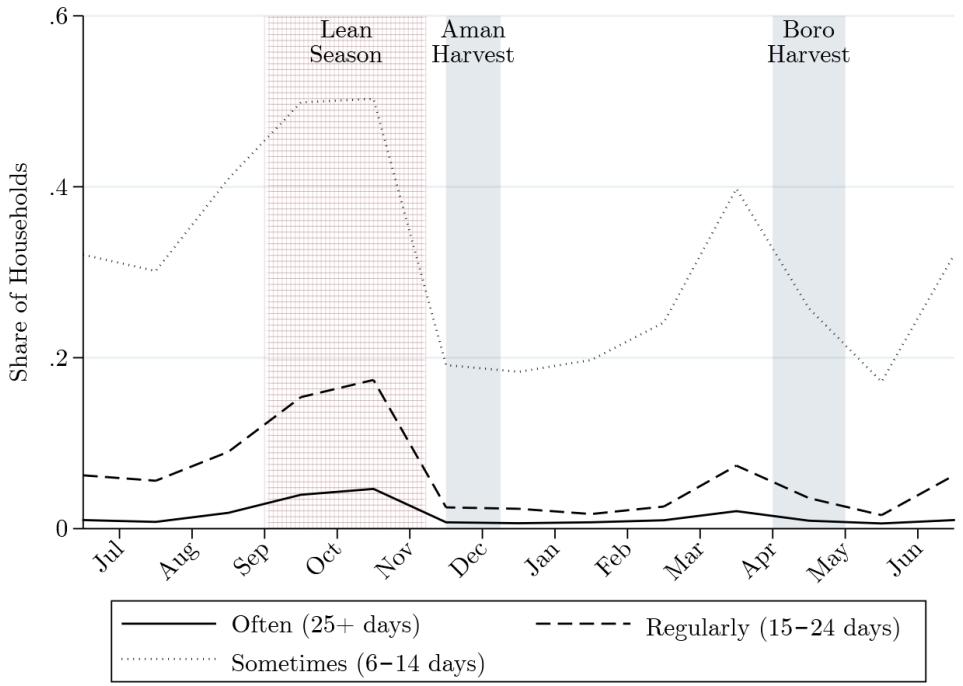
Rural economies in this area are characterized by strong seasonality tied to the agricultural cycle. The primary rice crop season lasts from planting in June through harvest in December. Labor demand, and correspondingly the agricultural wage, peaks at these endpoints but remains low between the start and end of the season. Low wages are accompanied by high food prices toward the end of the season as stocks dwindle ahead of the harvest ([Khandker, 2012](#)).

Low wages and high prices combine to create an annual period of heightened food insecurity during the agricultural lean season. September through early November in this part of the country is locally referred to as “monga”, a phenomenon known as the “hungry season” in many parts of the world. As shown in Figure 1, more than half of landless rural households surveyed report reducing meals or portion sizes during this period, and nearly a fifth do so for more than fifteen days per month. The fact that this level of deprivation occurs with annual regularity indicates that many vulnerable households have little capacity to use savings or other methods to smooth consumption over the year. Rural landless households comprise half to two-thirds of all households in the study region, so seasonal hunger is a widespread phenomenon. At the same time, formal coping support through government or nongovernmental organizations is limited. Among poor households, the vast majority of which experience some food deprivation during the period, only 32% report receiving support from these sources as part of their coping strategy ([Khandker and Mahmud, 2012](#)).

Many households turn to short-term, intra-national migration as a response to agricultural seasonality during this period. At baseline, a third to half of landless households in our region of study use migration to supplement earnings during the lean season months. The typical migration episode involves one household member, almost always male, traveling for work to destinations within Bangladesh. The typical migration episode lasts for 2–3 months, perhaps with one or two visits home in the interval. Migrants generally bringing their earnings back in cash. Comparably high rates of lean season migration can be observed in many rural parts of the world.

Agriculture is the most common destination sector of work for seasonal migrants from the Rangpur region. Roughly half of seasonal migrants find agricultural work in other parts of the country where the planting and harvest seasons are offset due to climate. Among the other half

Figure 1: Food Insecurity among Landless Rural Households



Notes: Self-reported monthly frequency of reducing meals or portion sizes from a 12-month recall survey.

that travel to urban destinations, a third work in the transportation sector (i.e. pulling cycle rickshaws) and another third find employment in low-skill construction, both sectors that are far less sensitive to agricultural seasonality. The capital city of Dhaka accounts for nearly 30% of migration to urban areas, with the rest spread throughout other cities around the country.

## 2.2 No Lean Season Migration Loan Program

While migration during the agricultural lean season is fairly common in the Rangpur region, many households are unable to access this option due to liquidity constraints. The issue arises because migration requires up-front financing—for transportation and initial lodging—to realize a subsequent stream of labor market returns. This financing requirement comes at exactly the time of year when households that rely on agricultural labor and their local social network have the least available cash on hand. The challenge of migration is compounded by the fact that those who stand to benefit the most from migration fall closest to the threshold for subsistence. Therefore, even though they have the highest marginal return to increasing consumption, they also face the greatest risk from financing a migration attempt that turns out to be unsuccessful.

The No Lean Season (NLS) program aims to enable seasonal migration for a larger fraction of the rural landless population through access to liquidity. NLS offers a short-term, zero-interest loan of BDT 1,000 (around \$12 USD) for migration during the agricultural lean season. This value

is sufficient to cover bus fare and a few nights' lodging for one household member establish in a destination labor market. Loans are offered in early September at the start of the agricultural lean season, with a duration of 3–6 months for repayment.

Eligibility for NLS loans is based on exposure to the agricultural lean season. Households are eligible if they own under 0.5 acres of land, meaning their primary earnings must come from the labor market, and they self-report having reduced or missed meals in prior lean seasons. Roughly sixty percent of rural households satisfy these eligibility criteria in our region of study.

NLS loans are issued with soft conditionality. The loans are marketed as intended for the purpose of migration, and, to receive the loan, recipients had to state their destination and employment plan. Upon disbursement, loan officers follow up with recipients to inquire about migration status and encourage travel for those that have not yet departed. However, there is no penalty for failure to migrate nor any other formal enforcement of the migration requirement.

### 2.3 Pilot Evidence of Program Success

The NLS program was validated in two independent rounds of piloting with randomized evaluation. The initial pilot to establish viability of migration loans took place in 2008. In this round, 1,292 households in 68 villages were offered migration loans<sup>2</sup>, out of a total study population of 1,826. In villages where loan offers were made, offers comprised on average 14% of the eligible population, reaching 19 households per village. Eligibility surveys and outreach to participants were conducted by the same evaluation team that collected survey data on outcomes, as is typical in pilot studies, and disbursement was made in close partnership with local microfinance organizations.

In this round of piloting, lean season migration was 22p.p. higher among households offered a migration loan relative to control. Induced migration generated large benefits for participant households, with LATE estimates indicating those enabled to send a migrant increased their lean-season food consumption by 550–700 calories per person per day. These migration episodes proved to be so beneficial that migration remained 8–10p.p. greater among treated households three years after the pilot had ended. This evaluation also included an information-only treatment arm where participants were told about types of jobs available at potential migration destinations, but which ultimately had little impact. Full details and evaluation results are reported by [Bryan et al. \(2014\)](#).

The second pilot evaluation round in 2014 maintained close to the same number of villages, but increased treatment intensity within village. In this round, loans were offered to 5,764 households in 95 villages with a comparably sized control group for comparison. Within villages where offers were made, treatment intensity was randomly assigned to be either 14% or 70% of eligible households; this variation was introduced to investigate the possibility increasing loan offer saturation would lead to within-village spillovers that crowded out program effectiveness. Loan eligibility was again determined by the evaluation team, and loan offers were made by a local microfinance organization

---

<sup>2</sup>589 households were offered conditional loans and 703 offered conditional grants, but the treatment effect was nearly identical between these two groups.

with close coordination and oversight from the evaluation team.

Loan offers in low-intensity villages increased migration by 25p.p. in 2014, closely matching results from 2008. In high-intensity villages, induced migration contributed an even larger 40p.p. In these villages, lean season migration eased labor market pressure, raising the wage rate relative to control and thereby lowering the returns to migration. However, this effect was more than offset by migrants traveling together and sharing fixed costs of housing and job search, leading to net crowd-in rather than crowd-out of new migrants. LATE estimates from this evaluation round indicate that induced migration raised household earnings by 50–60% during the lean season, with drops in food insecurity comparable to those found in 2008. Full details and results are reported by [Akram et al. \(2017\)](#).

These two successful pilots, which consistently demonstrate transformative returns to induced migration across different measures of household wellbeing, motivated the program to be implemented at scale by a local MFI operating throughout the region.

## 2.4 NLS Implementation at Scale

In the current paper, we present evaluation results from NLS implementation at scale that took place in 2017 and 2018.<sup>3</sup> This round differs from prior pilots in two key ways. First, the scope of the program was greatly expanded, with 158,014 loan-eligible households in 734 villages in 2017 and 143,721 loan-eligible households in 2018, comprising just over 5% of rural households in the region each year. Within each village where loan offers were made, all eligible households were offered a migration loan. Second, implementation by the local microfinance organization was decoupled from evaluation. Implementers coordinated with evaluators only in identifying villages where loan offers would be made and setting thresholds for household eligibility.<sup>4</sup> Once these design decisions were in place, the implementing organization conducted its own survey to determine household eligibility and make loan offers, and the evaluation team independently identified and surveyed a sample of households for evaluation.

Implementation at scale was conducted by a local MFI with 110 branch offices spread throughout the Rangpur division.<sup>5</sup> We define each branch's catchment area as the set of villages within a one-hour bike ride from the branch office. The MFI provided a census of all villages meeting that criterion, which serves as the sampling frame for both random assignment of treatment as well as for evaluation data collection.

We implement a two-level randomization to designate villages in which the MFI would make the loan offers. At the first level, we randomly assign microfinance branch offices into a treatment group that makes loan offers and a control group that does not. 40 branch offices were assigned

---

<sup>3</sup>This study was preregistered in the AEA RCT Registry under ID No. AEARCTR-0002685.

<sup>4</sup>Villages in which to make loan offers were selected by the research team to enable randomized evaluation, explained in detail next.

<sup>5</sup>The MFI was chosen as the implementing partner because it has a strong presence in the region and was involved in loan disbursements in the 2008 pilot study.

to treatment in 2017 and ten more were added in 2018, based on overall program implementation funding. Villages in the catchment area for control branches are designated as “pure control” with no loan offers made in or nearby. Randomization at this level was stratified by district to enable comparison of treatment effects between pilot and expansion districts.

At the second level of randomization, we select a subset of villages within each treated branch’s catchment area to make loan offers based on the individual branch’s loan capacity. We introduce a novel design to preserve random assignment while allowing us to distinguish between untreated villages highly exposed to nearby treatment and those further away. For this, we project villages in a MFI branch’s catchment area onto a circle with the branch office at the center. Then, we randomly select one village from this projection to be designated as the “spillover” village in which no offers will be made. Next, we assign an equal number of villages in either direction along the circle projection to the “treatment” group where loan offers are made. Between a third and a half of the villages in each branch’s catchment area are assigned to the treatment, according to anticipated branch capacity, and all eligible households in these villages are assigned to be offered a migration loan. Finally, the remaining villages within the branch’s catchment area are designated as “branch-control”, and no loan offers are made.

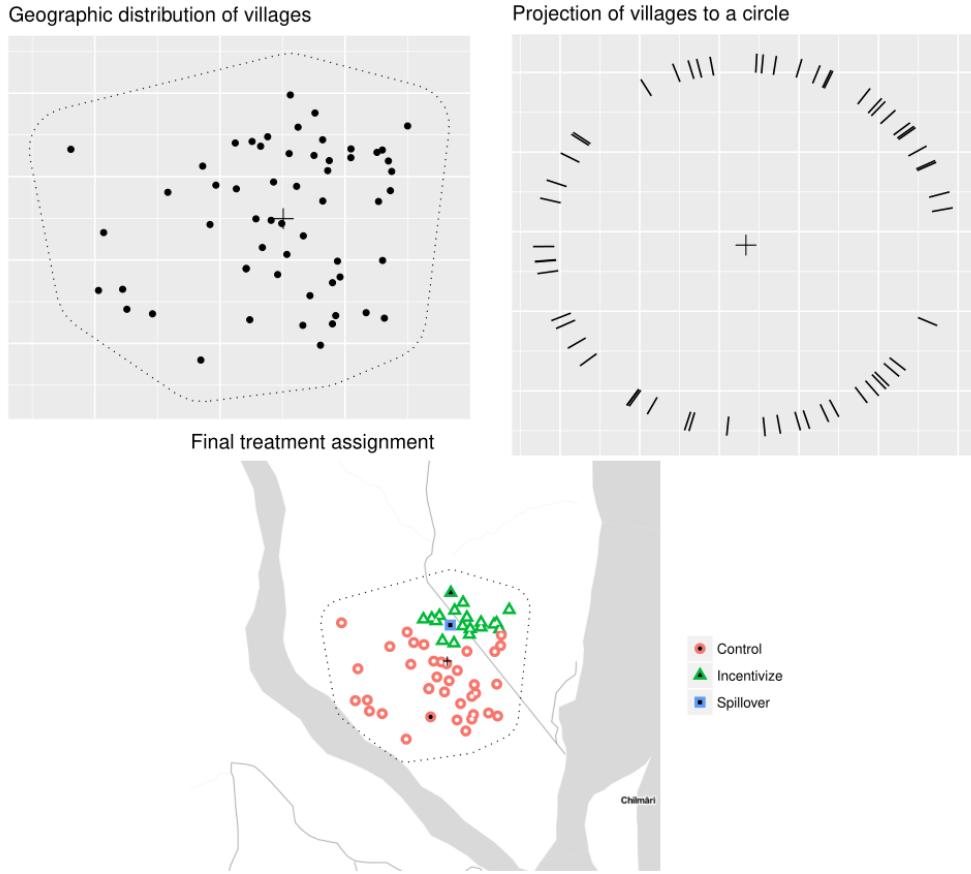
An example of the resulting village assignment is provided in Figure 2. This strategy effectively creates a pie-slice-shaped treated region originating from the branch office in the center of the catchment area. A single village close to the middle of the treated slice remains untreated to test for spillovers across villages. Because all assignment is randomized according to projected circle order, the probability of treatment remains uncorrelated with density of villages, proximity to the branch, or other geographic characteristics. Similarly, spillover villages are selected uniformly at random and therefore this assignment is uncorrelated with geographic density or proximity.

After randomization, each treatment branch office hired two new employees as migration loan officers to conduct outreach to eligible households and handle the loan portfolio for the branch. These officers first administered a census of households in each village assigned to treatment ahead of the lean season. The census included questions about land ownership and history of food security to determine loan eligibility. Loan officers then contacted households deemed eligible to advertise and promote migration loans. Actual loan disbursement took place at the branch office, and loans were promoted as migration loans. Any member of an eligible household could pick up the loan for their household, but loans were not disbursed to members of households either not in a village assigned to treatment or deemed ineligible through census.

## 2.5 Study Sample and Evaluation Data

Village-level randomization generates four types of villages: a treatment group where loan offers are made to all eligible households; a spillover group where no offers are made but the village is in the midst of a treated region; a branch-control group where no offers are made but the village is in the catchment area of a branch that makes offers; and a pure control group where no offers are made

Figure 2: Within-Branch Treatment Assignment According to Circle Order

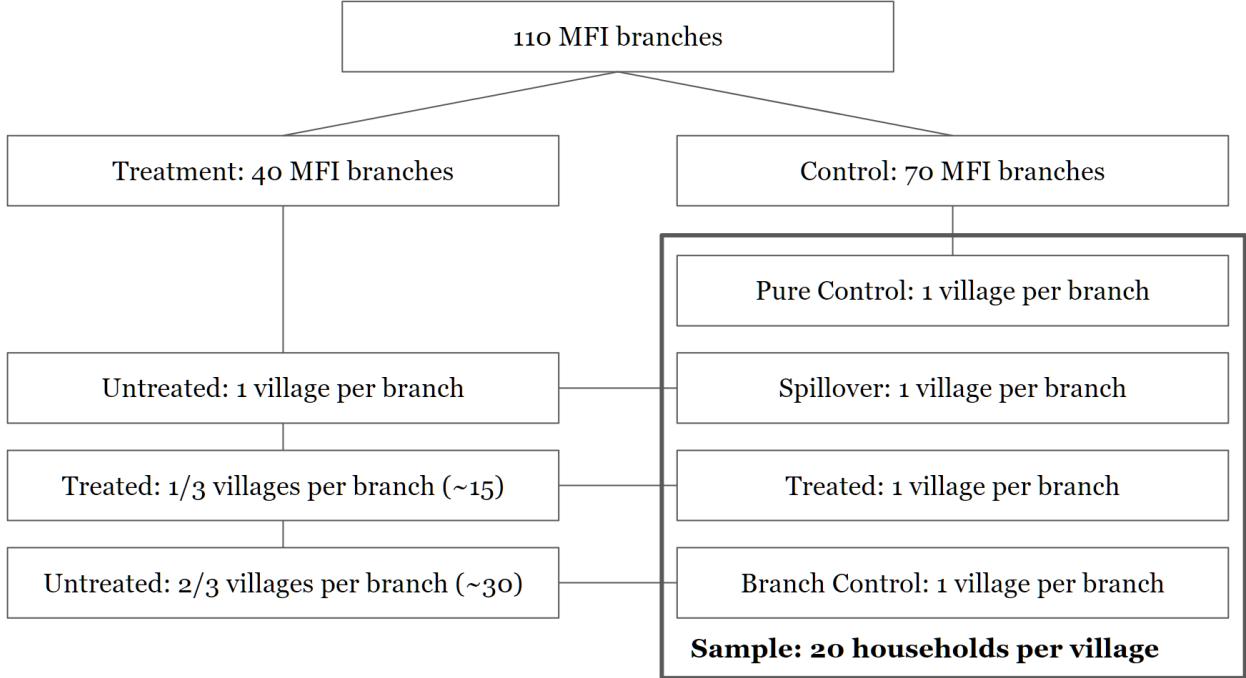


Notes: Example of village assignment in a treated branch according to circle order. Top left panel shows geographic distribution of villages in catchment area, with branch office represented by +. Top right panel shows circle projection of villages around branch office. Bottom panel shows resulting treatment assignment. Square represents randomly selected “spillover” village. Triangles represent “treated” villages according to circle order around spillover. Circles represent untreated villages designated as “branch control”. Randomization generates treated and untreated regions, with one untreated village in midst of treated region. Shaded shapes represent villages selected for evaluation surveys.

anywhere in the catchment area of the nearest branch. For evaluation purposes, we randomly select one treated and one branch-control village per treatment branch to conduct household surveys. We also conduct surveys in every spillover village as well as one randomly selected pure control village from each untreated branch. Within each survey village, we randomly sample twenty households out of the eligible population for surveying. Figure 3 characterizes the full randomization and survey strategy for 2017; the only adjustment in 2018 was to increase to 50 treatment branches.

To identify households that would be loan-eligible in untreated villages for surveying, the survey team used a random walk sampling strategy. Surveyors followed a skip pattern to select households, applied the same questions used to determine eligibility criteria in each selected household, and stopped once they had identified twenty would-be-eligible households. In 2017, survey households in treated villages were drawn randomly from the census conducted by the microfinance organization. This asymmetry raised a concern that differences between the treated and untreated survey samples

Figure 3: Village Randomization and Survey Assignment for 2017



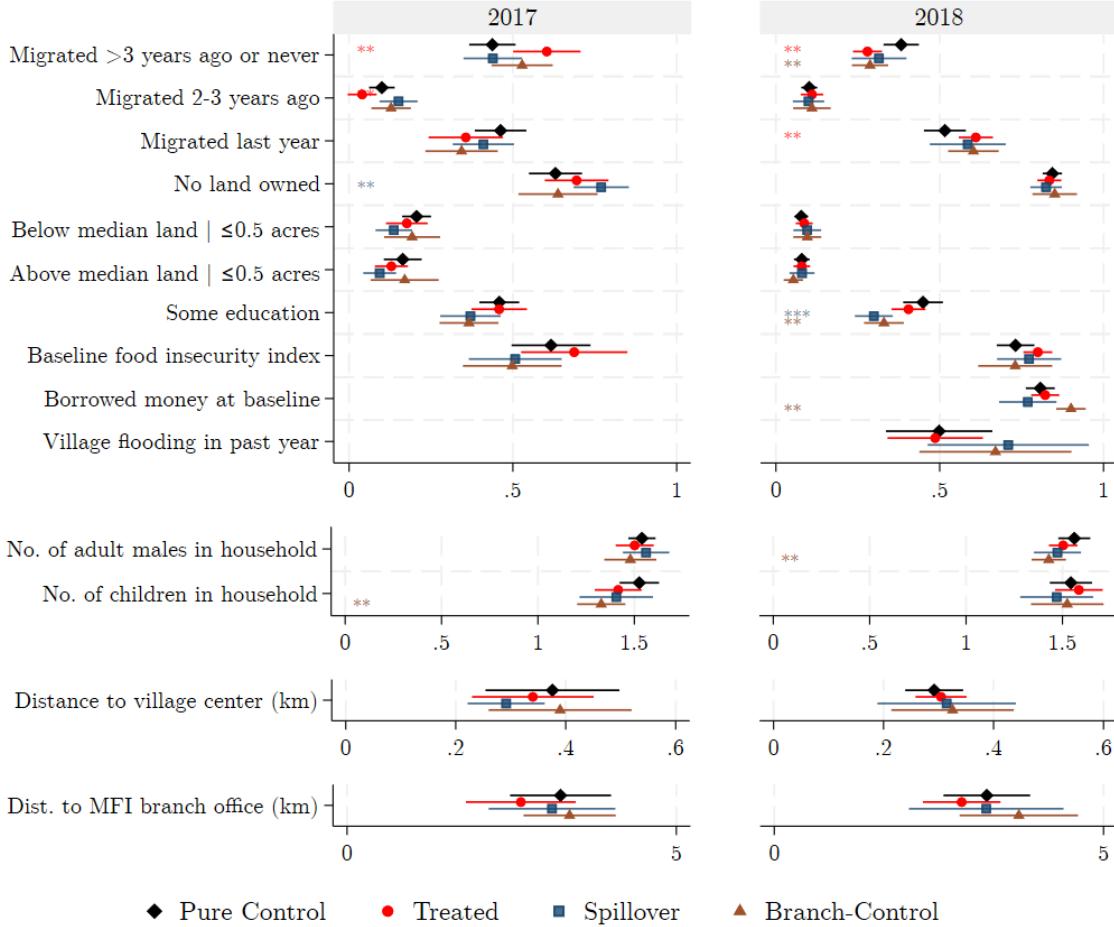
may have been induced by differential selection, especially if households responded differently to surveyors than to MFI officers. As a result, in 2018, sample households across all four village types were selected by surveyors through the same random walk strategy.

We conduct three surveys with each sample household in each evaluation year of the at-scale study. First, we administer a short survey on loan eligibility, migration history, and household composition in August–September prior to the lean season and any potential migration. Second we conduct a longer survey in the following January with questions about migration, earnings, and food consumption during the lean season. Finally, we conduct a third survey in April–May 2017 and June–August 2018 focused on subsequent migration, subsequent earnings, and overall financial status to evaluate whether lean season earnings may have persistent effects.

Figure 4 reports balance on baseline characteristics across treatment status in 2017 and 2018, with detailed point estimates reported in Appendix B. A joint F-test rejects balance across baseline household characteristics in 2017, when the method of sample selection differed between treated and untreated villages. In 2018, when sample selection was consistent across treatment assignment, we cannot reject equality between treatment and control jointly across all variables, though some individual imbalances remain. In all that follows, we verify that results are robust to controlling for baseline household covariates.

We supplement the primary analysis with three additional sources of data. First, we conduct a debriefing survey with migration loan officers following the 2017 round of evaluation. Second, we include recall questions on loan offers and microfinance institution activities following the 2018

Figure 4: Balance across Treatment Arms on Baseline Covariates



Notes: Variable means and standard errors for baseline household outcomes. Full numerical results are reported in Appendix B. Standard errors clustered at the village level. Stars at left side of each row indicate statistically significant deviations from control group. \*( $p < 0.10$ ), \*\*( $p < 0.05$ ), \*\*\*( $p < 0.01$ )

lean season. Third, we use administrative data on implementation from the microfinance institution itself. Table 1 provides details on the timing and data sources associated with each study round.

There was strict separation between implementation and evaluation, and survey teams did not provide feedback or monitoring to implementers as is common in pilot evaluations. In this study, the research was designed to evaluate the policy as it would normally operate without parallel research efforts. Because of this separation, we cannot match household survey data to administrative data at the household level. This prevents us from, e.g., using administrative data to validate recall about loan offers, or from using survey data to validate administrative loan and migration records.

Table 1: Study Rounds and Data Sources

| Year | Treatment Assignment   | Source              | Outcome Data   | N       |
|------|--|---------------------|--|---------|
| 2008 | Household randomization  | Household Survey    | Baseline; Loan; Migration; Food consumption                                    | 1,826   |
| 2014 | Village treatment intensity; Household randomization                 | Administrative      | Loan; Migration  | 11,369  |
|      |  | Household Survey    | Migration; Earnings  | 5,764   |
| 2017 | Branch randomization; Village randomization; All eligible households | Administrative      | Eligibility; Loan; Migration   | 158,014 |
|      |  | Loan Officer Survey | Incentive Structure  | 66      |
| 2018 | Branch randomization; Village randomization; All eligible households | Household Survey    | Baseline; Migration; Earnings; Food security                                   | 3,678   |
|      |  | Administrative      | Eligibility; Loan; Migration   | 143,721 |
|      |  | Household Survey    | Baseline; Loan; Offer Recall; Migration; Indebtedness; Earnings; Food security | 4,324   |

Notes: In 2017, eligibility screening data in the treatment villages was collected by the implementing partner and, in control villages, by the survey team. In 2018, eligibility screening data in both treatment and control villages was collected by the survey team. Administrative data from implementing partner cannot be linked to household surveys in 2017 and 2018.

### 3 Evaluation at Scale

We test for differences across treatment groups using the regression specification

$$Y_{ivs} = \sum_Z \beta^Z Z_v + \gamma X_i + \delta_s + \epsilon_{ivs} \quad (1)$$

where  $Y_{ivs}$  denotes an outcome  $Y$  for household  $i$  in village  $v$  and sub-district (upazila)  $s$ .<sup>6</sup>  $Z$  indicates the village treatment assignment for that household,  $X_i$  is a vector of household characteristics from Figure 4, and  $\delta_s$  controls for subdistrict fixed effects. The coefficients of interest  $\beta^Z$  report the difference in outcome between households in villages with treatment status  $Z$  and those assigned to pure control. We estimate effects separately for 2017 and 2018, and standard errors are clustered at the village level.

Our primary evaluation metric is the fraction of program-eligible households in which at least one member temporarily migrates during the study period. Estimated treatment effects across all rounds of program evaluation are presented in Table 2. The first two columns replicate results from the 2008 and 2014 pilot studies, respectively.<sup>7</sup> In isolation, being offered an NLS loan raises a household's propensity to send a migrant by 18–25 percentage points, from a base migration rate

<sup>6</sup>There are 58 subdistricts in Rangpur division. Subdistrict boundaries do not coincide with MFI branch catchment areas.

<sup>7</sup>Point estimates differ slightly from those reported in prior publications due to the inclusion of household controls and subdistrict fixed effects.

Table 2: Treatment Effect on Household Migration

|                                  | Evaluation Round |                |                |                |
|----------------------------------|------------------|----------------|----------------|----------------|
|                                  | 2008             | 2014           | 2017           | 2018           |
| Low Saturation                   | 0.18<br>(0.03)   | 0.25<br>(0.04) |                |                |
| High Saturation                  |                  | 0.40<br>(0.03) | 0.04<br>(0.05) | 0.12<br>(0.04) |
| $H_0 : \beta = \beta_{14, Low}$  |                  |                | 0.00           | 0.03           |
| $H_0 : \beta = \beta_{14, High}$ |                  |                | 0.00           | 0.00           |
| Control Mean                     | 0.36             | 0.34           | 0.40           | 0.39           |
| HH Controls                      | Yes              | No             | Yes            | Yes            |
| Subdistrict FEs                  | Yes              | Yes            | Yes            | Yes            |
| N                                | 1826             | 3600           | 1537           | 1901           |

Notes: Regression results from pilot and at-scale evaluation following (1). Outcome is a dummy for whether any member of the household migrated during the lean season. 2014 analysis excludes untreated households in treatment villages, and lacks household controls because baseline data on household characteristics is unavailable. 2017 and 2018 analysis restricts to treated and pure control villages in the original two districts of the pilot studies. Standard errors clustered at the village level.

around 35%. As the saturation of loan offers within a village increases from 14% to 70% of qualified households, reported in column 2, the fraction induced to migrate nearly doubles from 25% to 40%. This sizable increase indicates there is strong crowd-in of new migrants as the within-village loan offer and migration rates increase.

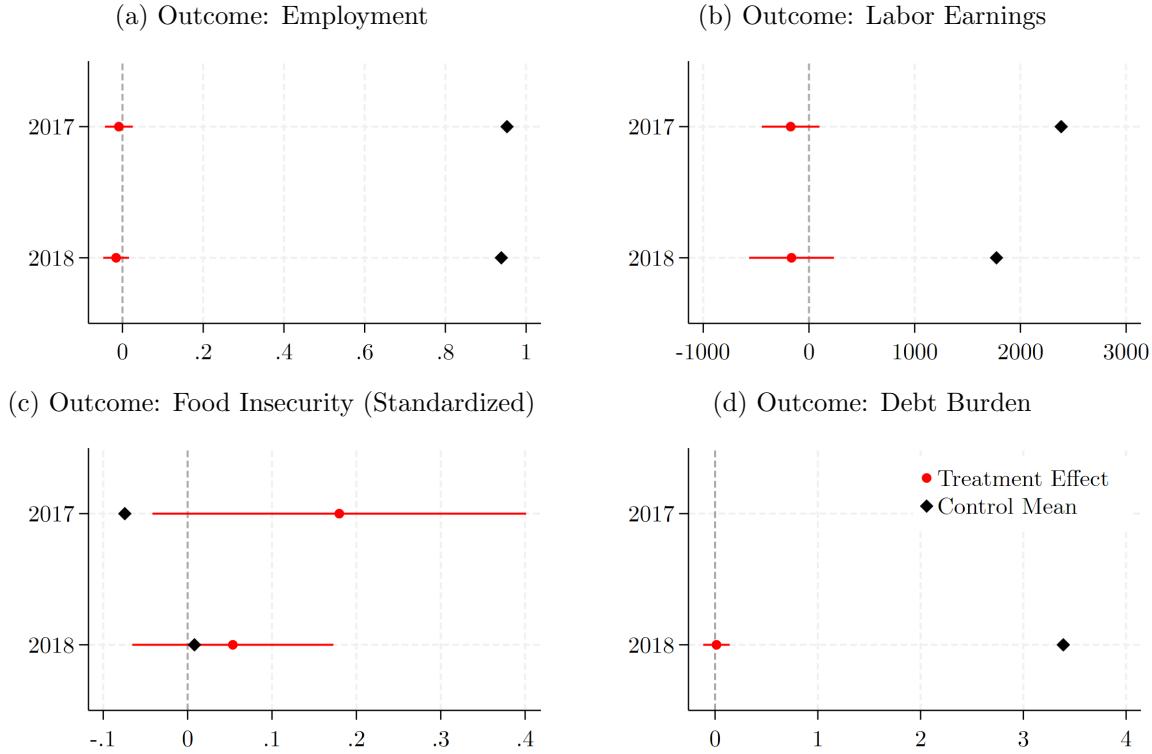
These large migration impacts fail to replicate at scale. The final two columns of Table 2 report estimated treatment effects from the 2017 and 2018 evaluation rounds comparing treatment to pure control in the districts where the original piloting took place.<sup>8</sup> In both years, the background rate of migration in control villages remains close to its pilot level. From this base, eligibility for an NLS loan only increases a household's likelihood of sending a migrant by 4 percentage points in 2017 despite full within-village saturation of loan eligibility, and this change is not statistically distinguishable from zero.<sup>9</sup> In 2018, the treatment effect climbs to 12 percentage points, still well short of the anticipated 40 percentage points from the 2014 high-saturation arm and even below the 2008 and 2014 low-saturation treatment effects. We can reject equality in treatment effect between every pilot and at-scale round at the 5% level except for 2008 and 2018.

Even without inducing new migration, it is possible that NLS benefited eligible households by offering a low-cost alternative to finance planned migration or other lean season activities. We investigate this possibility in Table 5, which reports the reduced-form effect of treatment on household earnings, consumption, and indebtedness. Results are summarized, with full results presented in Appendix C. Across a range of outcomes, reduced form effects are quantitatively small

<sup>8</sup>Analysis of treatment effects in expansion districts and spillover villages is presented in Section 5.

<sup>9</sup>In Appendix C we verify robustness to the way migration is measured.

Figure 5: Reduced Form Effects on Midline Household Outcomes



Notes: Regression results from at-scale evaluation following (1). Food insecurity is relative to a normalized index. Debt burden is self-reported concern about outstanding debt; only asked in 2018 survey. Full regression results reported in Appendix C. Standard errors clustered at village level; 95% confidence intervals depicted on graph.

relative to their mean value and statistically indistinguishable from zero. We find little evidence of a direct program impact on household wellbeing, indicating that the downstream results on earnings and consumption achieved in piloting primarily operated through NLS's role in facilitating migration.

It is worth noting that this failure to replicate at scale is not apparent in administrative data on program implementation. For comparison, we report implementation details in Table 3. The table reports the fraction of households lost at each stage of the NLS process. The first three columns describe the 2008 and 2014 pilot rounds, the fourth and fifth column the 2017 at-scale implementation, and the sixth column the 2018 implementation. The first three rows of Table 3 describe program scale and reach, and the next four rows report implementation details.

As evident from the first two rows, while the NLS program drastically expanded in scale from 2008 to 2018, similar fractions of the population in treated villages met the criteria for receiving a migration loan in 2008, 2014, and 2018. Not all eligible households were assigned to receive a loan offer the pilot rounds due to experimental randomization. This fraction is around one tenth in the 2008 pilot, and varies based on offer saturation in 2014. In 2017 and 2018, all eligible households in treatment villages were assigned to be offered a loan.

Table 3: Sources of Administrative Leakage from Loan Eligibility to Migration

|                                   | 2008   | 2014<br>Low | 2014<br>High | 2017<br>Full | 2017<br>Strict | 2018  |
|-----------------------------------|--------|-------------|--------------|--------------|----------------|-------|
| Population                        | 21,902 | 12,974      | 16,486       | 206,655      | 232,916        |       |
| % Eligible from Population        | 56.4   | 50.2        | 45.1         | 76.5         | 21.6           | 61.6  |
| % Offered from Eligible           | 9.8    | 13.7        | 65.0         | 100.0        | 100.0          | 100.0 |
| % Accepted Offer                  | 64.2   | 58.0        | 57.1         | 42.8         | 48.4           | 62.6  |
| % Received Loan from Accepted     | .      | .           | .            | 62.4         | 61.0           | 99.1  |
| % Migrated from Disbursed         | 65.9   | 66.3        | 64.9         | 91.3         | 91.5           | 70.0  |
| % Migrated with Loan from Offered | 42.3   | 38.4        | 37.1         | 23.3         | 25.7           | 41.0  |

Notes: Loan offers and migration from initial census population according to administrative implementation data. “2017 Strict” refers to the subset of the 2017 qualified population that satisfied the 2008/2018 eligibility criteria according to the (partially recorded) history of missing meals. In 2017 and 2018, 5% of loan recipients were not surveyed for migration status, so “% Migrated from Disbursed” may vary by up to 5%.

A greater portion of the treated population was deemed loan-eligible in 2017 because the criterion of having missed meals in past lean seasons was dropped in this year. The fifth column of Table 3 displays statistics for the subset of loan-eligible households that satisfy this additional criterion, though the fraction is low because the variable was sparsely recorded in administrative data. Nevertheless, implementation details within this subset are nearly identical to the full loan-eligible population, indicating that differential selection into loan eligibility was not the main source of difference in program administration in 2017. In 2018, past missed meals was reinstated as a condition of loan eligibility, and the fraction of households eligible for a loan is closer to that of the pilots.

Table 3 highlights two notable facts about recorded implementation quality. First, a greater fraction of loan recipients migrated in the at-scale implementation rounds than in the pilot. In 2008 and 2014, over a third of loan recipients did not subsequently migrate. This fraction fell to under 10% in 2017, and remained close to 30% in 2018. This fact indicates that the loss of program effectiveness—and the decline in induced migration in particular—were not caused by laxer enforcement of loan conditionality. If anything, conditionality appears to have been more strictly enforced by loan officers at scale. Instead, analysis in Section 4 focuses on outreach, acceptance, and disbursement.

Second, loan acceptance and disbursement rates differed substantially between 2017 and 2018. Implementation appears to have reached less of the recorded loan-eligible population in 2017 relative to the pilots. A smaller fraction of eligible households accepted a migration loan, even after accounting for the looser eligibility criteria, and there was considerable leakage between loan acceptance and actual disbursement, as fewer than two thirds of households that accepted an offer actually received a loan. These factors together resulted in less than a quarter of loan-eligible

households migrating with a loan in 2017, compared to around 45% in 2008 and 38% in 2014. Low acceptance and disbursement rates in 2017 were constrained by explicit loan targets assigned to implementing officers, which we discuss in greater detail in the next section.

Recognizing implementation deficiencies in 2017, program administrators adjusted the instructions given to loan officers in 2018 to explicitly incorporate language describing the intended goal of enabling induced migration. Following this change, official program implementation metrics from this year look nearly identical to the pilot rounds. Notably, close to 40% of loan-eligible households migrated with a loan in all years except 2017. Looking at administrative data alone would have led implementers to mistakenly conclude the program righted itself in 2018 and managed to replicate the pilot study’s effectiveness in facilitating migration at scale.

## 4 Delegation Risk in Program Scaling

In this section we present theory and evidence that the NLS program’s failure to replicate at scale was caused by delegation of local decision-making duties in a setting where measured implementation and impact may be misaligned. A related concern when scaling a policy is that implementers introduce design differences that alter the policy itself. NLS avoids this concern because the migration loan on offer remains consistent across all rounds of evaluation. Instead, we identify a threat to implementation that can lead program impacts to attenuate with scale *for the same policy*.

### 4.1 Theory of Implementation Compliance and Delegation Risk

We first develop a conceptual framework to characterize how program effectiveness may diminish with scale when authority is delegated to local implementing agents. To do so, we introduce a typology to describe sources of divergence between implementation and impact based on beneficiary response to a policy. We then identify conditions under which scaling may systematically direct implementation efforts away from the types of beneficiaries among whom we would anticipate the greatest impact.

#### 4.1.1 Classification of Implementation Compliance among Beneficiaries

Program beneficiaries differ in both their direct uptake of a development policy as well as their subsequent behavior contingent on policy implementation. In the case of NLS, this means that beneficiaries differ in whether they would accept and receive a migration loan if offered (uptake/service delivery), as well as whether they would migrate (behavior) with or without a loan. These possibilities generate five beneficiary types, summarized in Table 4:

1. *Opportunistic (OP)*: Accept a loan if offered; migrate with or without a loan.
2. *Self-Sufficient (SS)*: Reject a loan if offered; migrate with or without a loan.

Table 4: Classification of Beneficiary Types

| Type                 | Contingent Actions |                 |                  | Program Evaluation | Service Delivery |
|----------------------|--------------------|-----------------|------------------|--------------------|------------------|
|                      | Accept Offer       | Migrate w/ Loan | Migrate w/o Loan |                    |                  |
| Opportunistic (OP)   | Yes                | Yes             | Yes              | Always-Taker       | Yes              |
| Self-Sufficient (SS) | No                 | –               | Yes              | Always-Taker       |                  |
| Induced (IN)         | Yes                | Yes             | No               | Complier           | Yes              |
| Time-Wasting (TW)    | Yes                | No              | No               | Never-Taker        | Yes              |
| Uninterested (UN)    | No                 | –               | No               | Never-Taker        |                  |

3. *Induced (IN)*: Accept a loan if offered; migrate with a loan; do not migrate without a loan.
4. *Time-Wasting (TW)*: Accept a loan if offered; do not migrate with or without a loan.
5. *Uninterested (UN)*: Reject a loan if offered; do not migrate with or without a loan.

Beneficiary types are discussed here with reference to NLS loans, but this framework generalizes to any type of development policy with a target induced behavior and an intermediary service intended to encourage that behavior. Examples of such policies include targeted microfinance, conditional cash transfers, occupational training, and agricultural extension.

Our classification of beneficiaries is a refinement of the type space used in the LATE framework for program evaluation (Angrist et al., 1996). For instance, to calculate the LATE returns to migration, Bryan et al. (2014) and Akram et al. (2017) use treatment assignment ( $Z$ ), i.e. the offer of a migration loan, as an instrument for seasonal migration ( $D$ ), and then estimate the impact of migration on downstream economic outcomes ( $Y$ ). We introduce the intermediary stage of service delivery or program uptake, i.e. loan acceptance/disbursement, in between treatment assignment and subsequent behavior. Our extended classification maintains the LATE monotonicity assumption of No Defiers: A loan offer does not dissuade any beneficiary from migrating. We add an additional assumption of No Suggestion: A loan offer does not induce a non-migrant to migrate unless they accept and receive a loan.

Service delivery is an important intermediary outcome because it is the main observable metric from the perspective of a policy implementer. Impacts such as induced migration are generally not realized until well after implementation and cannot be quantified without reference to a control group or other credible counterfactual. Moreover, impact can only be characterized statistically; it is not possible to identify any specific treated individual as an induced migrant. By contrast, program uptake is immediately observable and service delivery can be precisely associated with a specific set of recipients.

The final two columns of Table 4 makes explicit the divergence between service delivery and program impact. The benefits of a development policy accrue to Induced beneficiaries—in this case, those that would not migrate without a loan but are enabled to do so. These correspond to

Compliers in policy evaluation and appear as successful service delivery to the implementer. By contrast, Always-Takers and Never-Takers—those who do not change their migration decision—see little benefit from NLS and therefore contribute little to measured impacts in program evaluation. However, subsets of both of these groups—Opportunistic and Time-Wasting—also take up the program and appear as service delivery successes. Delegation risk comes about when implementation at scale rewards service delivery in a way that prioritizes these latter two types over Induced beneficiaries.

#### 4.1.2 Delegation Risk and Program Scale

In principle, the mismatch between service delivery and program impact is present at any scale. We set forth three conditions that, together, will lead to the systematic exclusion of Induced beneficiaries in a way that is exacerbated by program expansion. These factors provide an explanation for the dilution of the NLS program’s role in enabling migration at scale relative to pilot.

First, decision-making authority must be delegated to implementing agents in the presence of capacity constraints or a fixed factor that does not scale with policy size. In particular, policies that employ humans to manage individual beneficiary cases must scale in at least one of two ways: either they hire more staff, increasing administrative complexity, or they increase the workload per staff member. In the former case, limitations in managerial capacity and the ability to recruit talent are likely to constrain program expansion,<sup>10</sup> and in the latter case, agent time per beneficiary falls. Either way, implementers face an increasing marginal cost of program outreach, must make decisions about how to allocate the constrained resource(s) across employees and beneficiaries as a program grows.<sup>11</sup>

Both constraints were present in the scale-up of NLS. Administrative complexity grew in 2017 and 2018, with two new loan officers hired per implementing branch to manage the migration loan portfolio. In addition, and the number of loan-eligible households per branch increased by more than an order of magnitude in comparison to the pilot. Therefore, both loan officer time devoted to outreach and administrative oversight of loan officer activities acted as constrained factors in policy scaling.

These constraints manifested in the need for loan officers to exercise discretion in their marketing and outreach efforts. At the start of each season, officers held a meeting announcing the NLS program for all eligible households in every treatment village. After this initial kickoff, officers conducted followup visits in their assigned regions and met with potential migrant households individually to continue marketing and to encourage migration with loan disbursement. Loan officers had freedom to allocate their limited time across outreach activities, and to select which villages and households to follow up with most intensely based on their perceived interest in the

---

<sup>10</sup>Das et al. (2024) present a closely related analysis of enumerator recruitment and oversight as a barrier to scaling data collection in academic research.

<sup>11</sup>This condition would not apply, for example, to a text message campaign that could readily reach more recipients at constant marginal cost.

program. By contrast, the pilot rounds were small enough that evaluators maintained careful oversight of a small enumerator team that made loan offers to treated households with uniformly high outreach intensity.

Second, delegation of decision-making risks undermining program effectiveness at scale when implementation values service delivery. This is extremely common feature of management practices in the development sector, as service delivery is an immediately observable metric that organizations use to monitor employee effort. Incentives for service delivery were certainly in place in NLS. In 2017, the MFI set an explicit loan target for each loan officer, and disbursement essentially ceased when the target was reached. In 2018, loan targets were lifted, and the MFI made an effort to promote intrinsic motivation to maximize program impact among loan officers by using the language of induced migration. Nevertheless, loan disbursement remained an implicit indicator of employee effort that played a role in loan officer performance evaluation and retention, and therefore constituted at least some part of loan officers' objective function.

Third, when authority is delegated to implementing agents who value service delivery, program impacts will be attenuated if the cost of outreach to Induced beneficiaries exceeds the cost of reaching Opportunistic or Time-Wasting beneficiaries. Heterogeneity in outreach costs may derive from both the supply of implementer effort as well as demand from beneficiaries themselves. From the implementer's perspective, it may be easier, for example, to find Opportunistic beneficiaries who already have migration plans in place before the first loan officer visit. On the beneficiary side, Opportunistic and Time-Wasting types do not change their migration behavior, and therefore may be ready to signal their demand sooner than Induced types that need time to scout opportunities and make plans before requesting a loan. In either case, tighter capacity constraints and greater implementer discretion would lead to more systematic omission of the Induced migrants that drive program benefits.

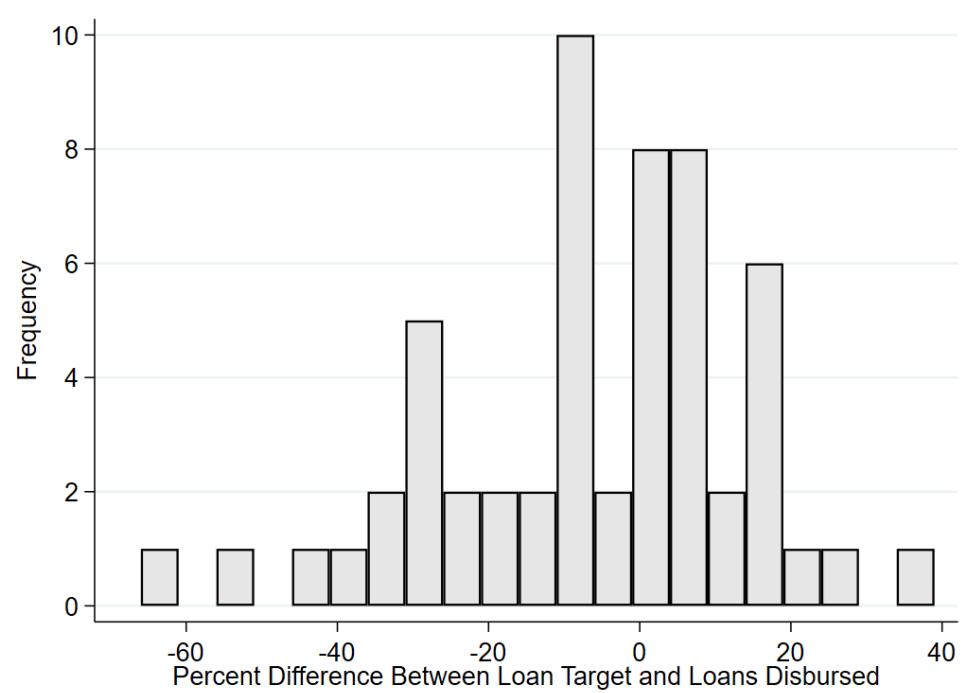
## 4.2 Evidence of Delegation Risk

We next present evidence that delegation risk led to the attenuation of NLS program impact at scale. We show that limited capacity constrained loan officer outreach efforts in treated villages, loan officers were selective in how they allocated their outreach, this selectivity led to a focus on Time-Wasting types, and that the relative exclusion of Induced migrants is sufficient to explain the decline in treatment effect from 2008 to 2018.

### 4.2.1 Capacity Constraints in Program Implementation

Low rates of loan acceptance and disbursement in 2017 reported in Table 3 were driven at least in part by binding capacity constraints. At the officer level, a debrief survey administered after the migration season reveals that 85% of officers were hired with an explicit target for the number of loans to disburse. Figure 6 plots the distribution of disbursements relative to the reported target for each officer. The figure shows that the majority of officers hit their target nearly exactly. The

Figure 6: 2017 Loan Disbursements Relative to Target



Notes: Loans disbursed by migration officer relative to target. Data from debrief survey with loan officers following 2017 migration season.

net result of these targets was that total loan disbursements across all branches in 2017 plateaued at the planned quantity of 40,000<sup>12</sup>, explaining why so many households that accepted a migration loan never received one.

Quantity targeting was abandoned in the 2018 implementation year, and loan officer instructions shifted to promote intrinsic motivation using the language of induced migration. Correspondingly, the total number of loan disbursements rose to nearly 90,000. Nevertheless, capacity constraints remained in loan officers' time as they decided where to allocate their marketing and recruitment efforts in the short window before the migration season. Moreover, while explicit loan requirements were dropped, incentives for implementation quantity implicitly remained in place through relational contracting as officers who did not disburse a sufficient number of loans were unlikely to be retained in subsequent years. As a result, loan officers faced the dual incentives of directing outreach to those without prior migration plans while still ensuring sufficient loan disbursement.

Administrative records reveal loan officers faced binding time constraints that limited their marketing and outreach. Officer records indicate the amount of time spent individually with each

<sup>12</sup>The planned number of loan disbursements fell short of the 64% acceptance rate realized in 2008 because the program made more loan offers than anticipated. At the time of implementation, microfinance branches realized they had the capacity to conduct eligibility surveys in more villages than they predicted, and expanded the scope of operations accordingly at the last minute. The anticipated number of loans had already been registered with the Bangladeshi government and could not be increased to match the new program size.

Table 5: Loan Officer Effort and Participant Offer Recall

|                         | (1)     | (2)         | (3)                     | (4)      |
|-------------------------|---------|-------------|-------------------------|----------|
|                         | Avg.    | Time per HH | Recall Offer at Endline |          |
| Distance to MFI Branch  | -0.951* |             | -0.214                  | -0.139   |
|                         | (0.525) |             | (0.212)                 | (0.197)  |
| Eligible HHs in Village |         | -0.116**    | 0.0416**                | 0.0302   |
|                         |         | (0.0570)    | (0.0192)                | (0.0199) |
| Avg. Time per HH        |         |             | 0.142**                 | 0.163*** |
|                         |         |             | (0.0687)                | (0.0593) |
| Outcome Mean            | 3.586   | 3.586       | 0.436                   | 0.436    |
| R-Squared               | 0.682   | 0.699       | 0.112                   | 0.166    |
| Obs                     | 632     | 633         | 613                     | 613      |
| Controls                | Yes     | Yes         | No                      | Yes      |
| Subdistrict FE          | Yes     | Yes         | No                      | Yes      |

Notes: Outcome for Columns (1) and (2) is village-level average number of minutes spent per household reported by loan officers in administrative records. Regressors are distance to MFI branch office (in 100ms), number of eligible households in village (in 100s), and village average minutes spent per household. All regressions include subdistrict fixed effects and household controls. Standard errors clustered at the village level in parentheses. \*( $p < 0.10$ ), \*\*( $p < 0.05$ ), \*\*\*( $p < 0.01$ )

eligible household to explain the NLS program and gauge interest during outreach visits. The mean time spent per household visit in this data is just over 2 minutes, and 99% of visits lasted for under 10 minutes.

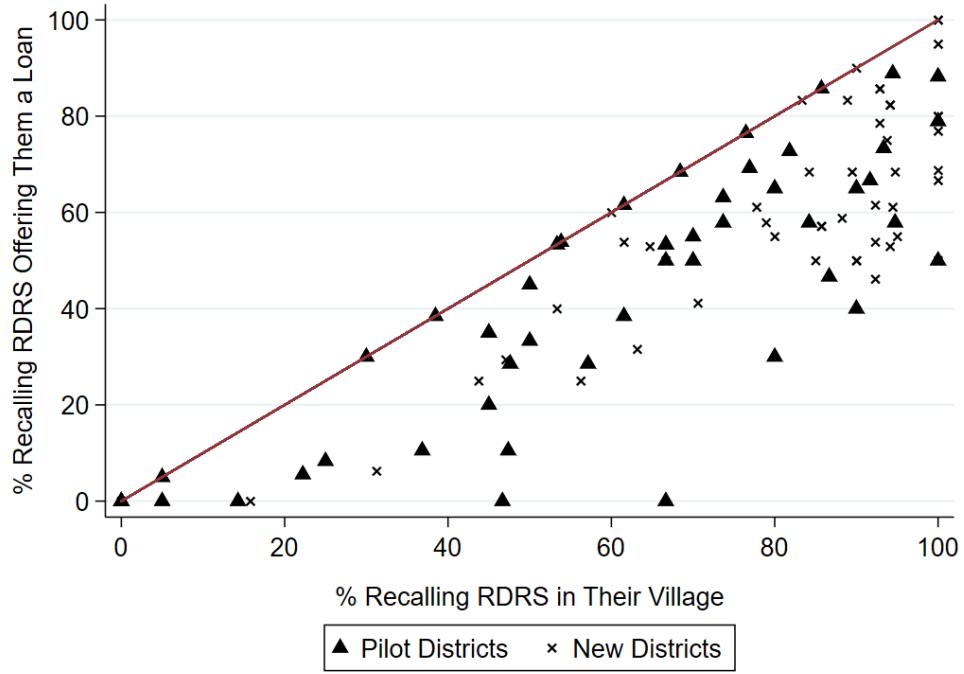
In Table 5, we show a negative relationship between the opportunity cost and intensity of outreach. Columns (1) and (2) relate the average time spent per household at the village level<sup>13</sup> to distance from the MFI branch and number of eligible households in the village, respectively. Travel distance and competing caseload both raise the opportunity cost of outreach to a household, and loan officers correspondingly spend less time per household visit under these conditions.

Household survey data following the 2018 evaluation corroborates the concern that migration officers directed different levels of effort to different households. In the endline survey, only around half of loan-eligible participants in treatment villages recall receiving a loan offer. Recall should not be considered a literal binary measure of program outreach, as all eligible households identified by the MFI were told about the loan at least once in the initial village meeting. Offer recall is positively correlated with household characteristics that predict migration, such as prior migration experience, as shown in Appendix D. It is natural that households already intending to migrate are more attentive to opportunities relating to migration.

Nevertheless, offer recall provides information on the intensity of loan officer engagement. As the third and fourth columns of Table 5 show, offer recall increases with the average duration of

<sup>13</sup>Due to data limitations, administrative records can only be matched to survey households at the village level, and not at the household level.

Figure 7: Households that Remember Own Offer and Offers in Village



Notes: Data from household recall survey following 2018 implementation. Each datapoint represents a village. x-axis: fraction of eligible households that remember loan offers made in their village. y-axis: fraction of eligible households that recall receiving a loan offer. 45-degree line in red.

meetings between loan officers and households in the village, and this relationship remains intact even after controlling for household characteristics that may predict attentiveness to the loan. in magnitude, a one standard deviation increase in loan officer time spent with households corresponds to a ten percentage point increase in the likelihood of recalling the loan offer at endline. Note also that, after adjusting for its effect on meeting duration, the number of eligible households in a village is positively correlated with loan offer recall, which is consistent with information-sharing leading to crowd-in of knowledge about local opportunities.

Discrepancies between respondents' recall of their own loan offer and loan officer activity in their village provide further evidence of heterogeneity in outreach rather than merely faulty memory. Figure 7 plots the fraction of households who remember being offered a loan by village on the y-axis against the fraction that remember loans being offered in their village on the x-axis. Nearly all the data lie below the 45-degree line, revealing a large number of eligible households who can recall NLS-related activity in their village but feel they were not invited to participate.

Table 6: Distribution and Effort Intensity by Implementation Type

| Type                 | Population Frequency |          |           | 2018 Outreach<br>Intensity |
|----------------------|----------------------|----------|-----------|----------------------------|
|                      | 2008                 | 2014 Low | 2014 High |                            |
| Opportunistic (OP)   | 19.9                 | 21.7     | 43.9      | 0.42                       |
| Self Sufficient (SS) | 16.1                 | 16.1     | 2.5       | 0.23                       |
| Induced (IN)         | 22.4                 | 24.1     | 30.5      | 0.66                       |
| Time Wasting (TW)    | 21.2                 | 14.3     | 21.4      | 0.83                       |
| Uninterested (UN)    | 20.4                 | 23.9     | 1.7       | 0.13                       |

Notes: Population frequencies in Columns 1–3 based on observed loan and migration behavior in 2008 and 2014 assuming uniform outreach intensity normalized to 1. Outreach intensity in Column 4 assumes type distribution in 2018 matches 2014. Calculation details provided in Appendix D.

### 4.3 Delegation, Outreach Intensity, and Treatment Effect Attenuation

Capacity constraints introduce the possibility of selective outreach, but it does not necessarily follow that effort is directed away from Induced migrants. Because Induced migrants eventually accept a loan, outreach to this group would in principle achieve implementation-based goals. Delegation risk poses a threat to implementation specifically when implementation prioritizes other, lower-cost alternatives to maximize loan disbursement.

To directly quantify selectivity in program outreach, we first calculate the implied type distribution in the population according to the 2008 and 2014 implementation data. For instance, 36% of households in the 2008 control group migrated, corresponding to the population fraction of Opportunistic and Self-Sufficient beneficiaries. In the treatment group, 16% declined a loan and still migrated, representing the portion that is Self-Sufficient. This value in turn implies that the other 20% must have been Opportunistic. Other population frequencies are calculated similarly, with results presented in Table 6 and detailed calculations in Appendix D.

This analysis adds detail to how both loan uptake and program impact changed as the saturation of loan offers to eligible households increased from 14% to 70% in 2014. In low-offer-saturation villages, the type distribution in 2014 closely resembles that of 2008. From an evaluation perspective, increasing offer saturation shrinks the fraction of Never-Takers by 14p.p., with a net 7p.p. growth in Compliers and another 7p.p. growth in Always-Takers. This shift reflects the crowd-in into migration reported by Akram et al. (2017) as some households that would not have migrated in a low-saturation village choose to migrate in a high-saturation village.

Further breaking down beneficiary types by policy response reveals a secondary effect of increasing loan offer saturation. Within non-compliers, the rate of loan acceptance grew to above 90% from a prior rate of 59% among Always-Takers and 37% among Never-Takers. That is, increasing the fraction of households offered a loan transformed many Self-Sufficient types into Opportunistic, and Uninterested into Time-Wasting. In effect, high saturation of loan offers crowds in both migration and service delivery regardless of migration status.

### 4.3.1 Outreach Intensity by Type

To verify that implementer outreach matters, we turn to the information-only arm from 2008. In this arm, participants received information about labor market conditions in potential migration destinations, which by itself had no effect on migration rates relative to control. Participants were also asked about initial demand for a migration loan, but with no subsequent followup as they were not ultimately eligible to receive one. Migrants who said they would accept and reject a loan offer make up 17.8% and 18.5% of this group, respectively, nearly mirroring the anticipated rates of Opportunistic and Self-Sufficient types. However, among non-migrants, those that would accept alone only comprise 17.8% of this group, while a full 45.8% are non-migrants who indicate they would not want a migration loan. The excess of what appear to be Uninterested types indicates that some Time-Wasting or Induced migrants need multiple points of outreach before deciding to accept a loan. Unfortunately, With these data alone, it is not possible to separately identify the relative frequencies of these types among the excess loan rejections.

Selective outreach in the face of capacity constraints is already apparent in high-saturation villages in 2014. There is a stark discrepancy between loan acceptance rates reported in administrative data in Table 3 and among the survey sample reported in Table 6. 97% of survey respondents assigned to high-saturation treatment accepted a loan, but only 65% of offers are accepted in total. This gap is reconciled by the fact that among households assigned to high-saturation treatment but *not* selected for surveying and evaluation, the loan acceptance rate is only 49%. From the net disbursement rate, it appears that loan capacity was comparable among high-saturation and low-saturation villages in 2014. However, outreach in high-saturation villages was targeted to the survey sample, possibly either because implementers felt greater oversight pressure among this sample or because surveying itself served as a form of marketing.<sup>14</sup>

For evaluation at scale in 2017 and 2018, the research team aimed to minimize any influence of surveying and evaluation on program implementation. To achieve this, we maintained strict separation between survey teams and the implementing organization so that evaluation would not provide feedback or seem like an assessment of implementation. In addition, surveyors collected minimal baseline data to prevent the baseline serving as program outreach, and held off on detailed data collection until after the agricultural lean season. We next explore how capacity constraints manifested in program implementation in the absence of researchers placing special focus and maintaining high contact with visible evaluation subset of beneficiaries.

For this calculation, we start from the assumption that outreach to all sample households assigned to treatment in pilot rounds was uniformly high. Equivalently, we can think of the pilot as the “best possible” implementation of NLS, at least among those surveyed for evaluation. Normalizing this outreach level to 1, we compute the implied outreach intensity for each beneficiary

---

<sup>14</sup>It is also worth noting that MFI administrative records of eligible households in 2014, which constituted the sampling frame for surveying, contained 13% fewer households than the number identified as eligible in an initial census, pointing to another potential source of implementation leakage.

type in 2018 using survey data on migration, loan acceptance, and offer recall. In 2018 endline survey data, we observe loan offer recall, loan acceptance conditional on recall, and migration with or without a loan for every sample household in treatment villages. The fraction of the population in each data cell can be written as a function of the type distribution and the outreach directed to each type.

Formally, for each beneficiary type  $X$ , let  $P_X$  represent the fraction of  $X$  type in the beneficiary population, and let  $\omega_X$  represent the average outreach intensity to type  $X$ .  $\omega_X = 1$  would indicate that outreach to type  $X$  matched that of the pilot, and lower values indicate that some fraction of type  $X$  did not receive sufficient outreach to be aware of the NLS program.

Outreach intensity to the three types that accept a loan when offered can be inferred from loan acceptance and migration within the treated population alone and do not rely on offer recall:

$$\begin{aligned} \text{No Loan, Migrate} &= (1 - \omega_{OP})P_{OP} + P_{SS} \\ \text{Accept, Migrate} &= \omega_{OP}P_{OP} + \omega_{IN}P_{IN} \\ \text{Accept, Remain} &= \omega_{TW}P_{TW} \end{aligned} \tag{2}$$

The first equation in (2) states that households who migrate without a loan in the treated group comprise all Self-Sufficient types ( $P_{SS}$ ) as well as Opportunistic types that did not receive sufficient outreach to take up the NLS program ( $(1 - \omega_{OP})P_{OP}$ ). The other two equations similarly describe the makeup of migrants and non-migrants who received sufficient outreach to be aware of the loan offer and accept it

Taking seriously the offer recall data allows for further inference on outreach intensity to types that reject a loan offer as well. Formally:

$$\begin{aligned} \text{Recall, Decline, Migrate} &= \omega_{SS}P_{SS} \\ \text{Recall, Decline, Remain} &= \omega_{UN}P_{UN} \end{aligned} \tag{3}$$

This second set of calculations must be interpreted with the caveat that  $\omega_{SS}$  and  $\omega_{UN}$  conflate outreach intensity with heterogeneity in propensity to recall a loan conditional on outreach intensity, which may be lower for the types that do not want a loan in the first place.

In both 2017 and 2018, the fraction of migrants among the control population (i.e.  $P_{OP} + P_{SS}$ ) was 38%, very close to the corresponding value among untreated households in the 2014 low-saturation villages. Therefore, we hold fixed the type distribution from that sample and solve this system of equations for  $\{\omega_X\}$  in 2018<sup>15</sup>. Full details of this calculation are provided in Appendix D. The implied outreach intensity by type is reported in the final column of Table 6.

Delegation risk at scale appears in the estimated outreach intensity toward Induced types relative to others. Outreach is lower for Opportunistic types, who likely already had plans to migrate

---

<sup>15</sup>We unfortunately cannot conduct this exercise in 2017 because the 2017 survey data do not include information on loan acceptance, and survey households cannot be matched to administrative loan data.

ahead of the lean season. This difference is consistent with loan officers attempting to faithfully implement the instructions given in 2018 by directing outreach to those who would not be able to migrate without a loan.

Among those without migration plans, however, outreach intensity is 25% greater for Time-Wasting types than Induced. Even though implementers seem to have sought out households without migration plans, their efforts to reach those who would be enabled to migrate were diluted by non-migrant households who accepted a loan but did not migrate. Unfortunately for implementation oversight, both groups look identical *ex ante*—they do not have migration plans and accept a migration loan—but only Induced migrants generate the program impact seen in piloting.

Heterogeneous outreach intensity may reflect selective effort from loan officers (supply) or variable response to outreach from beneficiary types (demand). That is, the fact that  $\omega_{IN} < \omega_{TW}$  may arise because, among non-migrants, loan officers have an easier time finding Time-Wasting types and convincing them to take a loan, and therefore direct more of their limited focus to this group. Alternatively, it may reflect the fact that Induced types require a greater level of interaction to take up a loan because they must first put effort into changing their migration plans and identifying a target destination. In the latter case, implementation at scale may lose Induced types faster than Time-Wasting even if loan officers lowered their outreach efforts uniformly across all beneficiaries with no intentional selectivity at all.

The fraction of loan recipients who subsequently migrate is nearly identical in survey (67.9%) and administrative (65.9%) data. This consistency corroborates the relative prevalence of Time-Wasting types among loan accepters inferred from the household survey. By contrast, 84% of survey respondents who remember the loan offer report accepting it, compared to a 64% acceptance rate among offered households in the administrative data. This gap is consistent with low recall among those that decline a loan, which would lead us to underestimate outreach intensity to Self-Sufficient and Uninterested types.

To reconcile these values with faulty survey recall alone, it would need to be the case that only a third of those that consciously decline a loan recall having done so by the time of the endline survey. Correspondingly, true outreach intensity to Self-Sufficient and Uninterested types would be three times greater on average than that implied by survey recall. Given both loan officers' apparent focus on those without migration plans and the evidence from Appendix D that attentiveness to loan offers is greater among those with migration plans, it is likely that Uninterested types have the lowest recall propensity after declining a loan. Indeed, if Self-Sufficient types were reached with the same outreach intensity as Opportunistic, and Uninterested with the same intensity as Time-Wasting, and we scale up the offer recall rate among these groups correspondingly, then the average loan acceptance rate implied by the survey data would be 63.2%, nearly matching the administrative data. This suggests that imperfect recall among those who decline a loan is a plausible explanation for the survey—administrative gap in acceptance rates.

### 4.3.2 Treatment Effect Attenuation

We quantify the importance of delegation risk by analyzing treatment effect heterogeneity by outreach intensity. To do so, we first identify the baseline household characteristics that predict outreach intensity by modeling recall propensity in the 2018 survey data as a flexible function of observables, estimated using the random forest algorithm of [Wager and Athey \(2018\)](#). The variables most commonly selected as relevant over multiple iterations of the algorithm include migration history, reflecting non-loan migration status; borrowing history, reflecting an existing relationship with the MFI; and distance to MFI branch, reflecting outreach cost.

Next, we multiply the coefficients from the random forest algorithm with household characteristics in 2008. The fitted values we generate serve as a prediction of what the outreach intensity to each household would have been in the 2008 pilot if program implementation had matched that of 2018<sup>16</sup>.

Finally, we estimate heterogeneous treatment effects in 2008 by predicted outreach intensity. Results from this exercise are presented in Table 7. The table reports the average point estimates and standard errors from regressions following (1) over 100 iterations of the algorithm to predict outreach intensity. The first row reflects the estimated treatment effect of NLS on migration among households in the lowest quintile of predicted outreach intensity. The next four rows report treatment effect heterogeneity in the next four quintiles of predicted outreach intensity in ascending order.

While the estimates are noisy, Table 7 indicates that the 2008 NLS treatment effect is greatest among those that would have had the highest outreach intensity in 2018. This relationship is once again consistent with loan officers attempting to identify and make offers to Induced migrants. However, the second-highest propensity quintile has the lowest treatment effect, suggesting crowd-out by Time-Wasting types. The two quintiles after that also have a lower treatment effect than the lowest quintile. These outcomes are robust to controlling for the household characteristics correlated with household recall, suggesting that the effect is not simply a function of likely migrants being more attentive to the loan offer.

The net effect of this heterogeneity is to lower the aggregate impact of NLS on migration relative to a uniform decline in outreach intensity. The final two rows of Table 7 present this comparison. We first compute

$$\tilde{\beta}_{2008}^{DR} = \frac{1}{N} \sum_i L_i * \left( \beta^0 + \sum_k \beta^k \mathbf{1}\{i \in Bin_k\} \right) \quad (4)$$

where  $L_i$  is the predicted outreach intensity for participant  $i$ . This is multiplied by participant  $i$ 's predicted treatment effect, where  $\beta^0$  is the estimated treatment effect among the lowest-intensity quintile, and  $\beta^k$  adds treatment effect heterogeneity in participant  $i$ 's prediction quintile. Averaging

---

<sup>16</sup>We can only conduct this exercise using the 2008 pilot data because the same baseline characteristics were not recorded in 2014.

Table 7: 2008 Treatment Effect Heterogeneity by 2018-Predicted Offer Recall Propensity

|                              | (1)               | (2)               |
|------------------------------|-------------------|-------------------|
| Treatment                    | 0.182<br>(0.056)  | 0.185<br>(0.054)  |
| Treatment X Prediction Bin 2 | -0.005<br>(0.077) | -0.008<br>(0.075) |
| Treatment X Prediction Bin 3 | -0.014<br>(0.080) | -0.019<br>(0.077) |
| Treatment X Prediction Bin 4 | -0.025<br>(0.083) | -0.028<br>(0.079) |
| Treatment X Prediction Bin 5 | 0.028<br>(0.083)  | 0.065<br>(0.081)  |
| Controls                     | No                | Yes               |
| Subdistrict FE               | Yes               | Yes               |
| Replications                 | 100               | 100               |
| $\tilde{\beta}_{2008}^{DR}$  | 0.060             | 0.063             |
| $\tilde{\beta}_{2008}^U$     | 0.078             | 0.078             |

Notes: Outcome is migration in 2008 pilot experiment. Prediction quantiles calculated based probability of recalling loan offer in 2018 using causal forests following [Wager and Athey \(2018\)](#). Recall probability is ascending by bin.  $\tilde{\beta}_{2008}^{DR}$  computed as predicted treatment effect given outreach intensity interacted with treatment effect heterogeneity.  $\tilde{\beta}_{2008}^U$  computed as 2008 treatment effect estimate scaled down by average likelihood of loan offer recall.

across the study population,  $\tilde{\beta}_{2008}^{DR}$  estimates the aggregate treatment effect we would expect to observe in 2008 had the pilot been implemented with the same heterogeneous outreach intensity as in 2018.

$\tilde{\beta}_{2008}^{DR}$  remains stable around 0.06 across iterations of the prediction algorithm, well below the observed 2018 treatment effect of 0.12. However, this is almost certainly overestimates expected attenuation because loan offer recall is an imperfect proxy, and likely an underestimate of, outreach intensity. For a more apt comparison, the final row of Table 7 simply scales down the estimated 2008 treatment effect by the average loan offer recall rate, labeled as  $\tilde{\beta}_{2008}^U$ . This value is the aggregate treatment effect we would expect to observe in 2008 had the pilot lowered outreach intensity uniformly across all types to the same degree as in 2018.

$\tilde{\beta}_{2008}^{DR} < \tilde{\beta}_{2008}^U$ , meaning delegation risk attenuates the NLS treatment effect over and above diminished program capacity alone. From the actual 2008 treatment effect of 0.18, the level of capacity constraint inferred from the recall data applied equally across the treated population would have lowered the treatment effect by over half to 0.078. Heterogeneity in outreach intensity, specifically the mis-targeting of Induced types, adds an additional 15–17% to treatment effect attenuation.

Note that, to the extent attentiveness to loan offers is lower among those that decline than among those that accept, imperfect recall would bias away from the explanatory power of delegation risk.

This is because the types that decline a loan—Self-Sufficient and Uninterested—have zero response to treatment. Low recall among these groups would lead them to be concentrated in the lowest propensity prediction bins and therefore be assigned the lowest weight in (4). Effectively, it would appear as though little outreach was directed to types with no treatment effect. At the same time, the average  $\tilde{\beta}_{2008}^U$  would decline by more because it assigns equal weight to the offer recall of all participants. Combined, these effects diminish the apparent importance of delegation risk.

The analysis on delegation risk in this section takes advantage of the fact that NLS implementation has a discrete, measurable intermediary outcome of loan disbursement. In other types of development policy, especially those centered around training, extension, or information treatments, there may be important unobserved intermediate outcomes such as participant attention, learning, or persuasion, that mediate program impacts. The mapping from measurable service delivery to induced behavior may be more complicated in such cases, and service delivery itself may not be well-defined. These difficulties would make it even more challenging for implementers and researchers alike to distinguish between beneficiary types and assess success of program implementation.

## 5 Additional Threats to External Validity of Pilot Evaluation

Evidence suggests that capacity constraints and selective outreach together can account for the decline in program effectiveness between pilot and at-scale implementations of NLS. In this section we discuss the additional importance of geographic heterogeneity and rule out two other possible explanations for why program effectiveness may have diminished from pilot to scale.

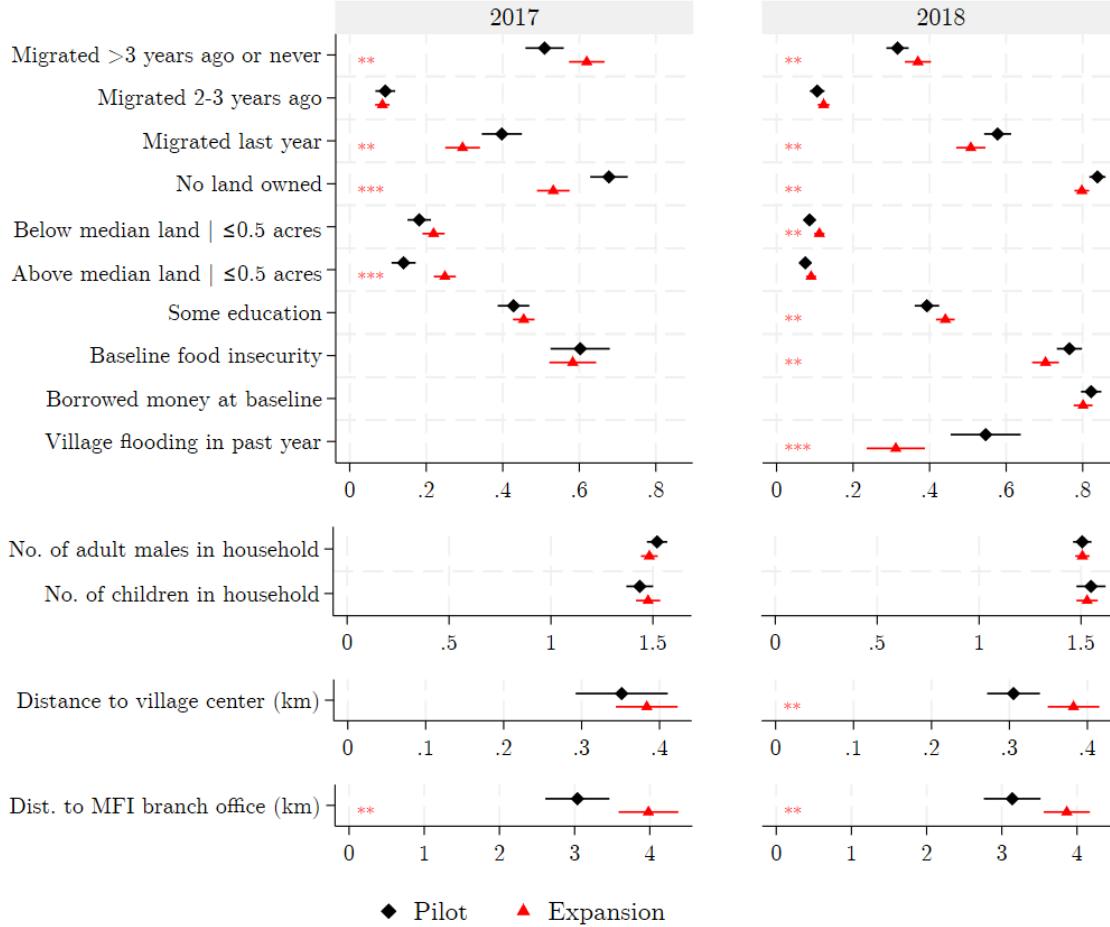
### 5.1 Geographic Expansion

The NLS program was piloted in villages randomly sampled from two districts within the Rangpur division of Bangladesh. The at-scale evaluation results presented thus far restrict analysis to these two districts, where we would expect the pilot results to be statistically representative.

In scaling up, the program both increased coverage in these pilot districts as well as expanded geographically to incorporate the remaining six districts in the division where the lean season economic cycle is similarly present. Figure 8 describes differences in baseline characteristics between the eligible population in the pilot and expansion districts in 2017 and 2018, respectively. Populations are similar across the region on most characteristics; a test of joint significance fails to reject equality across all variables excluding distance to the village center and to the MFI branch office. Notably, migration history among the eligible population is comparable between districts. The analogous macroeconomic conditions, parity on observables among the loan-eligible population, and geographic proximity of the expansion districts to the pilot set the expectation that program treatment effects will closely match.

Despite these similarities, there is striking heterogeneity in NLS program impacts between pilot and expansion districts. We present estimates of the treatment effect on migration following (1)

Figure 8: Balance across Pilot and Expansion Districts on Baseline Covariates



Notes: Variable means and standard errors for baseline household outcomes. Full numerical results are reported in Appendix B. Standard errors clustered at the village level. Stars at left side of each row indicate statistically significant deviations between pilot and expansion districts. \*( $p < 0.10$ ), \*\*( $p < 0.05$ ), \*\*\*( $p < 0.01$ )

separately by year and region in Table 8. The first row reveals that in the 2018 evaluation round, when loan offers in the original pilot districts raised household migration by 12p.p., the effect in expansion districts was small, statistically indistinguishable from zero, and if anything negative. Combining pilot and expansion districts to estimate a pooled effect would have led us to conclude the NLS effect on migration was substantially smaller in 2018 and statistically indistinguishable from zero in both years. Disappointingly, this geographic heterogeneity in treatment effect would not have been easy to predict given the observable similarity of targeted beneficiaries in the pilot and expansion districts at baseline.

One consistent point of difference between pilot and expansion districts is geographic dispersion. Households are on average 25% farther from the branch office and from each other in expansion districts. Dispersion has the potential to tighten the constraint on loan officers' time by raising the

Table 8: Effect on Migration in All Districts and Treatment Arms

|                | 2017 Districts  |                 |                 | 2018 Districts |                |                 |
|----------------|-----------------|-----------------|-----------------|----------------|----------------|-----------------|
|                | All             | Pilot           | New             | All            | Pilot          | New             |
| Treated        | 0.06<br>(0.03)  | 0.04<br>(0.05)  | 0.06<br>(0.04)  | 0.06<br>(0.03) | 0.12<br>(0.04) | -0.03<br>(0.03) |
| Spillover      | -0.03<br>(0.04) | -0.01<br>(0.06) | -0.05<br>(0.05) | 0.02<br>(0.04) | 0.13<br>(0.06) | -0.09<br>(0.04) |
| Branch Control | -0.02<br>(0.04) | -0.02<br>(0.06) | -0.02<br>(0.04) | 0.02<br>(0.03) | 0.09<br>(0.05) | -0.07<br>(0.04) |
| Control Mean   | 0.36            | 0.40            | 0.33            | 0.38           | 0.39           | 0.38            |
| HH Controls    | Yes             | Yes             | Yes             | Yes            | Yes            | Yes             |
| Upazila FEs    | Yes             | Yes             | Yes             | Yes            | Yes            | Yes             |
| N              | 3,678           | 1,537           | 2,141           | 4,324          | 1,901          | 2,423           |

Notes: Regression results in Pilot and Expansion districts following (1). Outcome is a dummy for whether any member of the household migrated during the lean season. Standard errors clustered at the village level.

cost of treatment per household, thereby exacerbating delegation risk. Applying (2) to expansion districts, again holding the type distribution fixed at 2014 pilot levels, would suggest that outreach to Induced types in these districts was only a quarter as intense as outreach to Opportunistic and under a fifth of outreach to Time Wasting types in these districts. However, it is likely that both outreach intensity and the type distribution itself diverges from pilot as the program expands geographically.

## 5.2 Population Changes over Time

The population under study in the at-scale evaluation does not just differ from the pilot population in geography, it also differs in time. In particular, we investigate two key factors that differentiate the at-scale evaluation years from pilot and may have altered the prevalence of program compliers. First, there has been a secular trend of economic growth in the region, leading the later evaluation rounds to take place among a generally wealthier population. Second, both 2017 and 2018 saw substantially greater rainfall than 2008 and 2014, especially during the rainy months of the lean season. These and other changes from year to year may affect the population responsiveness to NLS loan offers.

For analysis of the importance of changes in population characteristics, we restrict to the 2018 evaluation year to ensure that findings are not confounded by issues in implementation in 2017.

### 5.2.1 Household Characteristics

We first explore the importance of wealth and other household characteristics by estimating conditional treatment effects in the pilot data. As a demonstrative exercise, consider household calorie

consumption: calories per capita increased by nearly 25% over the intervening decade from 2008 to 2018. Appendix Figure S3 plots the full distribution of per capita calorie consumption across control households in 2008 and 2018.

To evaluate the importance of calorie consumption in responsiveness to treatment, we proceed in three steps. First, we divide households from the 2008 evaluation into  $k$  bins using data on calorie consumption collected before migration loan offers were made. Second, we estimate a bin-specific treatment effect in the 2008 data, effectively measuring the prevalence of induced migrants within each bin. Third, we construct a counterfactual 2018 treatment effect by taking a weighted average of bin-specific treatment effects weighted by the share of the population in each bin in the 2018 control group.<sup>17</sup>

This counterfactual describes the average migration response we would expect to observe if conditional treatment effects remained constant over time, and the only change from 2008 to 2018 was the distribution of baseline calorie consumption in the study population. We compare this counterfactual to the true treatment effect from 2018 to quantify how much of the gap can be explained by changes in calorie consumption leading to fewer induced migrants in the overall loan-eligible population.

We present results from this demonstrative exercise in Figure 9. The x-axis plots the number of bins used to estimate conditional treatment effects, and the y-axis plots the fraction of the difference between 2008 and 2018 that can be attributed to changes in the calorie distribution over time. The figure shows that with a small number bins, this exercise can explain on average only 10–15 percent of the attenuation in treatment effect, and the explanatory power shrinks as the number of calorie bins grows.

We generalize this exercise using a machine learning algorithm on all baseline characteristics available in 2008. Specifically, we implement a random forest algorithm following Wager and Athey (2018) to select baseline variables with the greatest explanatory power for conditional treatment effects in the 2008 evaluation. The algorithm selects from covariates among baseline education of household head, food insecurity, outstanding debt, cultivable land owned, household size, number of adult males, number of children, and distance to MFI branch. Summary statistics for these variables in each year are presented in Appendix Table S9. We then interact conditional treatment effects with the distribution of the algorithm-selected variables in the 2018 control group to construct a counterfactual 2018 treatment effect.

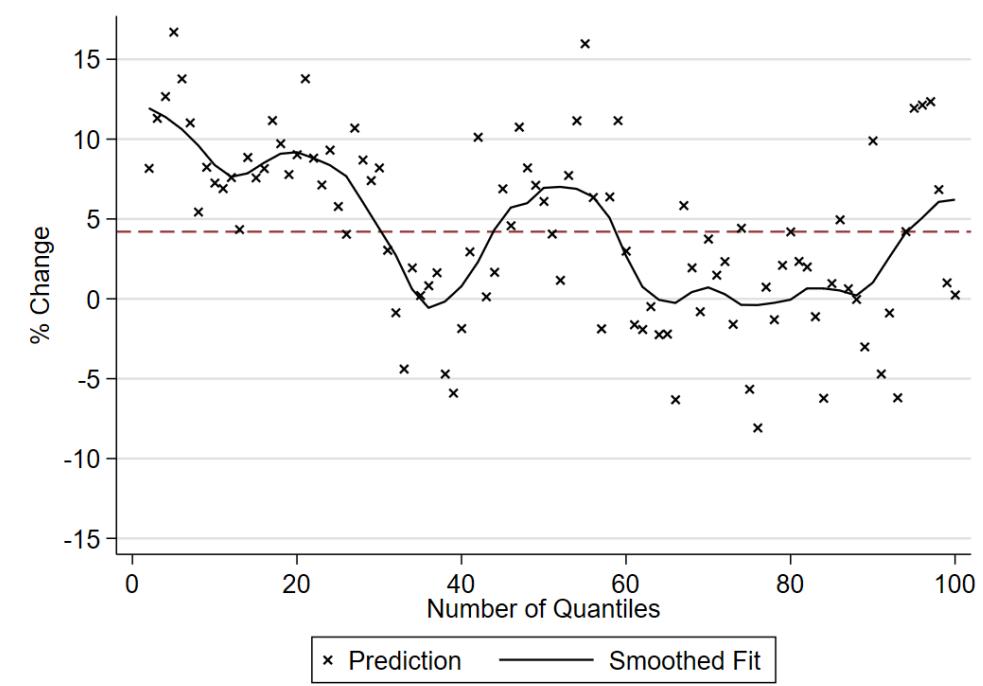
The machine learning algorithm generates a counterfactual 2018 treatment effect of 21.9%, almost identical to the actual estimated impact of 22.6%<sup>18</sup> in 2008. This small attenuation accounts for only 4% of the measured difference in treatment effect between 2008 and 2018. As a validity check, we verify that weighting the 2008 marginal treatment effects by the 2008 distribution of population characteristics recovers the 2008 treatment effect nearly exactly at 22.6%.

---

<sup>17</sup>We lack baseline calorie consumption data to directly estimate treatment effect heterogeneity in 2018.

<sup>18</sup>This value differs slightly from Table 2 because it is computed without controlling for household characteristics or subdistrict fixed effects.

Figure 9: Treatment Effect Attenuation Predicted by Change in Calorie Consumption over Time



Notes: x-axis plots number of quantile bins used to compute conditional treatment effects. y-axis plots fraction of difference between pilot and at-scale treatment effect explained by reweighting 2008 conditional treatment effects using 2018 population shares.

### 5.2.2 Rainfall and Flooding

We are unfortunately unable to include rainfall in this conditional treatment effect calculation because there is little overlap in support between rainfall in 2008 and in 2017/2018. Both 2017 and 2018 saw above-average levels of rainfall before and during the agricultural lean season, with historically high levels of flooding experienced in 2017. In contrast, 2008 was an abnormally dry year with most study village experiencing below-average rainfall. Rainfall data by year are presented in Appendix Figure S4. In principle, rainfall and flooding may interfere with induced migration by damaging roads and other migration infrastructure, raising the cost of leaving household assets unprotected, or hindering loan officer outreach and disbursement.

We explore the relationship between rainfall and NLS program impact by leveraging cross-sectional heterogeneity in local rainfall relative to historical average. To do so, we measure rainfall during the rainy season in each village over the period 2001 to 2019 using satellite data from the National Oceanic and Atmospheric Administration. We create a dummy for whether each village received above- or below-average rainfall during the lean season relative to this period. As a validation of this measure, recall survey responses show a strong correlation between our inference from satellite data and participants' self-reported memory of flooding over 2014–2018.

Table 9: 2018 Treatment Effect Heterogeneity by Rainfall

|  | (1)<br>Sent<br>Migrant | (2)<br>Sent<br>Migrant | (3)<br>Sent<br>Migrant |
|--|------------------------|------------------------|------------------------|
| Treated                                  | 0.07**<br>(0.03)       | 0.06<br>(0.12)         | 0.02<br>(0.09)         |
| Treated $\times$ Above Avg Rain=1        |                        |                        | 0.05<br>(0.09)         |
| Spillover                                | 0.05<br>(0.04)         | -0.01<br>(0.12)        | -0.05<br>(0.09)        |
| Spillover $\times$ Above Avg Rain=1      |                        |                        | 0.08<br>(0.10)         |
| Branch Control                           | 0.05<br>(0.04)         | 0.00<br>(0.12)         | -0.07<br>(0.09)        |
| Branch Control $\times$ Above Avg Rain=1 |                        |                        | 0.11<br>(0.10)         |
| Above Avg Rain=1                         |                        |                        | -0.04<br>(0.08)        |
| Mean Rainfall                            | 2222.021               | 1758.971               | 2122.001               |
| R-Squared                                | 0.140                  | 0.144                  | 0.132                  |
| Obs                                      | 3390                   | 934                    | 4324                   |
| Controls                                 | Yes                    | Yes                    | Yes                    |
| Upazila FE                               | Yes                    | Yes                    | Yes                    |
| Sample                                   | Above Average Rainfall | Below Average Rainfall | Full Sample            |

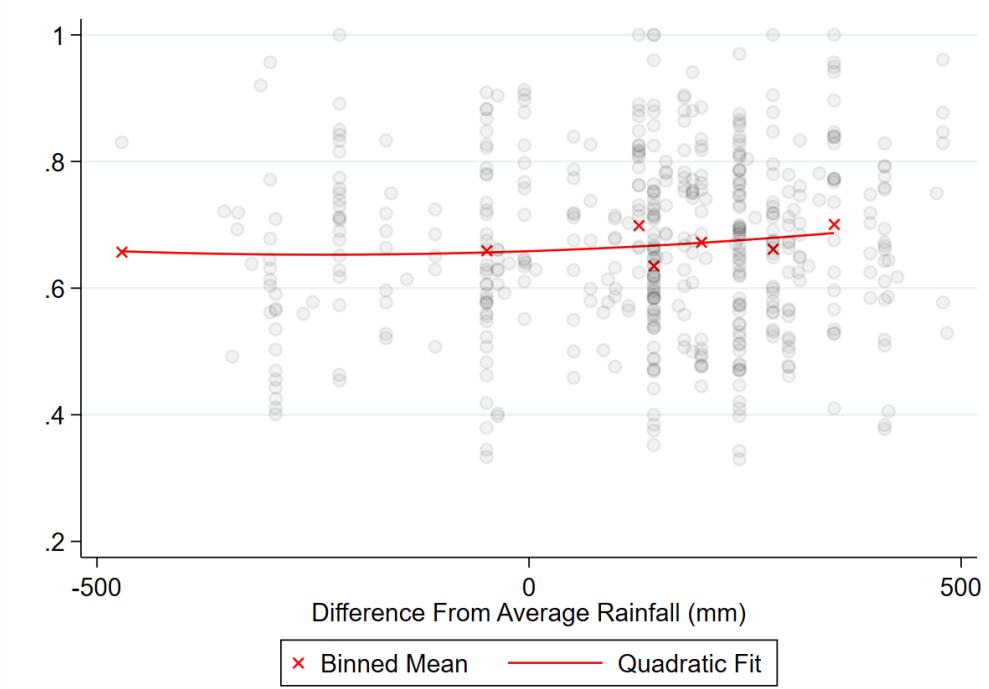
Notes: Outcome is a dummy for whether any member of the household migrated during the lean season. Column (1) presents results in villages with above-average rainfall relative to 2001–2019; Column (2) presents results in villages with below-average rainfall; and Column (3) combines all villages. Standard errors clustered at the village level in parentheses. \*( $p < 0.10$ ), \*\*( $p < 0.05$ ), \*\*\*( $p < 0.01$ )

First, we estimate heterogeneity in the NLS impact on migration in 2018 by village-level rainfall. Results are presented in Table 9. The first two columns estimate (1) separately for villages with above- and below-average rainfall, and the third column pools all study villages with an interaction term between treatment assignment and rainfall status.

Heterogeneity by rainfall is not statistically distinguishable from zero. The sign of the point estimates indicate that migration is slightly lower in villages with higher rainfall. However, the treatment effect of NLS is, if anything, larger in these villages, suggesting that high rainfall across the region is unlikely to have disrupted effectiveness.

As an alternative measure, we report the association between rainfall relative to average and

Figure 10: 2018 Loan Acceptance Rate by Rainfall Relative to Average



Notes: Each point represents a village in the treatment region. Red X's report bin averages, and the red line plots the best-fit quadratic trendline. x-axis: Rainfall during 2018 lean season relative to 2001–2019 average according to NOAA satellite data. y-axis: Fraction of eligible households that accept a loan in MFI administrative data.

migration loan disbursement among eligible households in treated villages according to MFI administrative records. Figure 10 plots this relationship by village, with deviation from average rainfall on the x-axis and the fraction of loan-eligible households that accept a loan on the y-axis. The correlation is once again positive and statistically indistinguishable from zero, confirming that an unlucky rainfall realization does not appear to have derailed implementation at scale.

Together, these facts suggest that observable changes in population and weather characteristics over time offer little explanatory power to account for the attenuation in treatment effect between evaluation rounds. Our analysis of change in treatment effect over time must be interpreted with the caveat that we draw inferences from cross-sectional heterogeneity in responsiveness to treatment. This approach is necessary because it is feasible experimentally estimate cross-sectional heterogeneity by randomizing over a large  $N$ , but much more difficult to experimentally estimate time-series heterogeneity by randomizing over a large  $T$ . Rosenzweig and Udry (2019) caution that cross-sectional variation may not fully capture the importance of macroeconomic conditions that differ over time. Nevertheless, we evaluate heterogeneity to the extent possible given the limitations of the experimental design. Overall, we find little evidence of time trends that led a program that was effective in 2008 and 2014 to suddenly lose potency over the next three years.

### 5.3 General Equilibrium Crowd-Out

Induced migration may be diminished at scale, relative to piloting, in the presence of general equilibrium factors that lead to crow-out. With respect to NLS, a large-scale labor supply shock that raises wages at the origin or lowers wages at common migration destinations could lower returns to migration, or a limited number of local migration opportunities might become saturated with broad loan availability. While [Akram et al. \(2017\)](#) find these factors to be less important than crowd-in through shared migration costs as the number of loan offers within a village increase, they may be more salient at scale given the large scope and regional coverage of the NLS program.

We explicitly test for regional spillovers across villages in our randomization design. Randomization generates one untreated, high-spillover village per MFI branch located in the middle of a group of treated villages. In addition, each branch catchment area contains a region of untreated villages, designated as branch-control, adjacent to the area where loans are offered. If migrants induced by the NLS implementation at scale crowd out other potential migrants from the area, we would expect to see depressed rates of migration in these villages relative to pure control where no loans are offered nearby.

Cross-village analysis reinforces findings from the 2014 evaluation that migration increases with NLS treatment intensity, indicating there is still crowd-in rather than crowd-out of new migrants. The second and third rows of Table 8 present estimates following (1) of the NLS program impact in spillover and branch-control villages relative to pure control. Among both groups, treatment effects are generally small and statistically insignificant. The only exception, reported in Column (5), is in pilot districts in 2018—precisely the area and year with the greatest induced migration in treated villages—where we see evidence of greater migration in spillover and branch-control villages as well.

Heterogeneity by distance to NLS branch office strengthens the case for crowd-in of new migrants. Distance to branch office has no relationship to migration on its own, but it serves as a proxy for branch-control exposure to spillovers due to our randomization design: branch-control villages are radially distinct from the treated sector within a branch’s catchment area, so proximity to the center of the catchment area also corresponds to proximity to the treated sector and therefore exposure to spillovers. Heterogeneity analysis among branch-control villages reveals that migration increases with proximity to branch office, again uniquely in pilot districts in 2018. These results suggest that, at the current scale, migration opportunities have not been exhausted and labor demand at migrant destinations is sufficiently elastic to absorb the labor supply shock.

## 6 Conclusion

In this paper, we report results from an evaluation of a large-scale migration loan program that fails to replicate the success achieved in pilot. We show that this failure to replicate cannot be attributed to crowd-out in general equilibrium, and are unlikely to result from population changes

over time. We introduce and provide evidence supporting a new theory of delegation risk caused by capacity-constrained implementers directing outreach intensity within the population of target beneficiaries.

The conditions that lead to delegation risk are fairly general. Any program with an increasing marginal cost or shadow cost of effort will lead implementers to be more selective as scale grows. There is room for this selectivity to be mistargeted when benefits are concentrated among a subset of the population that are Induced to undertake a target behavior, but other types that take up the policy without altering their behavior are easier to reach. In addition to directed lending programs such as NLS, many types of policy satisfy these conditions. For example, other forms of microfinance, conditional cash transfers, occupational or technological training, and agricultural extension all seek to enable or promote behaviors that some subset of the population would already engage in.

In general, contracts for implementing agents cannot directly reward effort focused on Induced beneficiaries because beneficiary type is typically unobservable. A common management practice in many development organizations is to focus instead on implementation and service delivery. This metric is both used to evaluate employee performance as well as reported to donors and other benefactors as a measure of impact. Unfortunately, contracts and career incentives built around implementation quantity can induce selection in exactly the wrong direction if it is easier Induced beneficiaries are difficult to identify in the target population.

It remains an open question how best to design agent incentives. The ideal contract would reward program impact. In the case of NLS, this would mean paying and retaining loan officers based on the Induced migration in their catchment area. Unfortunately, this is only possible with a credible counterfactual, such as from independently collected data on an experimental control group. More feasible alternatives may use performance bonuses and competition across agents, but such schemes risk triggering fairness concerns as outcomes are influenced by unobserved cross-sectional heterogeneity. Our experience also indicates that replacing implementation targets with intrinsic motivation can recover some, but not all, of a program's intended effect.

The presence of delegation risk adds complications to cost–benefit analysis. Pilot studies often draw a distinction between the cost of evaluation and implementation, and report only the latter for policy analysis. Our work suggests that evaluation plays an important monitoring and oversight role by providing detailed feedback on implementation quality, and surveying itself may provide additional program outreach. This monitoring cannot be divorced from the implementation and should be factored into program costs.

This lesson also applies when evaluating aid effectiveness. Rating agencies for charitable giving commonly focus on the fraction of an organization's budget devoted to beneficiaries.<sup>19</sup> We find that resources for effective targeting can be equally important as poorly targeted programs can fail

---

<sup>19</sup>For instance, Charity Navigator's methodology page explicitly states that charities "fulfill the expectations of givers when they allocate most of their budgets towards their charitable missions."

to deliver promised impacts. Organizations may be more effective when they spend resources on screening or outreach to target beneficiaries even if such spending ultimately raises the total cost of service delivery.

More broadly, the analysis in this paper highlights the complementary roles played by pilot experimentation and evaluation at scale. A pilot can be considered a proof-of-concept evaluating whether a market failure exists and if remedying it generates returns to beneficiaries. These are necessary but not sufficient conditions for policy success. Evaluation at scale reveals whether the remedy can be sustained as a general policy. Ideally, pilot experimentation would provide insight into the potential for scaleup. However, we would not want to sacrifice evidence about market failures by designing pilots that are too frequently derailed by implementation challenges.

## References

- Akram, Agha Ali, Shyama Chowdhury, and Ahmed Mushfiq Mobarak**, “Effects of Emigration on Rural Labor Markets,” NBER Working Paper 23929, National Bureau of Economic Research 2017.
- Al-Ubaydli, Omar, John A. List, and Dana L. Suskind**, “What Can We Learn from Experiments? Understanding the Threats to the Scalability of Experimental Results,” *American Economic Review*, May 2017, 107 (5), 282–86.
- , **John A List, and Dana Suskind**, “The Science of Using Science: Towards an Understanding of the Threats to Scaling Experiments,” Working Paper 25848, National Bureau of Economic Research May 2019.
- Alatas, Vivi, Arun G Chandrasekhar, Markus Möbius, Benjamin A Olken, and Cindy Paladines**, “When Celebrities Speak: A Nationwide Twitter Experiment Promoting Vaccination In Indonesia,” Working Paper 25589, National Bureau of Economic Research 2019.
- Allcott, Hunt**, “Site Selection Bias in Program Evaluation,” *The Quarterly Journal of Economics*, 2015, 130 (3), 1117–1165.
- Andersen, Simon Calmar and Ulrik Hvidman**, “Implementing Educational Interventions at Scale,” Working Paper 2020-039, Human Capital and Economic Opportunity Working Group May 2020.
- Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin**, “Identification of Causal Effects Using Instrumental Variables,” *Journal of the American Statistical Association*, 1996, 91 (434), 444–455.
- Araujo, M. Caridad, Marta Rubio-Codina, and Norbert Schady**, “70 to 700 to 70,000: Lessons from the Jamaica Experiment,” Working Paper IDB-WP-1230, Inter-American Development Bank April 2021.
- Balán, Pablo, Augustin Bergeron, Gabriel Tourek, and Jonathan L. Weigel**, “Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the Democratic Republic of the Congo,” *American Economic Review*, March 2022, 112 (3), 762–97.
- Banerjee, Abhijit, Raghabendra Chattopadhyay, Esther Duflo, Daniel Keniston, and Nina Singh**, “Improving Police Performance in Rajasthan, India: Experimental Evidence on Incentives, Managerial Autonomy, and Training,” *American Economic Journal: Economic Policy*, 2021, 13 (1), 36–66.
- Banerjee, A.V., S. Chassang, and E. Snowberg**, “Decision Theoretic Approaches to Experiment Design and External Validity,” in Abhijit Vinayak Banerjee and Esther Duflo, eds., *Handbook of Field Experiments*, Vol. 1 of *Handbook of Economic Field Experiments*, North-Holland, 2017, pp. 141–174.
- Bangladesh Bureau of Statistics**, “Preliminary Report on Household Income and Expenditure Survey 2016,” Technical Report, Government of the People’s Republic of Bangladesh 2017.
- Basu, Karna and Maisy Wong**, “Evaluating seasonal food storage and credit programs in east Indonesia,” *Journal of Development Economics*, 2015, 115 (C), 200–216.

- Bellés-Obrero, Cristina and María Lombardi**, “Teacher Performance Pay and Student Learning: Evidence from a Nationwide Program in Peru,” *Economic Development and Cultural Change*, 2022, 70 (4), 1631–1669.
- Bloem, Jeffrey R and Bruce Wydick**, “All i really need to know i learned in kindergarten? evidence from the philippines,” *Economic Development and Cultural Change*, 2023, 71 (2), 753–791.
- Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng’ang’a, and Justin Sandefur**, “Experimental evidence on scaling up education reforms in Kenya,” *Journal of Public Economics*, 2018, 168, 1–20.
- Brooks, Wyatt, Kevin Donovan, and Terence R Johnson**, “From micro to macro in an equilibrium diffusion model,” Technical Report, Working Paper 2020.
- Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak**, “Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh,” *Econometrica*, 2014, 82 (5), 1–43.
- Bulte, Erwin, Gonnie Beekman, Salvatore Di Falco, Joseph Hella, and Pan Lei**, “Behavioral Responses and the Impact of New Agricultural Technologies: Evidence from a Double-blind Field Experiment in Tanzania,” *American Journal of Agricultural Economics*, 2014, 96 (3), 813–830.
- Burke, Marshall, Lauren Falcao Bergquist, and Edward Miguel**, “Sell Low and Buy High: Arbitrage and Local Price Effects in Kenyan Markets\*,” *The Quarterly Journal of Economics*, 2018, 134 (2), 785–842.
- Cameron, Lisa, Susan Olivia, and Manisha Shah**, “Scaling up sanitation: Evidence from an RCT in Indonesia,” *Journal of Development Economics*, 2019, 138, 1–16.
- Chassang, Sylvain, Gerard Padró I Miquel, and Erik Snowberg**, “Selective Trials: A Principal-Agent Approach to Randomized Controlled Experiments,” *American Economic Review*, 2012, 102 (4), 1279–1309.
- Cunha, Jesse M, Giacomo De Giorgi, and Seema Jayachandran**, “The Price Effects of Cash Versus In-Kind Transfers,” *The Review of Economic Studies*, 2018, 86 (1), 240–281.
- Das, Navishti, Andrew Dillon, Srishti Gupta, and Deepika Nagesh**, “Design Choices for an Enumerator Recruitment Experiment,” Research Methods Notes, Global Poverty Research Lab 2024.
- de Ree, Joppe, Karthik Muralidharan, Menno Pradhan, and Halsey Rogers**, “Double for Nothing? Experimental Evidence on an Unconditional Teacher Salary Increase in Indonesia\*,” *The Quarterly Journal of Economics*, 2017, 133 (2), 993–1039.
- DellaVigna, Stefano and Elizabeth Linos**, “RCTs to Scale: Comprehensive Evidence From Two Nudge Units,” *Econometrica*, 2022, 90 (1), 81–116.
- Dhaliwal, Iqbal and Rema Hanna**, “The devil is in the details: The successes and limitations of bureaucratic reform in India,” *Journal of Development Economics*, 2017, 124, 1–21.

- Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan**, “Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India\*,” *The Quarterly Journal of Economics*, 2013, 128 (4), 1499–1545.
- , **Michael Kremer, and Jonathan Robinson**, “Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya,” *American Economic Review*, October 2011, 101 (6), 2350–90.
- Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael W Walker**, “General Equilibrium Effects of Cash Transfers: Experimental Evidence from Kenya,” *Econometrica*, 2022, *forthcoming*.
- Evans, David K. and Fei Yuan**, “How Big Are Effect Sizes in International Education Studies?,” Working Paper 545, Center for Global Development August 2020.
- Fink, Guünther, B. Kelsey Jack, and Felix Masiye**, “Seasonal Liquidity, Rural Labor Markets, and Agricultural Production,” *American Economic Review*, 2020, 110 (11), 3351–92.
- Ganimian, Alejandro J.**, “Growth-Mindset Interventions at Scale: Experimental Evidence From Argentina,” *Educational Evaluation and Policy Analysis*, 2020, 42 (3), 417–438.
- Giné, Xavier, Jessica Goldberg, and Dean Yang**, “Fingerprinting in Malawi: The Challenges of Scaling Up a Complicated Technological Solution in a Resource-Constrained Setting,” Technical Report, Innovations for Poverty Action 2021.
- Global Innovations Fund**, “Does “Sugar Daddies” replicate? The preliminary results are in for Botswana,” Technical Report 2018.
- Hertzberg, Andrew, Jose Maria Liberti, and Daniel Paravisini**, “Information and Incentives Inside the Firm: Evidence from Loan Officer Rotation,” *The Journal of Finance*, 2010, 65 (3), 795–828.
- Kerwin, Jason T. and Rebecca L. Thornton**, “Making the Grade: The Sensitivity of Education Program Effectiveness to Input Choices and Outcome Measures,” *The Review of Economics and Statistics*, 2021, 103 (2), 251–264.
- Khan, Adnan Q., Asim I. Khwaja, and Benjamin A. Olken**, “Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors,” *The Quarterly Journal of Economics*, 2015, 131 (1), 219–271.
- Khandker, Shahidur R.**, “Seasonality of income and poverty in Bangladesh,” *Journal of Development Economics*, 2012, 97 (2), 244–256.
- Khandker, Shahidur R and Wahiduddin Mahmud**, *Seasonal hunger and public policies: evidence from Northwest Bangladesh*, World Bank Publications, 2012.
- Khanna, Gaurav**, “Large-scale education reform in general equilibrium: Regression discontinuity evidence from india,” *Journal of Political Economy*, 2023, 131 (2), 549–591.
- Kraft, Matthew A., David Blazar, and Dylan Hogan**, “The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence,” *Review of Educational Research*, 2018, 88 (4), 547–588.

**List, John A**, *The voltage effect: How to make good ideas great and great ideas scale*, Crown Currency, 2022.

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

**Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar**, “Building State Capacity: Evidence from Biometric Smartcards in India,” *American Economic Review*, 2016, 106 (10), 2895–2929.

—, —, and —, “General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India,” Working Paper 23838, National Bureau of Economic Research 2017.

**Pritchett, Lant and Justin Sandefur**, “Learning from Experiments When Context Matters,” *American Economic Review*, 2015, 105 (5), 471–75.

**Rabb, Nathaniel, Megan Swindal, David Glick, Jake Bowers, Anna Tomasulo, Zayid Oyelami, Kevin H. Wilson, and David Yokum**, “Evidence from a statewide vaccination RCT shows the limits of nudges,” *Nature*, 2022, 604 (7904), E1–E7.

**Rivera, Juan A., Daniela Sotres-Alvarez, Jean-Pierre Habicht, Teresa Shamah, and Salvador Villalpando**, “Impact of the Mexican Program for Education, Health, and Nutrition (Progresa) on Rates of Growth and Anemia in Infants and Young ChildrenA Randomized Effectiveness Study,” *JAMA*, 2004, 291 (21), 2563–2570.

**Rosenzweig, Mark R and Christopher Udry**, “External Validity in a Stochastic World: Evidence from Low-Income Countries,” *The Review of Economic Studies*, 2019, 87 (1), 343–381.

**Schultz, T. Paul**, “School subsidies for the poor: evaluating the Mexican Progresa poverty program,” *Journal of Development Economics*, 2004, 74 (1), 199–250. New Research on Education in Developing Economies.

**Sraer, David and David Thesmar**, “A Sufficient Statistics Approach for Aggregating Firm-Level Experiments,” Working Paper 24208, National Bureau of Economic Research January 2018.

**Stephens, Emma C. and Christopher B. Barrett**, “Incomplete Credit Markets and Commodity Marketing Behaviour,” *Journal of Agricultural Economics*, 2011, 62 (1), 1–24.

**Vivaldi, Eva**, “How Much Can We Generalize From Impact Evaluations?,” *Journal of the European Economic Association*, 2020, 18 (6), 3045–3089.

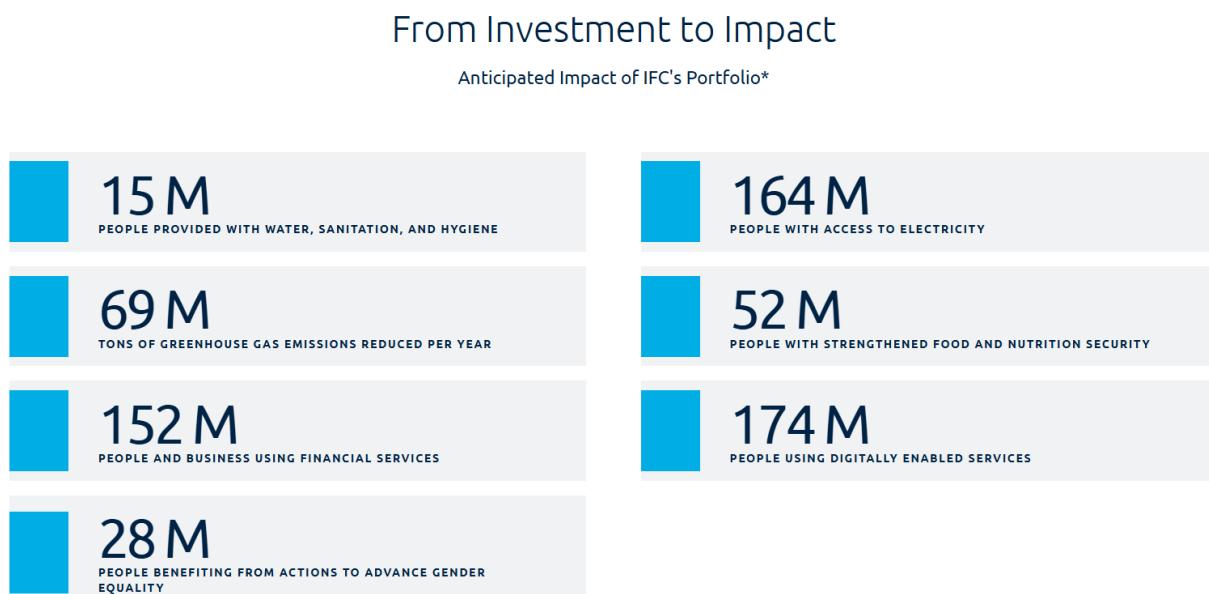
**Wager, Stefan and Susan Athey**, “Estimation and Inference of Heterogeneous Treatment Effects using Random Forests,” *Journal of the American Statistical Association*, 2018, 113 (523), 1228–1242.

## Supplementary Appendix for “Delegation Risk and Implementation at Scale”

### A Service Delivery as a Measure of Impact in Microfinance

Many major microfinance institutions advertise their quantity of services delivery—such as lending volume, number of loans, or number of clients—as a top-level measure of impact. By contrast, it is rare for such organizations to advertise their selectivity by making reference to counterfactual outcomes among their client base or providing evidence that their clients would otherwise be credit constrained. Figure S1 provide examples from websites and annual reports of several well-known microfinance institutions with a focus on economic development. This is not to claim that these institutions ignore targeting, selectivity, or impact. Indeed, many of the organizations provided as examples here create products tailored to specific constraints within the populations they serve and engage in rigorous program evaluation. Nevertheless, it is clear that, in public-facing headline measures of charitable impact, quantity of services delivered features far more prominently than appropriate targeting or reference to financial constraints in the microfinance sector.

Figure S1: Common Headline MFI Performance Metrics



International Finance Corporation. Source: <https://www.ifc.org/en/our-impact>, accessed 12/26/2025

# UNLOCKING POTENTIAL



**145 million**

people globally have accessed BRAC's education, healthcare, and economic development programs to date

[Read more →](#)



**\$6 billion**

disbursed in microloans, primarily to women, in 2024

[Learn more →](#)



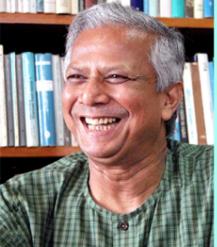
**1 in 5**

people in Bangladesh received support from BRAC in 2024

[Learn more →](#)

BRAC. Source: <https://www.bracusa.org/>, accessed 12/26/2025





**BRANCHES**

2,568

**EMPLOYEES**

24,103

**BORROWER MEMBERS**

10.81 Million

**VILLAGES COVERED (94%)**

81,678

**BDT- 3493.61 Billion | USD- 41.19 Billion**  
Cumulative Disbursed Loan  
(1976 – Nov'25)





Total Number of Borrower Members (In million)

| Period | Members (In million) |
|--------|----------------------|
| NOV'21 | 9.47                 |
| NOV'22 | 10.24                |
| NOV'23 | 10.45                |
| NOV'24 | 10.85                |
| NOV'25 | 10.81                |

Zones in profit

| Zone   | Count |
|--------|-------|
| Zone A | 28    |
| Zone B | 40    |
| Zone C | 40    |
| Zone D | 35    |





Branches in profit

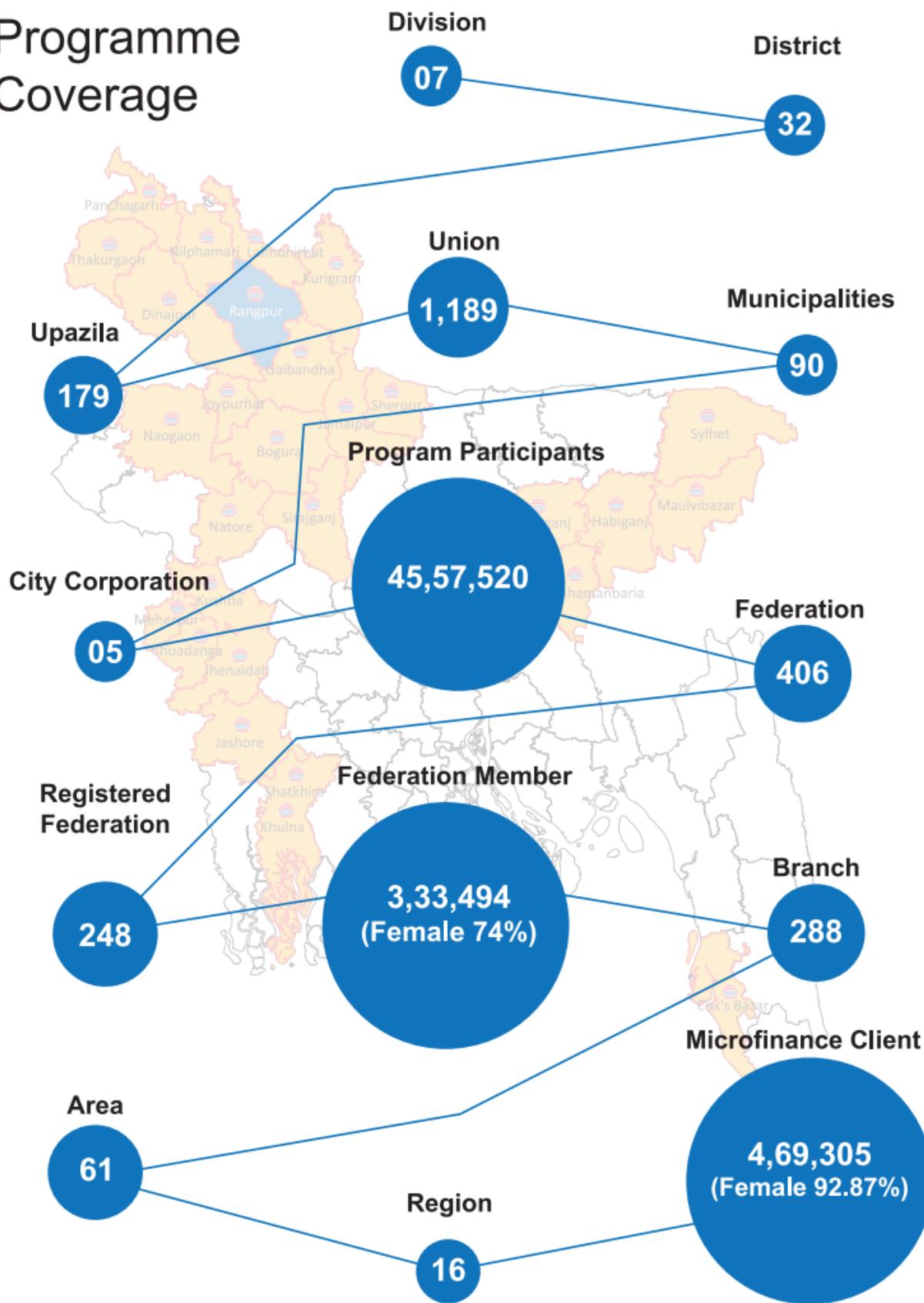
| Branch | Count |
|--------|-------|
| 1719   | 2502  |
| 2502   | 2503  |
| 2503   | 2216  |
| 2216   | 2005  |





Grameen Bank. Source: <https://grameenbank.org.bd/>, accessed 12/26/2025

# Programme Coverage



RDRS. Source: RDRS Bangladesh Annual Report 2023–24

## Financial inclusion of the poor especially women as on 31 March 2024

**177.5 million**

Estimated no. of families covered till 31 March 2024

**83.51%**

Out of total SHGs saving linked - exclusive Women SHGs

**97.04%**

Out of total SHGs – credit linked exclusive Women SHGs

**₹33.5 million**

Average loan amount outstanding/SHG as on as on 31 March 2024

**14.42 million**

Total no. of SHGs savings linked with banks as on 31 March 2024

**5.48 million**

Total no. of SHGs credit linked during the 2023-24

**₹38.2 million**

Average loan amount disbursed/SHG during 2023-24

NABARD Self-Help Group–Bank Linkage Program. Source: National Bank for Agricultural and Rural Development (India) Impact Report 2023–24

### OUR COMMUNITY IMPACT

#### Kivans help people improve their livelihoods



5 million people reached



\$2 billion in loans funded



96% loan repayment rate

Kiva. Source: <https://www.kiva.org/>, accessed 12/26/2025



2.7M

CLIENTS



2,226

BRANCHES



USD555.3m

OLP



2.0%

PAR &gt; 30

ASA International. Source: <https://www.asa-international.com/about-us/at-a-glance/>, accessed 12/26/2025

**850**

Water filters provided through our [water program](#)

**1,541**

Post-fistula women have started a income project

**1,900**

Original cows as living loans in our [cow program](#)

**73%**

[Loan program](#) participants reported increased savings

**391**

Students have stayed in school due to [scholarships](#)

Microfinancing Partners in Africa. Source: <https://microfinancingafrica.org/>, accessed 12/26/2025

## Impact highlights of the Jasiri Gender Bond

**74.3bn:**



of which  
**0.6bn**



The **volume of loans** in TZS disbursed to **women owned/controlled MSMEs**

**Financed** 2 schools and 2 hospitals

**89 SME loans:**



**3116 MSE** **loans:**

**3,205 total number of disbursed loans enabling women in business**

**340 Agri-loans**



The **number of women-owned/controlled agribusinesses financed**  
of which half financed small scale farmers

**209 new accounts**



**61 SMEs employing**  
**>30% women**



Number of **women** who **benefited** from the **Jasiri Gender Bond** financing and **joined NMB bank as new customers** during the allocation period

SME businesses that **employ more than 30% women** were **financed for a total over TZS 10 billion**

National Microfinance Bank, Tanzania. Source: NMB Jasiri Gender Bond Impact Report, 2023

## B Balance Tables

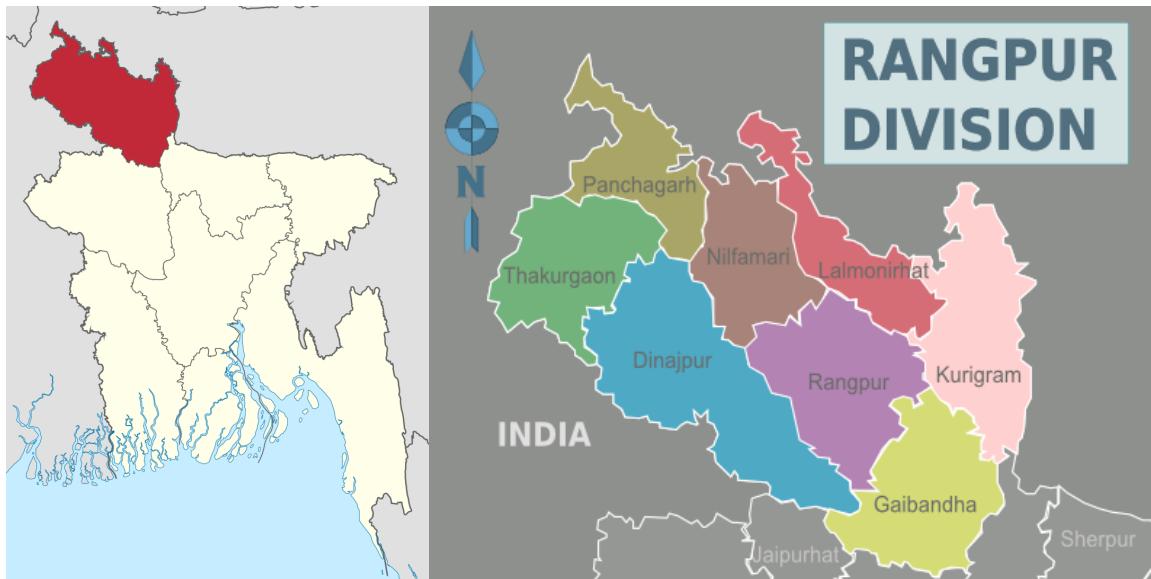
In this section we report balance on baseline covariates across different comparison groups.

### B.1 Balance by Treatment Assignment

Figure S2 presents a map of the region of study. The pilot evaluations in 2008 and 2014 took place in villages randomly sampled from the districts of Kurigram and Lalmonirhat. We divide our analysis of the NLS program at scale between these two original districts, of which we expect the pilots to be statistically representative, and the six expansion districts added in 2017 and 2018.

Tables S1 and S2 report balance on baseline covariates across treatment arms in the pilot districts in 2017 and 2018, respectively, corresponding to the estimates presented in Figure 4. Tables S3 and S4 present comparable statistics for expansion districts, and Tables S5 and S6 pool across all districts.

Figure S2: Map of Rangpur Division, Bangladesh



Source: TUBS (left); James Adams (right) accessed from Wikipedia.

Table S1: Balance on Baseline Covariates across Treatment Arms in 2017 Pilot Districts

| Variable                               | (1)<br>Pure Control<br>Mean/(SE) |                  |                  |                  | (2)<br>Treated<br>Mean/(SE) |  |  |  | (3)<br>Spillover<br>Mean/(SE) |  |  |  | (4)<br>Branch-Control<br>Mean/(SE) |  |  |  | (2)-(1)<br>Mean difference |  |  |  | (3)-(1)<br>Pairwise t-test |  |  |  | (4)-(1)<br>Mean difference |  |  |        |        |
|--|----------------------------------|------------------|------------------|------------------|-----------------------------|--|--|--|-------------------------------|--|--|--|------------------------------------|--|--|--|----------------------------|--|--|--|----------------------------|--|--|--|----------------------------|--|--|--------|--------|
|  |                                  |                  |                  |                  |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  |        |        |
| Migrated >3 years ago or never         | 0.437<br>(0.036)                 | 0.604<br>(0.052) | 0.439<br>(0.046) | 0.528<br>(0.048) |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | 0.001  | 0.091  |
| Migrated 2–3 years ago                 | 0.100<br>(0.020)                 | 0.040<br>(0.023) | 0.151<br>(0.030) | 0.128<br>(0.031) |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | 0.051  | 0.028  |
| Migrated a year ago                    | 0.462<br>(0.040)                 | 0.356<br>(0.058) | 0.410<br>(0.048) | 0.343<br>(0.056) |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | -0.106 | -0.119 |
| No land owned                          | 0.630<br>(0.041)                 | 0.695<br>(0.050) | 0.770<br>(0.043) | 0.638<br>(0.062) |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | 0.064  | 0.008  |
| Below median land   $\leq 0.5$ acres   | 0.206<br>(0.023)                 | 0.176<br>(0.032) | 0.137<br>(0.029) | 0.192<br>(0.044) |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | -0.030 | -0.069 |
| Above median land   $\leq 0.5$ acres   | 0.164<br>(0.029)                 | 0.129<br>(0.026) | 0.094<br>(0.026) | 0.170<br>(0.053) |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | -0.035 | -0.070 |
| Some education                         | 0.459<br>(0.031)                 | 0.458<br>(0.043) | 0.371<br>(0.047) | 0.366<br>(0.046) |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | -0.000 | -0.088 |
| No. of adult males in household        | 1.539<br>(0.036)                 | 1.502<br>(0.050) | 1.561<br>(0.061) | 1.479<br>(0.069) |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | -0.038 | -0.022 |
| No. of children in household           | 1.526<br>(0.052)                 | 1.416<br>(0.062) | 1.406<br>(0.097) | 1.328<br>(0.064) |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | -0.110 | -0.120 |
| Baseline food insecurity               | 0.617<br>(0.061)                 | 0.687<br>(0.083) | 0.507<br>(0.072) | 0.498<br>(0.077) |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | 0.071  | -0.109 |
| Distance to village center (km)        | 0.376<br>(0.062)                 | 0.340<br>(0.056) | 0.292<br>(0.036) | 0.390<br>(0.066) |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | -0.036 | -0.084 |
| Dist. to MFI branch office (km)        | 3.242<br>(0.391)                 | 2.639<br>(0.426) | 3.112<br>(0.491) | 3.379<br>(0.357) |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | -0.603 | -0.130 |
| Number of observations                 | 519                              | 550              | 278              | 265              |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | 1069   | 797    |
| Number of clusters                     | 32                               | 16               | 16               | 16               |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | 48     | 48     |
| F-test of joint significance (p-value) |                                  |                  |                  |                  |                             |  |  |  |                               |  |  |  |                                    |  |  |  |                            |  |  |  |                            |  |  |  |                            |  |  | 0.003  | 0.333  |

Notes: Standard errors clustered at the village level in parentheses. \*( $p < 0.10$ ), \*\*( $p < 0.05$ ), \*\*\*( $p < 0.01$ )

Table S2: Balance on Baseline Covariates across Treatment Arms in 2018 Pilot Districts

| Variable                               | (1)<br>Pure Control<br>Mean/(SE) | (2)<br>Treated<br>Mean/(SE) | (3)<br>Spillover<br>Mean/(SE) | (4)<br>Branch-Control<br>Mean/(SE) | (2)-(1)<br>Mean difference | (3)-(1)<br>Pairwise t-test<br>Mean difference | (4)-(1)<br>Mean difference |
|--|----------------------------------|-----------------------------|-------------------------------|------------------------------------|----------------------------|---|----------------------------|
| Migrated >3 years ago or never         | 0.382<br>(0.027)                 | 0.279<br>(0.023)            | 0.315<br>(0.043)              | 0.287<br>(0.028)                   | -0.103**                   | -0.068  | -0.095*                    |
| Migrated 2–3 years ago                 | 0.102<br>(0.013)                 | 0.110<br>(0.017)            | 0.100<br>(0.024)              | 0.110<br>(0.029)                   | 0.009                      | -0.002  | 0.008                      |
| Migrated a year ago                    | 0.516<br>(0.033)                 | 0.610<br>(0.027)            | 0.586<br>(0.059)              | 0.603<br>(0.039)                   | 0.094*                     | 0.070   | 0.087                      |
| No land owned                          | 0.844<br>(0.015)                 | 0.835<br>(0.019)            | 0.825<br>(0.025)              | 0.851<br>(0.035)                   | -0.009                     | -0.019  | 0.007                      |
| Below median land   ≤0.5 acres         | 0.077<br>(0.010)                 | 0.086<br>(0.013)            | 0.096<br>(0.022)              | 0.096<br>(0.022)                   | 0.009                      | 0.018   | 0.019                      |
| Above median land   ≤0.5 acres         | 0.079<br>(0.012)                 | 0.079<br>(0.013)            | 0.080<br>(0.020)              | 0.053<br>(0.015)                   | 0.000                      | 0.001   | -0.026                     |
| Some education                         | 0.449<br>(0.031)                 | 0.405<br>(0.026)            | 0.299<br>(0.030)              | 0.330<br>(0.031)                   | -0.044                     | -0.150***                                     | -0.119**                   |
| No. of adult males in household        | 1.561<br>(0.042)                 | 1.504<br>(0.038)            | 1.474<br>(0.062)              | 1.429<br>(0.045)                   | -0.058                     | -0.087  | -0.132*                    |
| No. of children in household           | 1.544<br>(0.056)                 | 1.585<br>(0.063)            | 1.470<br>(0.097)              | 1.525<br>(0.096)                   | 0.041                      | -0.074  | -0.019                     |
| Baseline food insecurity               | 0.732<br>(0.029)                 | 0.799<br>(0.022)            | 0.773<br>(0.050)              | 0.730<br>(0.058)                   | 0.068                      | 0.041   | -0.001                     |
| Borrowed money at baseline             | 0.807<br>(0.023)                 | 0.822<br>(0.022)            | 0.769<br>(0.045)              | 0.901<br>(0.023)                   | 0.015                      | -0.038  | 0.094**                    |
| Village flooding in past year          | 0.498<br>(0.083)                 | 0.486<br>(0.074)            | 0.709<br>(0.126)              | 0.670<br>(0.118)                   | -0.012                     | 0.211   | 0.172                      |
| Distance to village center (km)        | 0.292<br>(0.027)                 | 0.305<br>(0.024)            | 0.315<br>(0.064)              | 0.326<br>(0.057)                   | 0.013                      | 0.023   | 0.034                      |
| Dist. to MFI branch office (km)        | 3.228<br>(0.336)                 | 2.846<br>(0.301)            | 3.219<br>(0.599)              | 3.714<br>(0.459)                   | -0.381                     | -0.008  | 0.487                      |
| Number of observations                 | 570<br>40                        | 798<br>49                   | 251<br>15                     | 282<br>17                          | 1368<br>89                 | 821<br>55                                     | 852<br>57                  |
| Number of clusters                     |                                  |                             |                               |                                    |                            |   |                            |
| F-test of joint significance (p-value) |                                  |                             |                               |                                    | 0.250                      | 0.014   | 0.002                      |

Notes: Standard errors clustered at the village level in parentheses. \*( $p < 0.10$ ), \*\*( $p < 0.05$ ), \*\*\*( $p < 0.01$ )

Table S3: Balance on Baseline Covariates across Treatment Arms in 2017 Expansion Districts

| Variable                               | (1)<br>Pure Control<br>Mean/(SE) |                      | (2)<br>Treated<br>Mean/(SE) |                             | (3)<br>Spillover<br>Mean/(SE) |                 | (4)<br>Branch-Control<br>Mean/(SE) |                 | (2)-(1)<br>Mean difference |                 | (3)-(1)<br>Pairwise t-test |                 | (4)-(1)<br>Mean difference |                 |
|--|----------------------------------|----------------------|-----------------------------|-----------------------------|-------------------------------|-----------------|------------------------------------|-----------------|----------------------------|-----------------|----------------------------|-----------------|----------------------------|-----------------|
|  | Pure Control<br>Mean/(SE)        | Treated<br>Mean/(SE) | Spillover<br>Mean/(SE)      | Branch-Control<br>Mean/(SE) | Mean difference               | Mean difference | Mean difference                    | Mean difference | Pairwise t-test            | Mean difference | Pairwise t-test            | Mean difference | Pairwise t-test            | Mean difference |
| Migrated >3 years ago or never         | 0.437<br>(0.036)                 | 0.604<br>(0.052)     | 0.439<br>(0.046)            | 0.528<br>(0.048)            | 0.166*                        | 0.001           |                                    |                 |                            |                 |                            |                 |                            | 0.091           |
| Migrated 2–3 years ago                 | 0.100<br>(0.020)                 | 0.040<br>(0.023)     | 0.151<br>(0.030)            | 0.128<br>(0.031)            | -0.060*                       | 0.051           |                                    |                 |                            |                 |                            |                 |                            | 0.028           |
| Migrated a year ago                    | 0.462<br>(0.040)                 | 0.356<br>(0.058)     | 0.410<br>(0.048)            | 0.343<br>(0.056)            | -0.106                        | -0.052          |                                    |                 |                            |                 |                            |                 |                            | -0.119          |
| No land owned                          | 0.630<br>(0.041)                 | 0.695<br>(0.050)     | 0.770<br>(0.043)            | 0.638<br>(0.062)            | 0.064                         | 0.140*          |                                    |                 |                            |                 |                            |                 |                            | 0.008           |
| Below median land   $\leq 0.5$ acres   | 0.206<br>(0.023)                 | 0.176<br>(0.032)     | 0.137<br>(0.029)            | 0.192<br>(0.044)            | -0.030                        | -0.069          |                                    |                 |                            |                 |                            |                 |                            | -0.014          |
| Above median land   $\leq 0.5$ acres   | 0.164<br>(0.029)                 | 0.129<br>(0.026)     | 0.094<br>(0.026)            | 0.170<br>(0.053)            | -0.035                        | -0.070          |                                    |                 |                            |                 |                            |                 |                            | 0.006           |
| Some education                         | 0.459<br>(0.031)                 | 0.458<br>(0.043)     | 0.371<br>(0.047)            | 0.366<br>(0.046)            | -0.000                        | -0.088          |                                    |                 |                            |                 |                            |                 |                            | -0.093          |
| No. of adult males in household        | 1.539<br>(0.036)                 | 1.502<br>(0.050)     | 1.561<br>(0.061)            | 1.479<br>(0.069)            | -0.038                        | 0.022           |                                    |                 |                            |                 |                            |                 |                            | -0.060          |
| No. of children in household           | 1.526<br>(0.052)                 | 1.416<br>(0.062)     | 1.406<br>(0.097)            | 1.328<br>(0.064)            | -0.110                        | -0.120          |                                    |                 |                            |                 |                            |                 |                            | -0.198*         |
| Baseline food insecurity               | 0.617<br>(0.061)                 | 0.687<br>(0.083)     | 0.507<br>(0.072)            | 0.498<br>(0.077)            | 0.071                         | -0.109          |                                    |                 |                            |                 |                            |                 |                            | -0.118          |
| Distance to village center (km)        | 0.376<br>(0.062)                 | 0.340<br>(0.056)     | 0.292<br>(0.036)            | 0.390<br>(0.066)            | -0.036                        | -0.084          |                                    |                 |                            |                 |                            |                 |                            | 0.014           |
| Dist. to MFI branch office (km)        | 3.242<br>(0.391)                 | 2.639<br>(0.426)     | 3.112<br>(0.491)            | 3.379<br>(0.357)            | -0.603                        | -0.130          |                                    |                 |                            |                 |                            |                 |                            | 0.137           |
| Number of observations                 | 519                              | 550                  | 278                         | 265                         | 1069                          | 797             |                                    |                 |                            |                 |                            |                 |                            | 784             |
| Number of clusters                     | 32                               | 16                   | 16                          | 16                          | 48                            | 48              |                                    |                 |                            |                 |                            |                 |                            | 48              |
| F-test of joint significance (p-value) |                                  |                      |                             |                             | 0.003                         | 0.333           |                                    |                 |                            |                 |                            |                 |                            | 0.064           |

Notes: Standard errors clustered at the village level in parentheses. \*( $p < 0.10$ ), \*\*( $p < 0.05$ ), \*\*\*( $p < 0.01$ )

Table S4: Balance on Baseline Covariates across Treatment Arms in 2018 Expansion Districts

| Variable                               | (1)<br>Pure Control<br>Mean/(SE) | (2)<br>Treated<br>Mean/(SE) | (3)<br>Spillover<br>Mean/(SE) | (4)<br>Branch-Control<br>Mean/(SE) | (2)-(1)<br>Mean difference | (3)-(1)<br>Pairwise t-test<br>Mean difference | (4)-(1)<br>Mean difference |
|--|----------------------------------|-----------------------------|-------------------------------|------------------------------------|----------------------------|---|----------------------------|
| Migrated >3 years ago or never         | 0.382<br>(0.027)                 | 0.279<br>(0.023)            | 0.315<br>(0.043)              | 0.287<br>(0.028)                   | -0.103**                   | -0.068  | -0.095*                    |
| Migrated 2–3 years ago                 | 0.102<br>(0.013)                 | 0.110<br>(0.017)            | 0.100<br>(0.024)              | 0.110<br>(0.029)                   | 0.009                      | -0.002  | 0.008                      |
| Migrated a year ago                    | 0.516<br>(0.033)                 | 0.610<br>(0.027)            | 0.586<br>(0.059)              | 0.603<br>(0.039)                   | 0.094*                     | 0.070   | 0.087                      |
| No land owned                          | 0.844<br>(0.015)                 | 0.835<br>(0.019)            | 0.825<br>(0.025)              | 0.851<br>(0.035)                   | -0.009                     | -0.019  | 0.007                      |
| Below median land   ≤0.5 acres         | 0.077<br>(0.010)                 | 0.086<br>(0.013)            | 0.096<br>(0.022)              | 0.096<br>(0.022)                   | 0.009                      | 0.018   | 0.019                      |
| Above median land   ≤0.5 acres         | 0.079<br>(0.012)                 | 0.079<br>(0.013)            | 0.080<br>(0.020)              | 0.053<br>(0.015)                   | 0.000                      | 0.001   | -0.026                     |
| Some education                         | 0.449<br>(0.031)                 | 0.405<br>(0.026)            | 0.299<br>(0.030)              | 0.330<br>(0.031)                   | -0.044                     | -0.150***                                     | -0.119**                   |
| No. of adult males in household        | 1.561<br>(0.042)                 | 1.504<br>(0.038)            | 1.474<br>(0.062)              | 1.429<br>(0.045)                   | -0.058                     | -0.087  | -0.132*                    |
| No. of children in household           | 1.544<br>(0.056)                 | 1.585<br>(0.063)            | 1.470<br>(0.097)              | 1.525<br>(0.096)                   | 0.041                      | -0.074  | -0.019                     |
| Baseline food insecurity               | 0.732<br>(0.029)                 | 0.799<br>(0.022)            | 0.773<br>(0.050)              | 0.730<br>(0.058)                   | 0.068                      | 0.041   | -0.001                     |
| Borrowed money at baseline             | 0.807<br>(0.023)                 | 0.822<br>(0.022)            | 0.769<br>(0.045)              | 0.901<br>(0.023)                   | 0.015                      | -0.038  | 0.094**                    |
| Village flooding in past year          | 0.498<br>(0.083)                 | 0.486<br>(0.074)            | 0.709<br>(0.126)              | 0.670<br>(0.118)                   | -0.012                     | 0.211   | 0.172                      |
| Distance to village center (km)        | 0.292<br>(0.027)                 | 0.305<br>(0.024)            | 0.315<br>(0.064)              | 0.326<br>(0.057)                   | 0.013                      | 0.023   | 0.034                      |
| Dist. to MFI branch office (km)        | 3.228<br>(0.336)                 | 2.846<br>(0.301)            | 3.219<br>(0.599)              | 3.714<br>(0.459)                   | -0.381                     | -0.008  | 0.487                      |
| Number of observations                 | 570<br>40                        | 798<br>49                   | 251<br>15                     | 282<br>17                          | 1368<br>89                 | 821<br>55                                     | 852<br>57                  |
| Number of clusters                     |                                  |                             |                               |                                    |                            |   |                            |
| F-test of joint significance (p-value) |                                  |                             |                               |                                    | 0.250                      | 0.014   | 0.002                      |

Notes: Standard errors clustered at the village level in parentheses. \*( $p < 0.10$ ), \*\*( $p < 0.05$ ), \*\*\*( $p < 0.01$ )

Table S5: Balance on Baseline Covariates across Treatment Arms in 2017 All Districts

| Variable                               | (1)<br>Pure Control<br>Mean/(SE) |                      | (2)<br>Treated<br>Mean/(SE) |                             | (3)<br>Spillover<br>Mean/(SE) |                 | (4)<br>Branch-Control<br>Mean/(SE) |                 | (2)-(1)<br>Mean difference |                 | (3)-(1)<br>Pairwise t-test<br>Mean difference |                 | (4)-(1)<br>Mean difference |                 |
|--|----------------------------------|----------------------|-----------------------------|-----------------------------|-------------------------------|-----------------|------------------------------------|-----------------|----------------------------|-----------------|---|-----------------|----------------------------|-----------------|
|  | Pure Control<br>Mean/(SE)        | Treated<br>Mean/(SE) | Spillover<br>Mean/(SE)      | Branch-Control<br>Mean/(SE) | Mean difference               | Mean difference | Mean difference                    | Mean difference | Mean difference            | Mean difference | Mean difference                               | Mean difference | Mean difference            | Mean difference |
| Migrated >3 years ago or never         | 0.489<br>(0.027)                 | 0.676<br>(0.033)     | 0.528<br>(0.035)            | 0.534<br>(0.031)            | 0.186***                      | 0.039           |                                    |                 |                            |                 |   |                 | 0.045                      |                 |
| Migrated 2–3 years ago                 | 0.096<br>(0.013)                 | 0.060<br>(0.015)     | 0.129<br>(0.017)            | 0.095<br>(0.017)            | -0.036                        | 0.033           |                                    |                 |                            |                 |   |                 | -0.001                     |                 |
| Migrated a year ago                    | 0.414<br>(0.029)                 | 0.264<br>(0.035)     | 0.343<br>(0.034)            | 0.371<br>(0.036)            | -0.150**                      | -0.071          |                                    |                 |                            |                 |   |                 | -0.044                     |                 |
| No land owned                          | 0.551<br>(0.031)                 | 0.626<br>(0.032)     | 0.644<br>(0.035)            | 0.573<br>(0.040)            | 0.075                         | 0.093*          |                                    |                 |                            |                 |   |                 | 0.022                      |                 |
| Below median land   $\leq 0.5$ acres   | 0.247<br>(0.019)                 | 0.170<br>(0.019)     | 0.187<br>(0.021)            | 0.206<br>(0.026)            | -0.077**                      | -0.060*         |                                    |                 |                            |                 |   |                 | -0.041                     |                 |
| Above median land   $\leq 0.5$ acres   | 0.202<br>(0.022)                 | 0.204<br>(0.021)     | 0.169<br>(0.022)            | 0.221<br>(0.020)            | 0.002                         | -0.033          |                                    |                 |                            |                 |   |                 | 0.019                      |                 |
| Some education                         | 0.459<br>(0.019)                 | 0.445<br>(0.023)     | 0.444<br>(0.030)            | 0.405<br>(0.030)            | -0.014                        | -0.015          |                                    |                 |                            |                 |   |                 | -0.055                     |                 |
| No. of adult males in household        | 1.559<br>(0.025)                 | 1.439<br>(0.030)     | 1.530<br>(0.040)            | 1.490<br>(0.038)            | -0.120**                      | -0.029          |                                    |                 |                            |                 |   |                 | -0.069                     |                 |
| No. of children in household           | 1.514<br>(0.036)                 | 1.411<br>(0.037)     | 1.490<br>(0.064)            | 1.423<br>(0.054)            | -0.104*                       | -0.025          |                                    |                 |                            |                 |   |                 | -0.091                     |                 |
| Baseline food insecurity               | 0.550<br>(0.040)                 | 0.705<br>(0.048)     | 0.525<br>(0.044)            | 0.493<br>(0.047)            | 0.155*                        | -0.025          |                                    |                 |                            |                 |   |                 | -0.057                     |                 |
| Distance to village center (km)        | 0.376<br>(0.033)                 | 0.400<br>(0.036)     | 0.306<br>(0.023)            | 0.356<br>(0.034)            | 0.025                         | -0.070          |                                    |                 |                            |                 |   |                 | -0.019                     |                 |
| Dist. to MFI branch office (km)        | 3.634<br>(0.287)                 | 3.344<br>(0.288)     | 3.616<br>(0.328)            | 3.809<br>(0.270)            | -0.290                        | -0.018          |                                    |                 |                            |                 |   |                 | 0.175                      |                 |
| Number of observations                 | 1093                             | 1276                 | 621                         | 588                         | 2369                          | 1714            |                                    |                 |                            |                 |   |                 | 1681                       |                 |
| Number of clusters                     | 70                               | 40                   | 40                          | 40                          | 110                           | 110             |                                    |                 |                            |                 |   |                 | 110                        |                 |
| F-test of joint significance (p-value) |                                  |                      |                             |                             | 0.000                         | 0.182           |                                    |                 |                            |                 |   |                 | 0.189                      |                 |

Notes: Standard errors clustered at the village level in parentheses. \*( $p < 0.10$ ), \*\*( $p < 0.05$ ), \*\*\*( $p < 0.01$ )

Table S6: Balance on Baseline Covariates across Treatment Arms in 2018 All Districts

| Variable                               | (1)<br>Pure Control<br>Mean/(SE) | (2)<br>Treated<br>Mean/(SE) | (3)<br>Spillover<br>Mean/(SE) | (4)<br>Branch-Control<br>Mean/(SE) | (2)-(1)<br>Mean difference | (3)-(1)<br>Pairwise t-test<br>Mean difference | (4)-(1)<br>Mean difference |
|--|----------------------------------|-----------------------------|-------------------------------|------------------------------------|----------------------------|---|----------------------------|
| Migrated >3 years ago or never         | 0.398<br>(0.019)                 | 0.308<br>(0.019)            | 0.335<br>(0.031)              | 0.338<br>(0.033)                   | -0.091***                  | -0.063  | -0.060                     |
| Migrated 2–3 years ago                 | 0.121<br>(0.010)                 | 0.112<br>(0.011)            | 0.123<br>(0.016)              | 0.104<br>(0.016)                   | -0.009                     | 0.002   | -0.017                     |
| Migrated a year ago                    | 0.481<br>(0.020)                 | 0.580<br>(0.023)            | 0.541<br>(0.038)              | 0.558<br>(0.036)                   | 0.099**                    | 0.061   | 0.078                      |
| No land owned                          | 0.825<br>(0.011)                 | 0.818<br>(0.012)            | 0.796<br>(0.019)              | 0.809<br>(0.024)                   | -0.007                     | -0.029  | -0.016                     |
| Below median land   ≤0.5 acres         | 0.095<br>(0.008)                 | 0.096<br>(0.009)            | 0.106<br>(0.013)              | 0.119<br>(0.016)                   | 0.001                      | 0.011   | 0.024                      |
| Above median land   ≤0.5 acres         | 0.080<br>(0.008)                 | 0.086<br>(0.009)            | 0.098<br>(0.015)              | 0.072<br>(0.012)                   | 0.006                      | 0.018   | -0.008                     |
| Some education                         | 0.451<br>(0.017)                 | 0.433<br>(0.017)            | 0.373<br>(0.025)              | 0.369<br>(0.023)                   | -0.018                     | -0.078*                                       | -0.082**                   |
| No. of adult males in household        | 1.525<br>(0.025)                 | 1.496<br>(0.024)            | 1.480<br>(0.037)              | 1.518<br>(0.033)                   | -0.029                     | -0.045  | -0.007                     |
| No. of children in household           | 1.554<br>(0.036)                 | 1.546<br>(0.040)            | 1.513<br>(0.055)              | 1.509<br>(0.050)                   | -0.007                     | -0.040  | -0.044                     |
| Baseline food insecurity               | 0.731<br>(0.020)                 | 0.760<br>(0.019)            | 0.738<br>(0.034)              | 0.648<br>(0.041)                   | 0.028                      | 0.007   | -0.083                     |
| Borrowed money at baseline             | 0.806<br>(0.014)                 | 0.809<br>(0.016)            | 0.785<br>(0.028)              | 0.849<br>(0.020)                   | 0.003                      | -0.021  | 0.043                      |
| Village flooding in past year          | 0.322<br>(0.050)                 | 0.434<br>(0.051)            | 0.521<br>(0.084)              | 0.462<br>(0.082)                   | 0.112                      | 0.199*  | 0.139                      |
| Distance to village center (km)        | 0.342<br>(0.023)                 | 0.343<br>(0.017)            | 0.351<br>(0.034)              | 0.375<br>(0.037)                   | 0.001                      | 0.009   | 0.033                      |
| Dist. to MFI branch office (km)        | 3.662<br>(0.202)                 | 3.286<br>(0.212)            | 3.562<br>(0.326)              | 3.918<br>(0.313)                   | -0.376                     | -0.100  | 0.256                      |
| Number of observations                 | 1411                             | 1618                        | 641                           | 654                                | 3029                       | 2052  | 2065                       |
| Number of clusters                     | 100                              | 99                          | 39                            | 40                                 | 199                        | 139   | 140                        |
| F-test of joint significance (p-value) |                                  |                             |                               |                                    | 0.173                      | 0.095   | 0.003                      |

Notes: Standard errors clustered at the village level in parentheses. \*( $p < 0.10$ ), \*\*( $p < 0.05$ ), \*\*\*( $p < 0.01$ )

## B.2 Balance across Geography

Tables S7 and S8 report balance on baseline covariates between pilot (Kurigram and Lalmonirhat) and expansion (Dinajpur, Gaibandha, Nilfamari, Panchagarh, Rangpur, and Thakurgaon) in 2017 and 2018, respectively, corresponding to estimates presented in Figure 8.

Table S7: Balance on Baseline Covariates across Districts in 2017

| Variable                               | (1)<br>Pilot<br>Mean/(SE) | (2)<br>Expansion<br>Mean/(SE) | (2)-(1)<br>Pairwise t-test<br>Mean difference |
|--|---------------------------|-------------------------------|---|
| Migrated >3 years ago or never         | 0.620<br>(0.024)          | 0.509<br>(0.026)              | -0.110**                                      |
| Migrated 2–3 years ago                 | 0.085<br>(0.010)          | 0.093<br>(0.013)              | 0.008   |
| Migrated a year ago                    | 0.295<br>(0.023)          | 0.398<br>(0.027)              | 0.103**                                       |
| No land owned                          | 0.532<br>(0.022)          | 0.677<br>(0.025)              | 0.145***                                      |
| Below median land   ≤0.5 acres         | 0.219<br>(0.015)          | 0.182<br>(0.016)              | -0.037  |
| Above median land   ≤0.5 acres         | 0.249<br>(0.015)          | 0.141<br>(0.016)              | -0.108***                                     |
| Some education                         | 0.455<br>(0.014)          | 0.428<br>(0.021)              | -0.027  |
| No. of adult males in household        | 1.483<br>(0.021)          | 1.520<br>(0.026)              | 0.038   |
| No. of children in household           | 1.477<br>(0.030)          | 1.435<br>(0.034)              | -0.041  |
| Baseline food insecurity               | 0.583<br>(0.031)          | 0.602<br>(0.039)              | 0.019   |
| Distance to village center (km)        | 0.384<br>(0.020)          | 0.351<br>(0.030)              | -0.032  |
| Dist. to MFI branch office (km)        | 3.982<br>(0.203)          | 3.037<br>(0.217)              | -0.946**                                      |
| Number of observations                 | 1966                      | 1612                          | 3578  |
| Number of clusters                     | 110                       | 80                            | 190   |
| F-test of joint significance (p-value) |                           |                               | 0.000   |

Notes: Standard errors clustered at the village level in parentheses. \*( $p < 0.10$ ), \*\*( $p < 0.05$ ), \*\*\*( $p < 0.01$ )

Table S8: Balance on Baseline Covariates across Districts in 2018

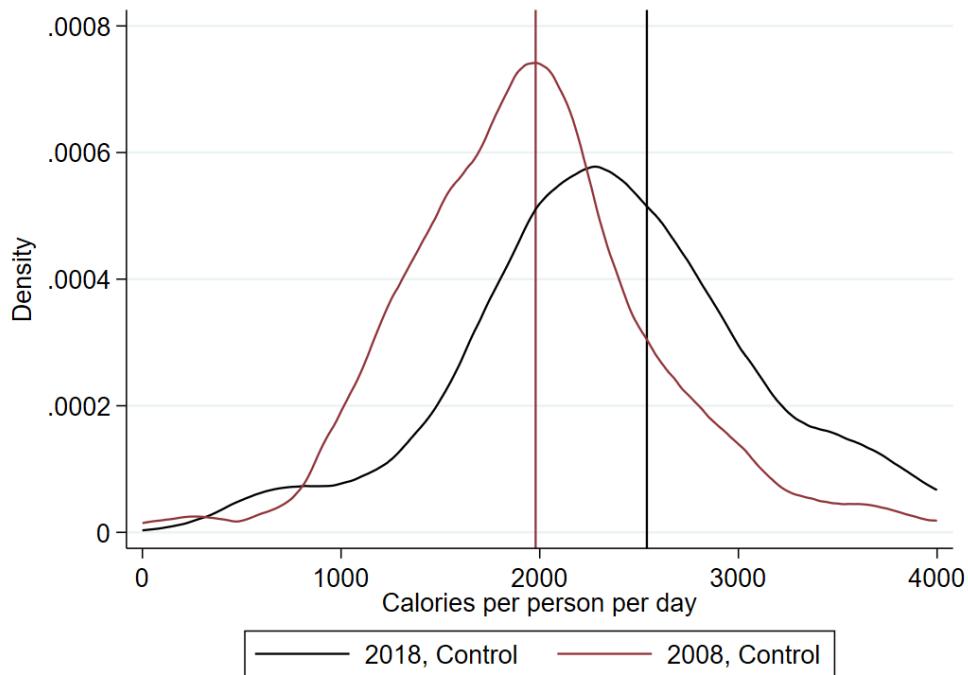
| Variable                               | (1)<br>Pilot<br>Mean/(SE) | (2)<br>Expansion<br>Mean/(SE) | (2)-(1)<br>Pairwise t-test<br>Mean difference |
|--|---------------------------|-------------------------------|---|
| Migrated >3 years ago or never         | 0.369<br>(0.017)          | 0.316<br>(0.015)              | -0.053*                                       |
| Migrated 2–3 years ago                 | 0.123<br>(0.008)          | 0.106<br>(0.010)              | -0.017  |
| Migrated a year ago                    | 0.508<br>(0.019)          | 0.578<br>(0.018)              | 0.070**                                       |
| No land owned                          | 0.798<br>(0.010)          | 0.839<br>(0.011)              | 0.041**                                       |
| Below median land   $\leq 0.5$ acres   | 0.112<br>(0.007)          | 0.086<br>(0.008)              | -0.026*                                       |
| Above median land   $\leq 0.5$ acres   | 0.090<br>(0.007)          | 0.075<br>(0.007)              | -0.015  |
| Some education                         | 0.441<br>(0.012)          | 0.393<br>(0.016)              | -0.048*                                       |
| No. of adult males in household        | 1.507<br>(0.018)          | 1.506<br>(0.023)              | -0.001  |
| No. of children in household           | 1.530<br>(0.027)          | 1.549<br>(0.036)              | 0.019   |
| Baseline food insecurity               | 0.703<br>(0.018)          | 0.765<br>(0.017)              | 0.063*  |
| Borrowed money at baseline             | 0.801<br>(0.013)          | 0.822<br>(0.014)              | 0.021   |
| Village flooding in past year          | 0.312<br>(0.039)          | 0.547<br>(0.047)              | 0.235***                                      |
| Distance to village center (km)        | 0.382<br>(0.017)          | 0.305<br>(0.017)              | -0.077**                                      |
| Dist. to MFI branch office (km)        | 3.865<br>(0.156)          | 3.139<br>(0.193)              | -0.726**                                      |
| Number of observations                 | 2423                      | 1901                          | 4324  |
| Number of clusters                     | 157                       | 121                           | 278   |
| F-test of joint significance (p-value) |                           |                               | 0.000   |

Notes: Standard errors clustered at the village level in parentheses. \*( $p < 0.10$ ), \*\*( $p < 0.05$ ), \*\*\*( $p < 0.01$ )

### B.3 Changes over Time

Figure S3 plots the distribution of calorie consumption at endline among experimentally untreated households in 2008 and 2018. There is a clear shift to the right over time, reflecting a secular trend of greater calorie consumption in the region overall.

Figure S3: Distribution of Per Capita Calorie Consumption by Year



Notes: Data from household endline calorie consumption per capita among households assigned to control in 2008 and 2018. Vertical lines represent group means.

Table S9 reports balance within pilot districts (Kurigram and Lalmonirhat) on baseline covariates that are present in both 2008 and 2018, used in the machine learning algorithm to predict conditional treatment effects. Households in the 2018 evaluation are generally larger and more educated, though the rate of land ownership among loan-eligible participants is lower than in 2008.

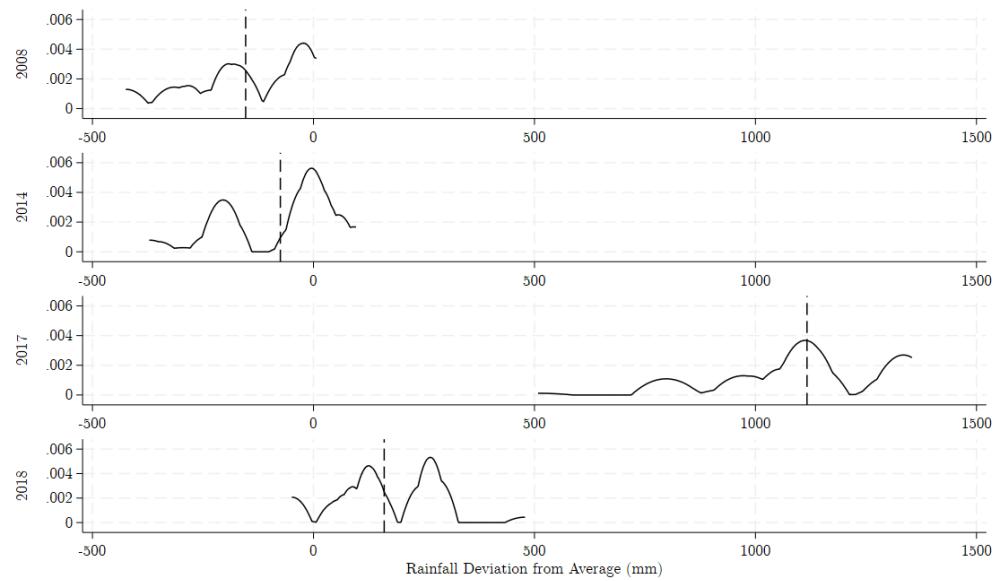
Figure S4 reports the distribution of rainfall deviations from average, as described in Appendix E, across villages in each study year. There is more rain across the board in 2017 and 2018 compared to the pilot years, especially in 2017 when the region recorded historically high rates of flooding. As a result, there is little overlapping support in realizations of rainfall across the study years.

Table S9: Balance on Baseline Covariates between 2008 and 2018

| Variable                             | (1)<br>2008<br>Mean/(SE) | (2)<br>2018<br>Mean/(SE) | (1)-(2)<br>Pairwise t-test<br>Mean difference |
|--------------------------------------|--------------------------|--------------------------|---|
| Dist. to MFI branch office (km)      | 3.860<br>(0.332)         | 3.136<br>(0.193)         | 0.724*  |
| Some education                       | 0.240<br>(0.011)         | 0.393<br>(0.016)         | -0.153***                                     |
| Baseline food insecurity             | 1.285<br>(0.037)         | 0.765<br>(0.017)         | 0.519***                                      |
| Borrowed money at baseline           | 0.664<br>(0.018)         | 0.822<br>(0.014)         | -0.158***                                     |
| No land owned                        | 0.595<br>(0.032)         | 0.839<br>(0.011)         | -0.244***                                     |
| Below median land   $\leq 0.5$ acres | 0.116<br>(0.011)         | 0.086<br>(0.008)         | 0.029**                                       |
| Above median land   $\leq 0.5$ acres | 0.290<br>(0.033)         | 0.075<br>(0.007)         | 0.214***                                      |
| No. of adult males in household      | 1.199<br>(0.016)         | 1.506<br>(0.023)         | -0.307***                                     |
| No. of children in household         | 1.059<br>(0.027)         | 1.549<br>(0.036)         | -0.490***                                     |
| Number of observations               | 1858                     | 1901                     | 3759  |
| Number of clusters                   | 100                      | 121                      | 221   |

Notes: Standard errors clustered at the village level in parentheses. \*( $p < 0.10$ ), \*\*( $p < 0.05$ ), \*\*\*( $p < 0.01$ )

Figure S4: Distribution of Deviations in Rainfall by Year



Notes: Solid lines represent smoothed distribution of rainfall deviations from average in 2018 program villages for each program year. Dashed vertical lines represent group mean in each year.

## C Evaluation Results and Robustness

In this section we present reduced form impact evaluation results of the effect of NLS on various outcomes following (1) according to the pre-analysis plan, explore robustness to the inclusion of controls, and discuss migration variable construction.

### C.1 Reduced Form Outcomes

We first present results for each primary evaluation outcome. Tables S10–S12 report treatment effects on migration, and Tables S13–S27 report treatment effects on the downstream outcomes of employment, labor earnings, net household income, food expenditures per capita, and a standardized index of food security, all measured at midline immediately after the lean season in both evaluation years. Tables S28–S39 report treatment effects on four additional downstream outcomes measured only in 2018: self-reported debt burden and calories per capita measured at midline immediately after the lean season, and employment and labor earnings measured at endline six months after the lean season. We present results separately for pilot district, expansion districts, and all evaluation districts combined. Each table reports results separately by year and checks robustness to dropping household controls and subdistrict fixed effects.

Table S10: Treatment Effects on Migration, Pilot Districts

| Outcome: Migration |                 |                 |                 |                |                |                |
|--------------------|-----------------|-----------------|-----------------|----------------|----------------|----------------|
|                    | 2017            |                 |                 | 2018           |                |                |
|                    | (1)             | (2)             | (3)             | (4)            | (5)            | (6)            |
| Treated            | 0.02<br>(0.06)  | 0.04<br>(0.05)  | 0.04<br>(0.05)  | 0.15<br>(0.04) | 0.14<br>(0.04) | 0.13<br>(0.04) |
| Spillover          | -0.02<br>(0.06) | -0.00<br>(0.06) | -0.01<br>(0.06) | 0.12<br>(0.06) | 0.13<br>(0.06) | 0.13<br>(0.06) |
| Branch-Control     | -0.04<br>(0.06) | -0.02<br>(0.06) | -0.02<br>(0.06) | 0.14<br>(0.04) | 0.11<br>(0.05) | 0.09<br>(0.05) |
| Control Mean       | 0.40            | 0.40            | 0.40            | 0.39           | 0.39           | 0.39           |
| HH Controls        | No              | No              | Yes             | No             | No             | Yes            |
| Subdistrict FEs    | No              | Yes             | Yes             | No             | Yes            | Yes            |
| N                  | 1,537           | 1,537           | 1,537           | 1,901          | 1,901          | 1,901          |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S11: Treatment Effects on Migration, Expansion Districts

|                 | Outcome: Migration |                 |                 |                 |                 |                 |
|-----------------|--------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | 2017               |                 |                 | 2018            |                 |                 |
|                 | (1)                | (2)             | (3)             | (4)             | (5)             | (6)             |
| Treated         | 0.03<br>(0.06)     | 0.05<br>(0.04)  | 0.06<br>(0.04)  | 0.07<br>(0.04)  | -0.04<br>(0.04) | -0.03<br>(0.03) |
| Spillover       | -0.08<br>(0.06)    | -0.05<br>(0.05) | -0.05<br>(0.05) | -0.03<br>(0.06) | -0.09<br>(0.04) | -0.09<br>(0.04) |
| Branch-Control  | -0.05<br>(0.06)    | -0.02<br>(0.05) | -0.02<br>(0.04) | 0.04<br>(0.05)  | -0.06<br>(0.04) | -0.07<br>(0.04) |
| Control Mean    | 0.33               | 0.33            | 0.33            | 0.38            | 0.38            | 0.38            |
| HH Controls     | No                 | No              | Yes             | No              | No              | Yes             |
| Subdistrict FEs | No                 | Yes             | Yes             | No              | Yes             | Yes             |
| N               | 2,141              | 2,141           | 2,141           | 2,423           | 2,423           | 2,423           |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S12: Treatment Effects on Migration, All Districts

|                 | Outcome: Migration |                 |                 |                |                |                |
|-----------------|--------------------|-----------------|-----------------|----------------|----------------|----------------|
|                 | 2017               |                 |                 | 2018           |                |                |
|                 | (1)                | (2)             | (3)             | (4)            | (5)            | (6)            |
| Treated         | 0.02<br>(0.04)     | 0.05<br>(0.03)  | 0.06<br>(0.03)  | 0.11<br>(0.03) | 0.06<br>(0.03) | 0.06<br>(0.03) |
| Spillover       | -0.06<br>(0.04)    | -0.03<br>(0.04) | -0.03<br>(0.04) | 0.03<br>(0.04) | 0.02<br>(0.04) | 0.02<br>(0.04) |
| Branch-Control  | -0.05<br>(0.04)    | -0.02<br>(0.04) | -0.02<br>(0.04) | 0.08<br>(0.04) | 0.03<br>(0.03) | 0.02<br>(0.03) |
| Control Mean    | 0.36               | 0.36            | 0.36            | 0.38           | 0.38           | 0.38           |
| HH Controls     | No                 | No              | Yes             | No             | No             | Yes            |
| Subdistrict FEs | No                 | Yes             | Yes             | No             | Yes            | Yes            |
| N               | 3,678              | 3,678           | 3,678           | 4,324          | 4,324          | 4,324          |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S13: Treatment Effects on Employment, Pilot Districts

|                 | Outcome: Employment |                 |                 |                 |                 |                 |
|-----------------|---------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | 2017                |                 |                 | 2018            |                 |                 |
|                 | (1)                 | (2)             | (3)             | (4)             | (5)             | (6)             |
| Treated         | -0.02<br>(0.02)     | -0.01<br>(0.02) | -0.01<br>(0.02) | -0.02<br>(0.02) | -0.01<br>(0.02) | -0.02<br>(0.02) |
| Spillover       | -0.01<br>(0.02)     | -0.00<br>(0.02) | -0.00<br>(0.02) | -0.02<br>(0.03) | -0.02<br>(0.02) | -0.03<br>(0.02) |
| Branch-Control  | -0.01<br>(0.02)     | -0.01<br>(0.02) | -0.00<br>(0.02) | -0.01<br>(0.03) | -0.01<br>(0.03) | -0.02<br>(0.03) |
| Control Mean    | 0.95                | 0.95            | 0.95            | 0.94            | 0.94            | 0.94            |
| HH Controls     | No                  | No              | Yes             | No              | No              | Yes             |
| Subdistrict FEs | No                  | Yes             | Yes             | No              | Yes             | Yes             |
| N               | 1,537               | 1,537           | 1,537           | 1,901           | 1,901           | 1,901           |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S14: Treatment Effects on Employment, Expansion Districts

|                 | Outcome: Employment |                 |                 |                 |                 |                 |
|-----------------|---------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | 2017                |                 |                 | 2018            |                 |                 |
|                 | (1)                 | (2)             | (3)             | (4)             | (5)             | (6)             |
| Treated         | -0.02<br>(0.02)     | -0.02<br>(0.02) | -0.01<br>(0.02) | -0.04<br>(0.01) | -0.03<br>(0.02) | -0.03<br>(0.02) |
| Spillover       | 0.01<br>(0.02)      | 0.01<br>(0.02)  | 0.01<br>(0.02)  | -0.02<br>(0.01) | -0.01<br>(0.02) | -0.01<br>(0.02) |
| Branch-Control  | 0.01<br>(0.02)      | 0.01<br>(0.02)  | 0.01<br>(0.02)  | -0.02<br>(0.02) | -0.00<br>(0.02) | -0.01<br>(0.02) |
| Control Mean    | 0.94                | 0.94            | 0.94            | 0.96            | 0.96            | 0.96            |
| HH Controls     | No                  | No              | Yes             | No              | No              | Yes             |
| Subdistrict FEs | No                  | Yes             | Yes             | No              | Yes             | Yes             |
| N               | 2,141               | 2,141           | 2,141           | 2,423           | 2,423           | 2,423           |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S15: Treatment Effects on Employment, All Districts

|                 | Outcome: Employment |                 |                 |                 |                 |                 |
|-----------------|---------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | 2017                |                 |                 | 2018            |                 |                 |
|                 | (1)                 | (2)             | (3)             | (4)             | (5)             | (6)             |
| Treated         | -0.02<br>(0.01)     | -0.02<br>(0.01) | -0.01<br>(0.01) | -0.03<br>(0.01) | -0.02<br>(0.01) | -0.02<br>(0.01) |
| Spillover       | 0.00<br>(0.01)      | 0.00<br>(0.01)  | 0.00<br>(0.01)  | -0.02<br>(0.01) | -0.01<br>(0.01) | -0.01<br>(0.01) |
| Branch-Control  | -0.00<br>(0.01)     | 0.00<br>(0.01)  | 0.00<br>(0.01)  | -0.02<br>(0.02) | -0.00<br>(0.02) | -0.01<br>(0.02) |
| Control Mean    | 0.94                | 0.94            | 0.94            | 0.95            | 0.95            | 0.95            |
| HH Controls     | No                  | No              | Yes             | No              | No              | Yes             |
| Subdistrict FEs | No                  | Yes             | Yes             | No              | Yes             | Yes             |
| N               | 3,678               | 3,678           | 3,678           | 4,324           | 4,324           | 4,324           |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S16: Treatment Effects on Labor Earnings, Pilot Districts

|                 | Outcome: Labor Earnings |                     |                     |                     |                     |                     |
|-----------------|-------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
|                 | 2017                    |                     |                     | 2018                |                     |                     |
|                 | (1)                     | (2)                 | (3)                 | (4)                 | (5)                 | (6)                 |
| Treated         | -226.59<br>(166.10)     | -203.39<br>(147.49) | -173.09<br>(139.24) | -92.28<br>(204.32)  | -185.99<br>(218.59) | -165.42<br>(204.72) |
| Spillover       | 106.59<br>(352.94)      | 135.55<br>(301.16)  | 294.56<br>(312.95)  | -210.67<br>(199.09) | -317.38<br>(237.80) | -230.83<br>(245.98) |
| Branch-Control  | -450.58<br>(128.84)     | -424.51<br>(146.21) | -221.57<br>(150.71) | -282.05<br>(205.10) | -284.54<br>(213.50) | -229.58<br>(226.43) |
| Control Mean    | 2385.83                 | 2385.83             | 2385.83             | 1773.59             | 1773.59             | 1773.59             |
| HH Controls     | No                      | No                  | Yes                 | No                  | No                  | Yes                 |
| Subdistrict FEs | No                      | Yes                 | Yes                 | No                  | Yes                 | Yes                 |
| N               | 1,537                   | 1,537               | 1,537               | 1,901               | 1,901               | 1,901               |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S17: Treatment Effects on Labor Earnings, Expansion Districts

|                 | Outcome: Labor Earnings |                    |                    |                     |                     |                     |
|-----------------|-------------------------|--------------------|--------------------|---------------------|---------------------|---------------------|
|                 | 2017                    |                    |                    | 2018                |                     |                     |
|                 | (1)                     | (2)                | (3)                | (4)                 | (5)                 | (6)                 |
| Treated         | -329.56<br>(234.93)     | -48.68<br>(133.15) | 155.54<br>(151.05) | -186.82<br>(183.10) | -126.66<br>(199.85) | -114.50<br>(184.05) |
| Spillover       | -53.61<br>(230.92)      | 185.06<br>(164.80) | 266.26<br>(165.32) | -68.40<br>(215.55)  | -41.18<br>(226.99)  | -42.85<br>(198.46)  |
| Branch-Control  | -184.41<br>(216.55)     | 55.42<br>(139.03)  | 6.28<br>(145.69)   | -116.54<br>(205.69) | 15.01<br>(212.30)   | -50.85<br>(189.96)  |
| Control Mean    | 2456.59                 | 2456.59            | 2456.59            | 2041.51             | 2041.51             | 2041.51             |
| HH Controls     | No                      | No                 | Yes                | No                  | No                  | Yes                 |
| Subdistrict FEs | No                      | Yes                | Yes                | No                  | Yes                 | Yes                 |
| N               | 2,141                   | 2,141              | 2,141              | 2,423               | 2,423               | 2,423               |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S18: Treatment Effects on Labor Earnings, All Districts

|                 | Outcome: Labor Earnings |                     |                     |                     |                     |                     |
|-----------------|-------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
|                 | 2017                    |                     |                     | 2018                |                     |                     |
|                 | (1)                     | (2)                 | (3)                 | (4)                 | (5)                 | (6)                 |
| Treated         | -284.89<br>(150.09)     | -146.55<br>(103.35) | -49.89<br>(105.01)  | -164.10<br>(138.81) | -160.87<br>(149.56) | -120.16<br>(140.82) |
| Spillover       | 13.15<br>(191.96)       | 127.00<br>(165.99)  | 202.64<br>(166.24)  | -120.79<br>(155.89) | -170.51<br>(162.98) | -123.31<br>(156.94) |
| Branch-Control  | -289.53<br>(136.08)     | -175.69<br>(108.46) | -112.32<br>(109.53) | -195.20<br>(151.72) | -131.39<br>(153.14) | -128.24<br>(148.10) |
| Control Mean    | 2425.00                 | 2425.00             | 2425.00             | 1933.28             | 1933.28             | 1933.28             |
| HH Controls     | No                      | No                  | Yes                 | No                  | No                  | Yes                 |
| Subdistrict FEs | No                      | Yes                 | Yes                 | No                  | Yes                 | Yes                 |
| N               | 3,678                   | 3,678               | 3,678               | 4,324               | 4,324               | 4,324               |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S19: Treatment Effects on Household Income, Pilot Districts

|                 | Outcome: Net Income   |                       |                       |                       |                       |                       |
|-----------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
|                 | 2017                  |                       |                       | 2018                  |                       |                       |
|                 | (1)                   | (2)                   | (3)                   | (4)                   | (5)                   | (6)                   |
| Treated         | -2038.76<br>(2754.10) | -864.49<br>(1717.47)  | -460.32<br>(1740.34)  | -1394.69<br>(2030.78) | -2316.67<br>(2208.00) | -2044.46<br>(2109.12) |
| Spillover       | -4030.17<br>(2522.40) | -2723.38<br>(2069.73) | -1719.50<br>(1924.19) | -4535.27<br>(2490.49) | -6368.85<br>(2511.78) | -5323.47<br>(2407.75) |
| Branch-Control  | 711.93<br>(2332.13)   | 1889.30<br>(1804.56)  | 3381.26<br>(1940.71)  | -1531.75<br>(2251.65) | -2600.65<br>(2113.39) | -2159.12<br>(2034.79) |
| Control Mean    | 39031.00              | 39031.00              | 39031.00              | 37733.29              | 37733.29              | 37733.29              |
| HH Controls     | No                    | No                    | Yes                   | No                    | No                    | Yes                   |
| Subdistrict FEs | No                    | Yes                   | Yes                   | No                    | Yes                   | Yes                   |
| N               | 1,456                 | 1,456                 | 1,456                 | 1,901                 | 1,901                 | 1,901                 |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S20: Treatment Effects on Household Income, Expansion Districts

|                 | Outcome: Net Income  |                       |                       |                      |                      |                      |
|-----------------|----------------------|-----------------------|-----------------------|----------------------|----------------------|----------------------|
|                 | 2017                 |                       |                       | 2018                 |                      |                      |
|                 | (1)                  | (2)                   | (3)                   | (4)                  | (5)                  | (6)                  |
| Treated         | 2568.76<br>(2087.63) | -1579.66<br>(2083.92) | -446.69<br>(2174.36)  | 642.39<br>(2440.84)  | 3753.12<br>(2867.47) | 4741.91<br>(2658.14) |
| Spillover       | 2983.97<br>(2340.55) | -1674.80<br>(2133.09) | -1572.17<br>(2204.56) | 1494.28<br>(3347.62) | 3746.62<br>(3527.71) | 4214.72<br>(3303.33) |
| Branch-Control  | 1930.79<br>(2053.23) | -2516.91<br>(2093.30) | -2577.12<br>(2078.69) | 2397.06<br>(3301.40) | 6234.94<br>(3375.49) | 5491.09<br>(3264.49) |
| Control Mean    | 40254.42             | 40254.42              | 40254.42              | 41586.61             | 41586.61             | 41586.61             |
| HH Controls     | No                   | No                    | Yes                   | No                   | No                   | Yes                  |
| Subdistrict FEs | No                   | Yes                   | Yes                   | No                   | Yes                  | Yes                  |
| N               | 2,021                | 2,021                 | 2,021                 | 2,423                | 2,423                | 2,423                |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S21: Treatment Effects on Household Income, All Districts

|                 | Outcome: Net Income  |                       |                       |                      |                       |                      |
|-----------------|----------------------|-----------------------|-----------------------|----------------------|-----------------------|----------------------|
|                 | 2017                 |                       |                       | 2018                 |                       |                      |
|                 | (1)                  | (2)                   | (3)                   | (4)                  | (5)                   | (6)                  |
| Treated         | 739.23<br>(1729.02)  | -1092.35<br>(1330.94) | -243.30<br>(1368.08)  | -706.14<br>(1669.22) | 408.92<br>(1819.22)   | 881.46<br>(1713.33)  |
| Spillover       | 244.30<br>(1852.04)  | -1876.46<br>(1447.26) | -1417.60<br>(1423.02) | -818.99<br>(2377.39) | -1229.85<br>(2300.81) | -403.10<br>(2136.92) |
| Branch-Control  | 1468.39<br>(1551.06) | -502.27<br>(1398.28)  | 23.43<br>(1450.29)    | 598.08<br>(2207.71)  | 1700.99<br>(2091.36)  | 1604.27<br>(2026.98) |
| Control Mean    | 39712.55             | 39712.55              | 39712.55              | 40029.99             | 40029.99              | 40029.99             |
| HH Controls     | No                   | No                    | Yes                   | No                   | No                    | Yes                  |
| Subdistrict FEs | No                   | Yes                   | Yes                   | No                   | Yes                   | Yes                  |
| N               | 3,477                | 3,477                 | 3,477                 | 4,324                | 4,324                 | 4,324                |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S22: Treatment Effects on Food Expenditure per Capita, Pilot Districts

|                 | Outcome: Food Expenditure |                     |                    |                     |                    |                    |
|-----------------|---------------------------|---------------------|--------------------|---------------------|--------------------|--------------------|
|                 | 2017                      |                     |                    | 2018                |                    |                    |
|                 | (1)                       | (2)                 | (3)                | (4)                 | (5)                | (6)                |
| Treated         | 72.20<br>(267.17)         | 33.31<br>(223.51)   | 83.55<br>(207.24)  | -405.11<br>(307.13) | -14.13<br>(283.27) | -62.32<br>(277.75) |
| Spillover       | 245.24<br>(242.96)        | 208.75<br>(209.98)  | 337.81<br>(199.53) | -610.93<br>(441.76) | 103.91<br>(399.62) | 179.82<br>(399.38) |
| Branch-Control  | -251.72<br>(226.62)       | -293.80<br>(238.72) | -53.70<br>(240.23) | -426.60<br>(425.93) | 198.70<br>(292.38) | 278.91<br>(309.26) |
| Control Mean    | 5041.56                   | 5041.56             | 5041.56            | 3155.60             | 3155.60            | 3155.60            |
| HH Controls     | No                        | No                  | Yes                | No                  | No                 | Yes                |
| Subdistrict FEs | No                        | Yes                 | Yes                | No                  | Yes                | Yes                |
| N               | 1,464                     | 1,464               | 1,464              | 1,463               | 1,463              | 1,463              |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S23: Treatment Effects on Food Expenditure per Capita, Expansion Districts

|                 | Outcome: Food Expenditure |                     |                     |                    |                    |                    |
|-----------------|---------------------------|---------------------|---------------------|--------------------|--------------------|--------------------|
|                 | 2017                      |                     |                     | 2018               |                    |                    |
|                 | (1)                       | (2)                 | (3)                 | (4)                | (5)                | (6)                |
| Treated         | -407.96<br>(184.67)       | -242.38<br>(176.97) | -120.01<br>(179.53) | 406.81<br>(222.31) | 268.23<br>(240.15) | 261.03<br>(241.40) |
| Spillover       | 9.79<br>(207.09)          | 146.70<br>(190.73)  | 101.35<br>(179.64)  | 112.81<br>(325.64) | 92.73<br>(243.42)  | 147.29<br>(252.26) |
| Branch-Control  | -117.09<br>(209.22)       | 12.95<br>(188.85)   | 19.23<br>(184.49)   | 262.69<br>(270.65) | 312.62<br>(227.68) | 425.74<br>(235.27) |
| Control Mean    | 5230.99                   | 5230.99             | 5230.99             | 2233.09            | 2233.09            | 2233.09            |
| HH Controls     | No                        | No                  | Yes                 | No                 | No                 | Yes                |
| Subdistrict FEs | No                        | Yes                 | Yes                 | No                 | Yes                | Yes                |
| N               | 2,072                     | 2,072               | 2,072               | 2,053              | 2,053              | 2,053              |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S24: Treatment Effects on Food Expenditure per Capita, All Districts

|                 | Outcome: Food Expenditure |                     |                    |                     |                    |                    |
|-----------------|---------------------------|---------------------|--------------------|---------------------|--------------------|--------------------|
|                 | 2017                      |                     |                    | 2018                |                    |                    |
|                 | (1)                       | (2)                 | (3)                | (4)                 | (5)                | (6)                |
| Treated         | -208.96<br>(153.89)       | -129.74<br>(140.67) | -2.35<br>(134.44)  | 109.82<br>(186.89)  | 124.29<br>(186.41) | 103.76<br>(184.80) |
| Spillover       | 111.47<br>(156.62)        | 172.19<br>(140.95)  | 222.47<br>(132.03) | -160.93<br>(265.66) | 77.16<br>(220.19)  | 125.92<br>(220.53) |
| Branch-Control  | -165.23<br>(155.85)       | -110.76<br>(151.65) | 15.46<br>(148.57)  | 7.03<br>(238.59)    | 252.63<br>(182.88) | 339.58<br>(187.74) |
| Control Mean    | 5147.39                   | 5147.39             | 5147.39            | 2582.21             | 2582.21            | 2582.21            |
| HH Controls     | No                        | No                  | Yes                | No                  | No                 | Yes                |
| Subdistrict FEs | No                        | Yes                 | Yes                | No                  | Yes                | Yes                |
| N               | 3,536                     | 3,536               | 3,536              | 3,516               | 3,516              | 3,516              |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S25: Treatment Effects on Food Insecurity (Standardized), Pilot Districts

| Outcome: Food Insecurity (Standardized) |                 |                 |                 |                |                |                |
|---|-----------------|-----------------|-----------------|----------------|----------------|----------------|
|   | 2017            |                 |                 | 2018           |                |                |
|   | (1)             | (2)             | (3)             | (4)            | (5)            | (6)            |
| Treated                                 | 0.15<br>(0.14)  | 0.17<br>(0.12)  | 0.18<br>(0.11)  | 0.14<br>(0.09) | 0.08<br>(0.06) | 0.05<br>(0.06) |
| Spillover                               | 0.26<br>(0.14)  | 0.27<br>(0.10)  | 0.20<br>(0.10)  | 0.32<br>(0.14) | 0.17<br>(0.10) | 0.14<br>(0.09) |
| Branch-Control                          | -0.03<br>(0.10) | -0.01<br>(0.10) | -0.06<br>(0.10) | 0.33<br>(0.13) | 0.21<br>(0.11) | 0.20<br>(0.10) |
| Control Mean                            | -0.07           | -0.07           | -0.07           | 0.01           | 0.01           | 0.01           |
| HH Controls                             | No              | No              | Yes             | No             | No             | Yes            |
| Subdistrict FEs                         | No              | Yes             | Yes             | No             | Yes            | Yes            |
| N                                       | 1,537           | 1,537           | 1,537           | 1,767          | 1,767          | 1,767          |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S26: Treatment Effects on Food Insecurity (Standardized), Expansion Districts

| Outcome: Food Insecurity (Standardized) |                 |                 |                 |                |                 |                 |
|---|-----------------|-----------------|-----------------|----------------|-----------------|-----------------|
|   | 2017            |                 |                 | 2018           |                 |                 |
|   | (1)             | (2)             | (3)             | (4)            | (5)             | (6)             |
| Treated                                 | -0.07<br>(0.11) | -0.02<br>(0.09) | -0.04<br>(0.08) | 0.10<br>(0.08) | -0.06<br>(0.08) | -0.05<br>(0.08) |
| Spillover                               | -0.15<br>(0.11) | -0.12<br>(0.10) | -0.13<br>(0.09) | 0.19<br>(0.11) | 0.10<br>(0.09)  | 0.08<br>(0.09)  |
| Branch-Control                          | 0.10<br>(0.11)  | 0.13<br>(0.09)  | 0.13<br>(0.09)  | 0.33<br>(0.12) | 0.19<br>(0.09)  | 0.20<br>(0.09)  |
| Control Mean                            | 0.06            | 0.06            | 0.06            | -0.01          | -0.01           | -0.01           |
| HH Controls                             | No              | No              | Yes             | No             | No              | Yes             |
| Subdistrict FEs                         | No              | Yes             | Yes             | No             | Yes             | Yes             |
| N                                       | 2,141           | 2,141           | 2,141           | 2,297          | 2,297           | 2,297           |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S27: Treatment Effects on Food Insecurity (Standardized), All Districts

| Outcome: Food Insecurity (Standardized) |                |                |                |                |                |                |
|---|----------------|----------------|----------------|----------------|----------------|----------------|
|   | 2017           |                |                | 2018           |                |                |
|   | (1)            | (2)            | (3)            | (4)            | (5)            | (6)            |
| Treated                                 | 0.02<br>(0.09) | 0.08<br>(0.07) | 0.07<br>(0.07) | 0.12<br>(0.06) | 0.02<br>(0.05) | 0.01<br>(0.05) |
| Spillover                               | 0.02<br>(0.09) | 0.06<br>(0.07) | 0.02<br>(0.07) | 0.24<br>(0.09) | 0.14<br>(0.06) | 0.12<br>(0.06) |
| Branch-Control                          | 0.06<br>(0.08) | 0.10<br>(0.07) | 0.07<br>(0.07) | 0.33<br>(0.09) | 0.21<br>(0.07) | 0.21<br>(0.07) |
| Control Mean                            | 0.00           | 0.00           | 0.00           | 0.00           | 0.00           | 0.00           |
| HH Controls                             | No             | No             | Yes            | No             | No             | Yes            |
| Subdistrict FEs                         | No             | Yes            | Yes            | No             | Yes            | Yes            |
| N                                       | 3,678          | 3,678          | 3,678          | 4,064          | 4,064          | 4,064          |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S28: Treatment Effects on Debt Burden, Pilot Districts

| Outcome: Debt Burden |                 |                 |                 |
|----------------------|-----------------|-----------------|-----------------|
|                      | 2017            |                 |                 |
|                      | (1)             | (2)             | (3)             |
| Treated              | 0.02<br>(0.08)  | 0.03<br>(0.07)  | 0.01<br>(0.07)  |
| Spillover            | -0.02<br>(0.12) | -0.06<br>(0.11) | -0.09<br>(0.11) |
| Branch-Control       | 0.11<br>(0.09)  | 0.12<br>(0.09)  | 0.07<br>(0.09)  |
| Control Mean         | 3.39            | 3.39            | 3.39            |
| HH Controls          | No              | No              | Yes             |
| Subdistrict FEs      | No              | Yes             | Yes             |
| N                    | 1,808           | 1,808           | 1,808           |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S29: Treatment Effects on Debt Burden, Expansion Districts

| Outcome: Debt Burden |                 |                 |                 |
|----------------------|-----------------|-----------------|-----------------|
| 2017                 |                 |                 |                 |
|                      | (1)             | (2)             | (3)             |
| Treated              | 0.02<br>(0.05)  | 0.05<br>(0.04)  | 0.04<br>(0.05)  |
| Spillover            | -0.10<br>(0.09) | -0.03<br>(0.06) | -0.04<br>(0.07) |
| Branch-Control       | 0.01<br>(0.06)  | 0.03<br>(0.05)  | 0.02<br>(0.06)  |
| Control Mean         | 3.51            | 3.51            | 3.51            |
| HH Controls          | No              | No              | Yes             |
| Subdistrict FEs      | No              | Yes             | Yes             |
| N                    | 2,289           | 2,289           | 2,289           |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S30: Treatment Effects on Debt Burden, All Districts

| Outcome: Debt Burden |                 |                 |                 |
|----------------------|-----------------|-----------------|-----------------|
| 2017                 |                 |                 |                 |
|                      | (1)             | (2)             | (3)             |
| Treated              | 0.01<br>(0.04)  | 0.04<br>(0.04)  | 0.03<br>(0.04)  |
| Spillover            | -0.06<br>(0.07) | -0.04<br>(0.06) | -0.06<br>(0.06) |
| Branch-Control       | 0.05<br>(0.05)  | 0.07<br>(0.05)  | 0.05<br>(0.05)  |
| Control Mean         | 3.46            | 3.46            | 3.46            |
| HH Controls          | No              | No              | Yes             |
| Subdistrict FEs      | No              | Yes             | Yes             |
| N                    | 4,097           | 4,097           | 4,097           |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S31: Treatment Effects on Calories per Capita, Pilot Districts

| Outcome: Calories per Capita |                     |                    |                    |
|------------------------------|---------------------|--------------------|--------------------|
| 2017                         |                     |                    |                    |
|                              | (1)                 | (2)                | (3)                |
| Treated                      | -52.92<br>(82.33)   | -54.83<br>(74.50)  | -29.29<br>(73.12)  |
| Spillover                    | -110.39<br>(105.20) | -74.12<br>(90.39)  | -108.09<br>(83.21) |
| Branch-Control               | -76.82<br>(90.64)   | -56.48<br>(103.77) | -46.96<br>(89.97)  |
| Control Mean                 | 2451.13             | 2451.13            | 2451.13            |
| HH Controls                  | No                  | No                 | Yes                |
| Subdistrict FEs              | No                  | Yes                | Yes                |
| N                            | 1,901               | 1,901              | 1,901              |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S32: Treatment Effects on Calories per Capita, Expansion Districts

| Outcome: Calories per Capita |                    |                    |                    |
|------------------------------|--------------------|--------------------|--------------------|
| 2017                         |                    |                    |                    |
|                              | (1)                | (2)                | (3)                |
| Treated                      | 93.51<br>(80.85)   | 112.70<br>(94.28)  | 98.31<br>(94.16)   |
| Spillover                    | -165.91<br>(77.49) | -163.43<br>(89.46) | -160.15<br>(88.94) |
| Branch-Control               | -8.96<br>(77.20)   | -7.96<br>(82.53)   | 4.44<br>(88.11)    |
| Control Mean                 | 2537.69            | 2537.69            | 2537.69            |
| HH Controls                  | No                 | No                 | Yes                |
| Subdistrict FEs              | No                 | Yes                | Yes                |
| N                            | 2,423              | 2,423              | 2,423              |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S33: Treatment Effects on Calories per Capita, All Districts

| Outcome: Calories per Capita<br>2017 |                    |                    |                    |
|--------------------------------------|--------------------|--------------------|--------------------|
|                                      | (1)                | (2)                | (3)                |
| Treated                              | 13.57<br>(57.76)   | 23.88<br>(59.48)   | 25.47<br>(59.70)   |
| Spillover                            | -143.10<br>(62.81) | -138.34<br>(63.83) | -157.44<br>(60.66) |
| Branch-Control                       | -40.58<br>(59.63)  | -35.42<br>(65.78)  | -19.41<br>(64.03)  |
| Control Mean                         | 2502.72            | 2502.72            | 2502.72            |
| HH Controls                          | No                 | No                 | Yes                |
| Subdistrict FEs                      | No                 | Yes                | Yes                |
| N                                    | 4,324              | 4,324              | 4,324              |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S34: Treatment Effects on Endline Employment, Pilot Districts

| Outcome: Employment (Endline)<br>2017 |                 |                 |                 |
|---------------------------------------|-----------------|-----------------|-----------------|
|                                       | (1)             | (2)             | (3)             |
| Treated                               | -0.00<br>(0.04) | -0.01<br>(0.04) | -0.01<br>(0.03) |
| Spillover                             | 0.01<br>(0.05)  | 0.01<br>(0.05)  | 0.02<br>(0.05)  |
| Branch-Control                        | -0.06<br>(0.05) | -0.05<br>(0.05) | -0.05<br>(0.05) |
| Control Mean                          | 0.64            | 0.64            | 0.64            |
| HH Controls                           | No              | No              | Yes             |
| Subdistrict FEs                       | No              | Yes             | Yes             |
| N                                     | 1,901           | 1,901           | 1,901           |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S35: Treatment Effects on Endline Employment, Expansion Districts

| Outcome: Employment (Endline) |                |                |                |
|-------------------------------|----------------|----------------|----------------|
| 2017                          |                |                |                |
|                               | (1)            | (2)            | (3)            |
| Treated                       | 0.06<br>(0.03) | 0.06<br>(0.03) | 0.07<br>(0.03) |
| Spillover                     | 0.01<br>(0.03) | 0.02<br>(0.03) | 0.02<br>(0.03) |
| Branch-Control                | 0.01<br>(0.04) | 0.04<br>(0.05) | 0.04<br>(0.05) |
| Control Mean                  | 0.80           | 0.80           | 0.80           |
| HH Controls                   | No             | No             | Yes            |
| Subdistrict FEs               | No             | Yes            | Yes            |
| N                             | 2,423          | 2,423          | 2,423          |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S36: Treatment Effects on Endline Employment, All Districts

| Outcome: Employment (Endline) |                 |                 |                 |
|-------------------------------|-----------------|-----------------|-----------------|
| 2017                          |                 |                 |                 |
|                               | (1)             | (2)             | (3)             |
| Treated                       | 0.01<br>(0.03)  | 0.02<br>(0.02)  | 0.03<br>(0.02)  |
| Spillover                     | 0.01<br>(0.03)  | 0.01<br>(0.03)  | 0.01<br>(0.03)  |
| Branch-Control                | -0.02<br>(0.04) | -0.00<br>(0.03) | -0.01<br>(0.03) |
| Control Mean                  | 0.74            | 0.74            | 0.74            |
| HH Controls                   | No              | No              | Yes             |
| Subdistrict FEs               | No              | Yes             | Yes             |
| N                             | 4,324           | 4,324           | 4,324           |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S37: Treatment Effects on Endline Labor Earnings, Pilot Districts

| Outcome: Labor Earnings (Endline) |                     |                     |                     |
|-----------------------------------|---------------------|---------------------|---------------------|
| 2017                              |                     |                     |                     |
|                                   | (1)                 | (2)                 | (3)                 |
| Treated                           | -26.25<br>(201.29)  | 1.76<br>(223.34)    | 27.09<br>(225.30)   |
| Spillover                         | -233.15<br>(201.01) | -224.51<br>(220.02) | -93.92<br>(230.03)  |
| Branch-Control                    | -394.57<br>(168.68) | -386.35<br>(190.58) | -334.72<br>(209.13) |
| Control Mean                      | 1327.70             | 1327.70             | 1327.70             |
| HH Controls                       | No                  | No                  | Yes                 |
| Subdistrict FEs                   | No                  | Yes                 | Yes                 |
| N                                 | 1,901               | 1,901               | 1,901               |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S38: Treatment Effects on Endline Labor Earnings, Expansion Districts

| Outcome: Labor Earnings (Endline) |                    |                    |                    |
|-----------------------------------|--------------------|--------------------|--------------------|
| 2017                              |                    |                    |                    |
|                                   | (1)                | (2)                | (3)                |
| Treated                           | -19.82<br>(108.45) | -69.43<br>(118.99) | -34.43<br>(117.28) |
| Spillover                         | -4.37<br>(124.71)  | -27.62<br>(132.60) | -5.12<br>(135.15)  |
| Branch-Control                    | 45.50<br>(135.85)  | 22.06<br>(146.42)  | 22.32<br>(143.91)  |
| Control Mean                      | 1381.17            | 1381.17            | 1381.17            |
| HH Controls                       | No                 | No                 | Yes                |
| Subdistrict FEs                   | No                 | Yes                | Yes                |
| N                                 | 2,423              | 2,423              | 2,423              |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

Table S39: Treatment Effects on Endline Labor Earnings, All Districts

| Outcome: Labor Earnings (Endline) |                     |                     |                     |
|-----------------------------------|---------------------|---------------------|---------------------|
| 2017                              |                     |                     |                     |
|                                   | (1)                 | (2)                 | (3)                 |
| Treated                           | -27.76<br>(110.68)  | -32.76<br>(132.90)  | -9.76<br>(133.81)   |
| Spillover                         | -93.29<br>(111.22)  | -115.63<br>(128.92) | -58.59<br>(131.91)  |
| Branch-Control                    | -145.71<br>(113.10) | -167.72<br>(126.60) | -157.07<br>(130.96) |
| Control Mean                      | 1359.57             | 1359.57             | 1359.57             |
| HH Controls                       | No                  | No                  | Yes                 |
| Subdistrict FEs                   | No                  | Yes                 | Yes                 |
| N                                 | 4,324               | 4,324               | 4,324               |

Notes: Regression results from at-scale evaluation following (1). Standard errors clustered at village level; 95% confidence intervals depicted on graph.

## C.2 Migration Outcome Measure

The measure of migration reported in the paper uses the self-reported departure date for each migration episode. To construct the variable, we pool all migration episodes reported at both midline and endline, and then count only those that take place in the same months used to measure migration in the 2008 and 2014 evaluations. Tables S40 and S41 report results using migration reported at midline and endline, respectively. The control mean differs from Table 8 because the midline survey covers a shorter migration window and the endline survey covers a longer one. Nevertheless, the pattern of estimated treatment effects remains consistent across migration measures.

Table S40: Estimated Treatment Effect on Migration, Midline

|                | 2017 Districts  |                 |                 | 2018 Districts |                |                 |
|----------------|-----------------|-----------------|-----------------|----------------|----------------|-----------------|
|                | All             | Pilot           | New             | All            | Pilot          | New             |
| Treated        | 0.04<br>(0.03)  | 0.04<br>(0.05)  | 0.03<br>(0.04)  | 0.08<br>(0.03) | 0.13<br>(0.04) | 0.01<br>(0.03)  |
| Spillover      | -0.05<br>(0.03) | -0.03<br>(0.05) | -0.08<br>(0.04) | 0.05<br>(0.04) | 0.14<br>(0.06) | -0.03<br>(0.04) |
| Branch Control | -0.03<br>(0.03) | -0.01<br>(0.06) | -0.05<br>(0.04) | 0.02<br>(0.04) | 0.07<br>(0.05) | -0.04<br>(0.05) |
| Control Mean   | 0.31            | 0.35            | 0.28            | 0.34           | 0.36           | 0.33            |
| HH Controls    | Yes             | Yes             | Yes             | Yes            | Yes            | Yes             |
| Upazila FEs    | Yes             | Yes             | Yes             | Yes            | Yes            | Yes             |
| N              | 4,428           | 1,839           | 2,589           | 4,324          | 1,901          | 2,423           |

Table S41: Estimated Treatment Effect on Migration, Endline

|                | 2017 Districts |                |                 | 2018 Districts |                |                 |
|----------------|----------------|----------------|-----------------|----------------|----------------|-----------------|
|                | All            | Pilot          | New             | All            | Pilot          | New             |
| Treated        | 0.04<br>(0.04) | 0.06<br>(0.05) | 0.01<br>(0.04)  | 0.05<br>(0.03) | 0.10<br>(0.04) | -0.02<br>(0.03) |
| Spillover      | 0.00<br>(0.04) | 0.07<br>(0.06) | -0.06<br>(0.05) | 0.04<br>(0.04) | 0.14<br>(0.06) | -0.06<br>(0.05) |
| Branch Control | 0.02<br>(0.04) | 0.03<br>(0.06) | 0.00<br>(0.05)  | 0.04<br>(0.04) | 0.13<br>(0.06) | -0.06<br>(0.05) |
| Control Mean   | 0.47           | 0.50           | 0.45            | 0.47           | 0.47           | 0.47            |
| HH Controls    | Yes            | Yes            | Yes             | Yes            | Yes            | Yes             |
| Upazila FEs    | Yes            | Yes            | Yes             | Yes            | Yes            | Yes             |
| N              | 3,678          | 1,537          | 2,141           | 4,324          | 1,901          | 2,423           |

## C.3 Migration History

In the two scale-up rounds, households were asked at baseline (prior to the migration subsidy offers) about their migration history—whether anyone in the household had migrated temporarily within the last 1–3 years. As shown in the balance tables above, the difference in migration history between incentivized and

control villages is statistically significant but switches signs between the two rounds. This may be caused by response bias.

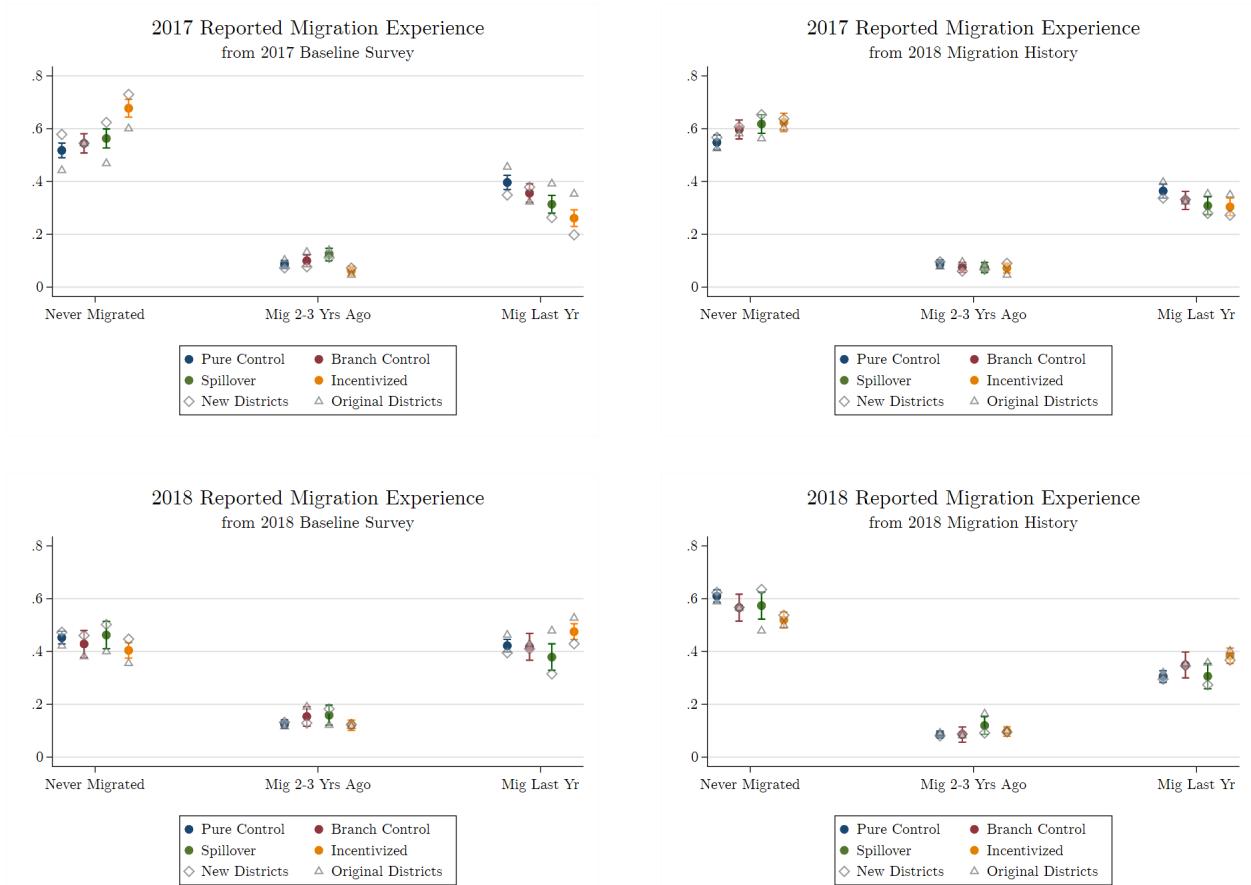
In 2017, baseline data was collected in incentivized villages by MFI migration officers as part of their eligibility screening. Loan officers were not told the eligibility criteria in the first year of at-scale implementation when they initially conducted eligibility surveys. Nevertheless, the implementing MFI knew that this data was being collected to screen for eligibility in a migration loan program. This foreknowledge may have influenced the way that questions were asked, the program was discussed, and what eligibility was assumed to depend on, in turn affecting recorded responses related to previous migration. In contrast, non-incentivized villages did not have migration officers present at baseline and the same questions were asked by our partner research organization instead. In 2017, reported recent migration was slightly lower among households in incentivized villages (interviewed by the migration officers) compared to the control groups (surveyed by research enumerators) (Figure S5, top left).

In 2018, the baseline survey was conducted by research enumerators in all villages. In that round, households in incentivized villages were more likely to report having migrated in the previous year, compared to control villages (Table S6 and Figure S5, top left). We do not have a clear explanation for this difference, but note that while branches were re-randomized into treatment and control between the two rounds, for treatment branches, the same villages were incentivized in 2017 and 2018. There may have been some learning from experience in this subset, as households inferred eligibility from what they observed in the previous round in their and nearby villages.

In 2018 endline data collection, we collected self-reported data on household migration in 2014–2018, and can compare responses given then with those given prior to program implementation in each round. These migration rates are less dispersed, with differences less significant between incentivized and control villages (Figure S5, right column). This is not to say that the endline information is free from response bias, however, as respondents may still have considered the possibility of future rounds of the program, particularly as the endline was carried out just 3 months prior to the subsequent lean season.

All results are robust to including or excluding baseline migration history as a control.

Figure S5: Self-Reported Migration History at Baseline and Endline



Notes: Left panels: migration history self-reported at baseline, collected by migration officers in 2017 and by research enumerators in 2018. Right panels: migration history self-reported in 2018 evaluation endline (conducted in June–August 2019), collected by research enumerators.

## D Type Distribution and Outreach Intensity Calculations

In this section we provide detail on the calculations used to compute the distribution of compliance types in 2008 and 2014, and the outreach intensity by type in 2018.

### D.1 Calculation of Type Distribution

To calculate the distribution of compliance types in the 2008 and 2014 implementation rounds, we combine administrative data on loan offers and disbursement with survey data on migration. In the 2008 evaluation, both household surveys and loan offer/disbursement were managed by the enumeration team, so both outcomes come from an integrated data set. In the 2014 evaluation, researchers initially identified the eligible population and assigned treatment status. They then shared the shared this information with the MFI for implementation and surveyed a subset of these households for evaluation. For our analysis, we link the administrative dataset on loans with the survey dataset on migration. The administrative dataset also records migration status for the subset of eligible households that accept a loan, which we use to validate the survey migration measure.

The frequency of Self-Sufficient, Time-Wasting, and Uninterested types can be measured directly in data on treated households based on their response to the loan offer and subsequent migration decision:

$$P_{SS} = \mathbb{P}\{\text{Decline, Migrate}|\text{Treated}\} \quad (5)$$

$$P_{TW} = \mathbb{P}\{\text{Accept, Not Migrate}|\text{Treated}\} \quad (6)$$

$$P_{UN} = \mathbb{P}\{\text{Decline, Not Migrate}|\text{Treated}\} \quad (7)$$

Those that accept a loan and then migrate in the treatment are a combination of Opportunistic and Induced types. To separate the relative frequency of each, we calculate the migration rate among untreated households, which comprises Self-Sufficient and Opportunistic types:

$$P_{SS} + P_{OP} = \mathbb{P}\{\text{Migrate}|\text{Untreated}\} \quad (8)$$

Equations (5) and (8) together pin down the value of  $P_{OP}$ . Finally, the frequency of Induced types can be computed as the residual

$$P_{IN} = 1 - (P_{SS} + P_{OP} + P_{TW} + P_{UN}) \quad (9)$$

Note that, just as with Compliers in program evaluation, we cannot specifically identify any individual as an Induced migrant; we can only statistically infer their frequency in the population.

The frequencies in the study populations used for these calculations are provided in Table S42.

### D.2 Calculation of Outreach Intensity

To calculate outreach intensity in 2018, we hold the type distribution  $\{P_*$  fixed and invert the system of equations described by (2) and (3) to solve for  $\{\omega_*\}$ . Outreach to Time-Wasting types is again uniquely

Table S42: Loan and Migration Frequencies in Pilot Rounds

|                                | 2008 | 2014 Low | 2014 High |
|--------------------------------|------|----------|-----------|
| Decline, Migrate — Treated     | 16.1 | 16.1     | 2.5       |
| Accept, Not Migrate — Treated  | 21.1 | 14.2     | 21.4      |
| Decline, Not Migrate — Treated | 20.4 | 23.9     | 1.8       |
| Migrate — Untreated            | 36.0 | 37.8     | 46.4      |

identified in the treatment group

$$\omega_{TW} = \frac{\mathbb{P}\{\text{Accept, Not Migrate|Treated}\}}{P_{TW}} \quad (10)$$

(11)

Opportunistic types in the treatment group would accept a loan if they received sufficient outreach. Therefore, those in the treatment group who migrate without a loan comprise Self-Sufficient types as well as Opportunistic types who received insufficient outreach, as described by the first line in (2). Inverting this yields

$$\omega_{OP} = \frac{P_{OP} + P_{SS} - \mathbb{P}\{\text{No Loan, Migrate|Treated}\}}{P_{OP}} \quad (12)$$

Finally, those who migrate with a loan consist of both Opportunistic and Induced types who received sufficient outreach. To isolate outreach to Induced types, can substitute the solution to (12) into the middle line of (2) to compute

$$\omega_{IN} = \frac{\mathbb{P}\{\text{Loan, Migrate|Treated}\} - \omega_{OP}P_{OP}}{P_{IN}} \quad (13)$$

These calculations assume Opportunistic, Induced, and Time-Wasting types would accept a loan if they received sufficient outreach to attend to the offer, so we can make inference from their measured behavior without relying on recall about loan offers.

Both Self-Sufficient and Uninterested types decline a loan when offered. Therefore, loan acceptance itself is insufficient to calculate outreach intensity to these types because we cannot tell whether a household without a loan intentionally declined an offer or simply did not receive enough outreach to attend to it. If we take the loan offer recall data seriously, we can calculate the implied outreach intensity to these types as

$$\omega_{SS} = \frac{\mathbb{P}\{\text{Recall, Decline, Migrate|Treated}\}}{P_{SS}} \quad (14)$$

$$\omega_{UN} = \frac{\mathbb{P}\{\text{Recall, Decline, Not Migrate|Treated}\}}{P_{UN}} \quad (15)$$

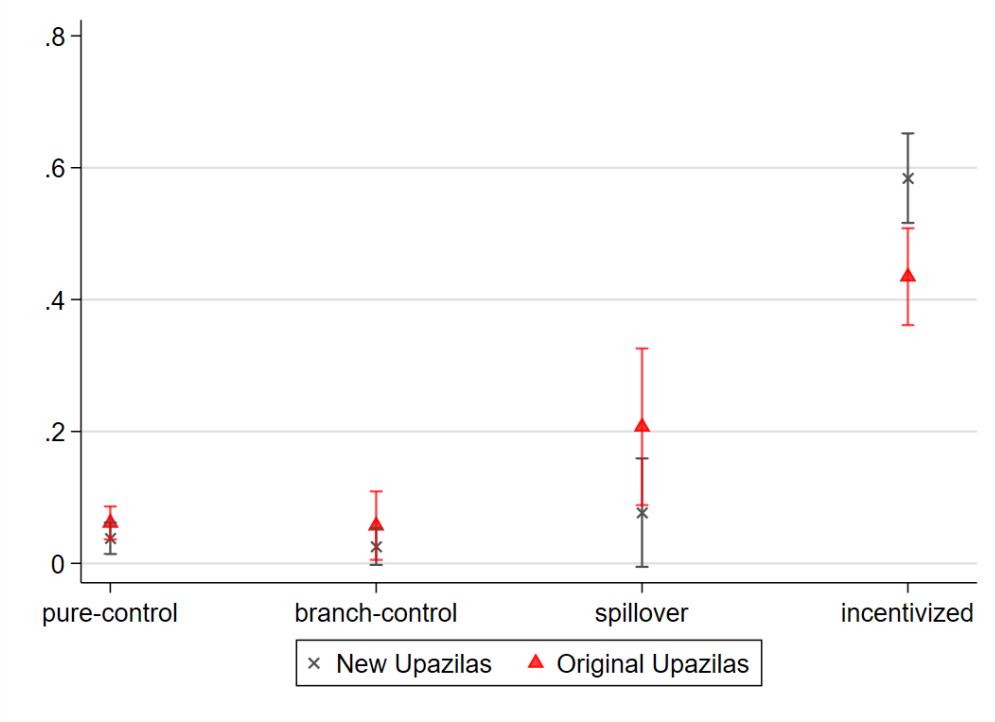
These values can be interpreted as lower bounds on outreach intensity because some fraction of those that do not recall receiving a loan offer may have chosen to decline the offer at the start of the lean season but did not remember the interaction by the time of endline surveying.

The frequencies in the study population used for these calculations are given in Table S43.

Table S43: Loan, Migration, and Offer Recall Frequencies in 2018

|  | Pilot | Expansion |
|--|-------|-----------|
| Loan, Migrate — Treated                | 24.9  | 27.6      |
| Loan, Not Migrate — Treated            | 11.8  | 16.2      |
| No Loan, Migrate — Treated             | 28.7  | 21.5      |
| Recall, Decline, Migrate — Treated     | 3.8   | 3.3       |
| Recall, Decline, Not Migrate — Treated | 3.1   | 4.0       |

Figure S6: Fraction of Households that Remember Loan Offer by Village



Fraction of households per village that remember receiving loan offer.

### D.3 Determinants of Loan Offer Recall

Figure S6 reports the fraction of households that remember being offered a migration loan during the prior lean season. While this fraction is substantially higher in treated villages than untreated, it is not close to 1. Only 40–60% of eligible households recall being offered a migration loan.

We use the survey data on household recall of loan offers to test for selectivity in recall. To do this, we first calculate the frequency of baseline characteristics among migrants in control villages and among non-migrants in treated villages in 2008. The difference between these two populations reflects the difference between regular migrants—both Opportunistic and Self-Sufficient types—and non-migrants—Time-Wasting and Uninterested types. For each characteristic, we compute the t-statistic for the difference as a measure of how informative the characteristic is about a household’s propensity to accept a loan. If migrant households are more attentive to migration opportunities, then these t-statistics should predict offer recall.

Table S44: Baseline Differences between Always- and Never-Takers

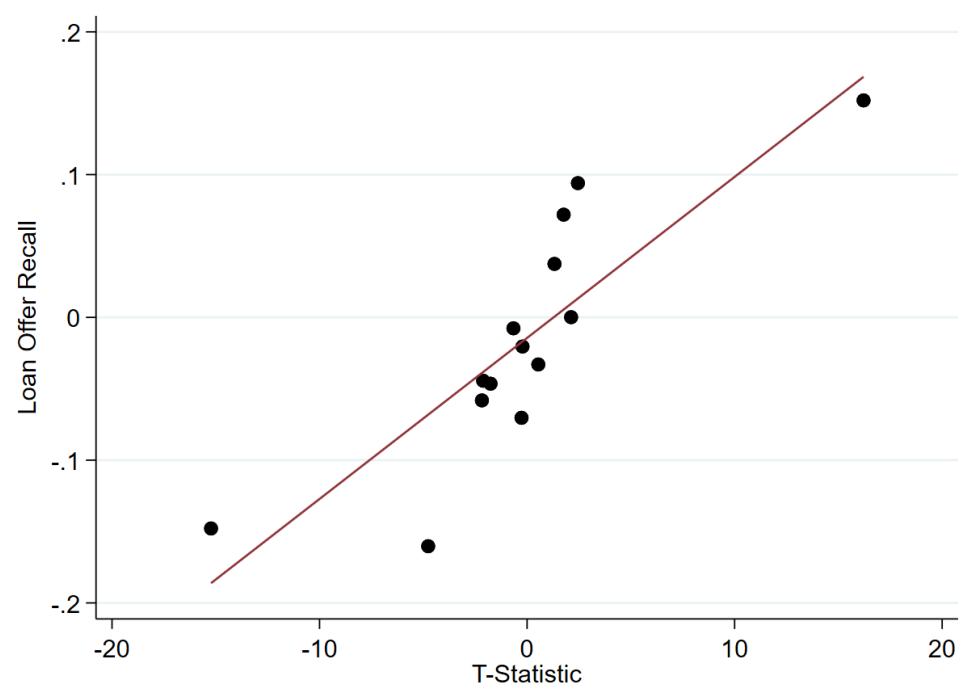
| Household Characteristics                   | Always Taker vs Never Taker       |                                   | Remember Offer from RDRS |
|---|-----------------------------------|-----------------------------------|--------------------------|
|   | Difference                        | T-statistic                       |                          |
| Never migrated or migrated over 3 years ago | -0.3723                           | -15.228                           | -0.148<br>(0.024)        |
| Migrated 2-3 years ago                      | -0.0385                           | -2.118                            | -0.044<br>(0.042)        |
| Migrated a year ago                         | 0.4108                            | 16.214                            | 0.152<br>(0.024)         |
| No land                                     | 0.0381                            | 1.762                             | 0.072<br>(0.028)         |
| Below med   on 0 < land $\leq$ 50           | -0.0045                           | -0.267                            | -0.070<br>(0.038)        |
| Above med   on 0 < land $\leq$ 50           | -0.0336                           | -2.177                            | -0.058<br>(0.040)        |
| Completed some education                    | -0.0485                           | -1.760                            | -0.046<br>(0.023)        |
| Zero adult males                            | -0.0533                           | -4.764                            | -0.160<br>(0.068)        |
| One adult male                              | -0.0181                           | -0.657                            | -0.008<br>(0.026)        |
| Two adult males                             | 0.0326                            | 1.320                             | 0.037<br>(0.028)         |
| Three or more adult males                   | 0.0388                            | 2.122                             | 0.000<br>(0.034)         |
| Borrowed money at baseline                  | 0.0539                            | 2.449                             | 0.094<br>(0.035)         |
| Baseline food insecurity                    | -0.0054                           | -0.221                            | -0.020<br>(0.029)        |
| Village flooding during the last year       | 0.0145                            | 0.541                             | -0.033<br>(0.052)        |
| Obs   | 1363                              | 1363                              | 1618                     |
| Sample                                      | Always Takers<br>and Never Takers | Always Takers<br>and Never Takers | Incentivized<br>Villages |

The first two columns of Table S44 report differences and t-statistics for the difference between always- and never-takers in each characteristic available at baseline. Migration history is the strongest predictor of current migration plans by far, with households that migrated recently more likely to be regular migrants and those that did not more likely to be never-migrants. The presence of adult males in the household also increases the likelihood that a household will send a migrant, as does borrowing at baseline. Land ownership and education of the household head are both negatively associated with regular migration.

Next, we regress household recall of receiving a loan offer on these baseline characteristics for households in treated villages in 2018. Regression results are reported in the final column of Table S44. We plot these regression coefficients against t-statistics in Figure S7, and the results are striking. There is a nearly monotonic relationship between how informative a characteristic is about migration status and how strongly that characteristic predicts whether a household remembers being offered a migration loan.

This relationship could come about either through differential attentiveness—as migrant households pay

Figure S7: Loan Offer Targeting to Always-Takers



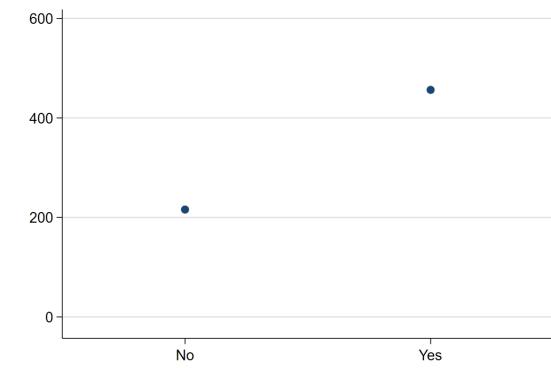
Each datapoint represents a household characteristic. x-axis: t-statistic for difference in that characteristic between always-takers and never-takers in 2008 pilot. y-axis: impact of that characteristic on likelihood of remembering a loan offer in 2018 endline.

more attention to migration opportunities—or through heterogeneity in loan officer outreach—if migration officers actively seek out those with existing migration plans to maximize their rate of loan disbursement. The analysis in Section 4 controls for these household characteristics to avoid contamination from selective attentiveness.

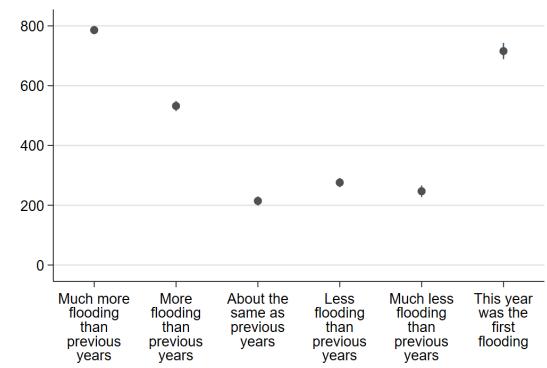
## E Measure of Rainfall

We use NOAA/NCEI CMORPH gridded satellite precipitation (daily,  $0.25^\circ$ ) to construct monthly rainfall totals between 2001 and 2019, which we then merge with household GPS coordinates to form household-by-month rainfall panels for analysis. We then aggregate the data to measure total rainfall in the 12 months prior to the typical start date of the NLS program. For each year we then construct a measure of deviations from the average annual rainfall over the 19 years in our panel. Figure S8 demonstrates that our measure of deviations in rainfall are highly correlated with self reported measures of flooding from 2014-2018 that we collect in the 2018 household endline survey. Our results in Figure S8 and Table 9 are similar if we instead use the absolute level of rainfall rather than deviations from the average.

Figure S8: Rainfall and Self Reported Flooding



(a) Reported Presence of Flooding



(b) Reported Flooding Degree

Each point corresponds to the mean relative rainfall for households and its associated 95% confidence interval. Panel (a) displays the relationship between relative rainfall and whether a household reported flooding in their village that year. Panel (b) plots the relationship between relative rainfall and the degree of reported flooding (conditional on reporting flooding). Reports of flooding correspond to 2014-2018.