Chapter 4

Experimental Design

Takashi Kurosaki

4.1 Introduction

This chapter describes the experimental design we employ in order to evaluate the impact of flexible microfinance interventions. As shown in Chapter 1 and 2, microcredit's success in poverty reduction has an important limitation that it is typically not offered to the poorest of the poor (those who are often called the "ultrapoor"). Lending to the ultrapoor may be too risky for microfinance institutions even with group lending designs while the ultrapoor may refrain from borrowing since they are afraid of accumulating debt and are not certain whether they can repay regularly as scheduled in a typical Grameen-style microcredit scheme.

In an attempt to enhance the welfare level of the ultrapoor, this study aims at understanding what will happen if a typical Grameen-style microcredit scheme is "relaxed" in terms of repayment scheduling. As shown in Chapter 3 as well as existing studies such as Shonchoy (2011) and Khandker (2012), seasonal deprivation associated with *Monga* and natural-disaster-related deprivation are serious threats that affect the poor in Bangladesh generally, and the ultrapoor in northern Bangladesh particularly. *Monga* refers to the period from September to November after the planting and before the harvesting of *Aman* paddy (main rice variety of Bangladesh), when farmers face seasonal unemployment. *Monga* sufferings could intensify when monsoon floods are larger than usual. In northern Bangladesh, to cope with such seasonal and disaster-driven deprivation, temporary migration plays an important role (Shonchoy, 2011; Berg and Emran, 2011). However, if a deprived person is a microcredit borrower under the regular design and subject to inflexible repayment schedule including weekly meetings, such a person cannot use the opportunity of temporary migration. Or expecting this constraint, a person

who expects suffering from such deprivation in the near future may intentionally refrain from borrowing a microcredit in order to be ready for temporary migration when the need arises. It is likely, therefore, that flexible microcredit can overcome these problems, allowing the ultrapoor to fully utilize opportunities from both temporary migration and microcredit.

Despite this potential, rigorous impact evaluation of such flexibility in microcredit designs is lacking in the literature. Among the few existing studies, Shoji (2010) evaluates the effectiveness of Bangladeshi microfinance in introducing a contingent repayment system beginning in 2002, which allowed rescheduling of savings and installments during natural disasters for affected members. In using evidence from a flood in 2004 based on an instrumental variable approach, he found that rescheduling plays the role of a safety net by substantially decreasing the probability that people skip meals during negative shocks and the effect is even higher on the landless and females. In a context more similar to ours, Czura *et al.* (2011) evaluates the impact of flexible repayment scheduling on borrowers' investment, consumption level, consumption smoothing, and repayment in the case of microcredit for dairy activities (i.e., microcredit to purchase milking cows) in India. Notably, they adopt a randomized controlled trial (RCT) design in their evaluation and find that flexible microcredit helps households to smooth consumption while increases default possibility. Except for this study, we are aware of no rigorous study on the impact of repayment flexibility in South Asia based on a RCT design. Furthermore, if we restrict our attention to studies in the context of *Monga*-related seasonal deprivation in northern Bangladesh, we have found neither descriptive nor qualitative study on flexible microcredit.

We have thus initiated RCT experiments in northern Bangladesh since early 2011. In this chapter, we explain how we designed the experiments and how we will investigate the impact of flexibility using the RCT design. RCT experiments are implemented in partnership with an NGO named Gana Unnayan Kendra (GUK). GUK is a local NGO based in Gaibandha District, northern Bangladesh, where the population of the ultrapoor is concentrated due to severe *Monga* and frequesnt floods from Jamuna River.

In the followings, we first describe our RCT design in Section 2. Section 3 explains the field implementation of our surveys and RCT interventions. More detailed description of GUK and its past activities is given in this section as well. By the time of this writing, two surveys have been completed: benchmark and short *Monga* survey for 1,440 households who had been given one of the six statuses (one control and five

treatment statuses). The benchmark survey was implemented before interventions while the short *Monga* survey was carried out during the interventions in late 2011. Using the benchmark survey data, Section 4 examines how successful our RCT design was in guaranteeing exogenous variation in microcredit repayment designs. Section 5 summarizes the chapter.

4.2 RCT Strategy

4.2.1 Inflexible microcredit as the control

A typical Grameen-style microcredit scheme proceeds as the followings (Armendariz and Morduch, 2010). Persons eligible for microcredit first form a group who are expected to help each other in times of difficulty (there is an argument whether this is only in terms of moral support or this implies the joint liability in the legally-enforceable sense, but this argument is irrelevant in our context). Not all members can borrow a microcredit immediately. It is usually the case that only some of them are given credit after all members save small amount of money on regular basis and the rest of them are given credit after the first borrowers successfully repay several installments and all members continue to save the small amount regularly. The weekly repayment begins without a long grace period. In a typical Grameen microcredit, the first lending is a small amount and it is to be repaid in 50 weekly installments within a year.

Several reasons are offered as rationales for this rigid design of the repayment schedule (Armendariz and Morduch, 2010). The success of frequent repayment in minimizing default and delay could be attributable to the early warning mechanism, the lender's capture of non-microfinance income flow of the borrower, and the borrower's commitment to save regularly. Repayment in group meetings in front of others also forces regular repayment of those borrowers who would like to maintain their reputations in the village.

Probably because of these mechanisms, the classic Grameen-type microcredit has been successful in maintaining high repayment rates.[1] However, attending weekly meetings regularly puts a high burden on the borrowers in terms of opportunity costs of time. Relaxing several of the classic Grameen-type features is thus being demanded from

[1] See Kurosaki and Khan (2012) for an exceptional case where a MFI suffered from high default rates despite adopting a Grameen-type credit scheme. In their case, due to weak enforcement of the contingent renewal rule, strategic default prevailed among borrowers.

borrowers. Academic research has been responded to this request as well, which have attempted to identify the key element that was the most critically important in guaranteeing the high repayment rates. For example, using the field experiment approach, Gine and Karlan (2011) evaluate the impact of removing group liability in the Philippines and find no adverse impact on repayment as long as public and frequent repayment systems are maintained. On the other hand, recent studies on the comparison between weekly versus monthly installments based on a RCT design show a mixed result: In India, Field and Pande (2008) show no differences between microfinance schemes with weekly and monthly repayment frequencies, as long as the repayment was done in a public meeting style, while in Indonesia, Feigenberg et al. (2011) find that repayment performance was better under weekly repayment installments than under monthly installments.

Given this background, we adopt the following repayment pattern as the control. Borrowers obtain credit of 3,000 Takas[2] and begin repayment after a short grace period of 2 weeks. The repayments are in 50 installments, each of which is 75 Takas, implying the gross interest payment of 750 Takas, which is spread through the borrowing period of approximately one year. Each of the weekly installments is to be repaid in a weekly meeting by the borrower (the borrower is obliged to attend the weekly meeting, including the period of *Monga*). We call this design as a traditional or inflexible microcredit scheme and denote it as "Control" or TT.

4.2.2 Flexible microcredit as the treatment

During the agriculturally lean period of *Monga*, microcredit borrowers may face difficulty in arranging money required for the repayment. They may want to migrate temporarily to cities to obtain additional wage income. To facilitate the demand for repayment flexibility in this context, the treatment relaxes the repayment schedule in two ways during *Monga*. The specific period designated as *Monga* in the treatment is the period from September 20 to December 20.

Under the first treatment, "Flexible 1," the repayment is temporarily put on moratorium during the designated *Monga* period. During the moratorium period, Flexible 1 groups do not pay any installment. After the *Monga* period is over, borrowers begin paying 100 Takas per week, so that their total repayment amount and repayment period be the same as those of the control group.

[2] One hundred Bangladeshi Takas were equivalent to approximately 99 Japanese Yen or US$1.22. 3,000 Takas were therefore approximately US$37.

Under the second treatment, the repayment is changed into two monthly installments of 300 Takas during the designated *Monga* period. After the *Monga* period is over, borrowers resume paying 75 Takas per week, so that their total repayment amount and repayment period be the same as those of the control group.

In experimenting with the two treatments, we add another dimension of experiments. We adopt two different ways of informing borrower groups at the time of group formation, in order to create exogenous variation in the information structure, as implemented by Karlan and Zinman (2009) in the context of consumer credit in South Africa. In the first set of experiments, borrower groups are formed given advertising for a traditional microcredit. Then when the *Monga* period begins, one third of them are surprised with Flexible 1 and another third are surprised with Flexible 2 (the remaining one third belong to the Control category). Each of these three treatment arms are abbreviated as TF1, TF2, and TT.

In the second set of experiments, borrower groups are formed given advertising for a flexible microcredit from the beginning. This group is also divided into three groups with different types of treatment. One third of them are given Flexible 1 treatment, another third are given Treatment 2, and the remaining one third are given the combination of Flexible 1 and the income generation activities (IGA) support. Under the IGA support, instead of giving cash, we provide microcredit borrowers with a productive asset as per their choice within the credit amount and advice for utilizing the asset, but we do not give any further subsidy. Each of these three treatment arms are abbreviated as F1, F2, and F3. If asymmetric information in borrower selection is a serious issue in the current context, borrowers under F1, F2, or F3 should be systematically different from borrowers under TF1, TF2, and TT. How large is the difference is an empirical question.

4.2.3 Randomization of treatment arms

As shown in the previous subsection, we have six treatment arms, including the one corresponding to the Control. In order to avoid unequal treatment of members within a group, we randomized the treatment status at the borrower group level. Since the counterpart NGO usually forms one group in one village, this implies that our randomization takes place at the village level.

Out of the list of 90 villages that were under potential treatment by the counterpart NGO, we randomly selected 36 villages for organizing under the traditional microcredit advertising and other 36 villages for organizing under the flexible microcredit advertising. Each group of 36 villages was further divided into three

sub-groups of 12 villages, each for different treatment arms. The distribution of six types of treatment arms is shown in Figure 4.1. In the randomization, we stratified villages based on the distance from the nearby station and location type of the village (see below).

In each village, a borrower group, known as *Samity*, was formed, comprising 20 members. Out of them, 15 members were those who satisfied the NGO's microcredit criterion and those who would be interested to receive microcredit. These members were given microcredit of 3,000 Takas in early September, 2011. The remaining five members did not receive microcredit in 2011 and remained in the group as a kind of observers.

This randomization implies the following sample distribution. There are 72 sample villages and 1,440 sample households, one sixth of which falls into one of the six treatment arm categories including the Control. Three fourth of the sample households (1,080 households) were actual borrowers of microcredit. There are 15 borrower households for each of the six treatment arm categories (Figure 4.1).

4.2.4 Empirical strategies to identify the impact of treatment

Let Y_{hj} be the level of some welfare indicator for household h who was treated with treatment j. We need to have an estimate for $E[Y_{hj}]$-$E[Y_{hk}]$ to evaluate the impact of treatment j relative to treatment k (treatment k may be the Control or a different type of treatment). By definition, however, for each h, we observe only one treatment status. Therefore, this estimate cannot be readily available.

Nevertheless, if our randomization is implemented appropriately, the heterogeneity across households becomes orthogonal to the treatment status. Under that condition, the simple test of difference in means between treatment arm type j and k is sufficient to show the impact of treatment j relative to treatment k. In other words, we can use $Avg[Y_j]$-$Avg[Y_k]$ as an unbiased estimate for $E[Y_{hj}]$-$E[Y_{hk}]$. In the fourth section of this chapter, we provide preliminary results regarding whether our randomization was ideally implemented.

Figure 4.1 shows how we interpret the differences across different treatment arms. The total effect of flexible microcredit over the traditional (inflexible) microcredit is shown by the difference between F (F1, F2, or F3) and TT. The total effect can be decomposed into selection and behavioral effects. The selection effect corresponds to the fact that (observationally or unobservationally) different households may be attracted to form a *Samity*, depending on how the advertising for microcredit is done (that is, whether the flexibility was given to borrowers as surprise or not). The behavioral effect corresponds to changes in borrower's behavior due to flexibility, such as combining

temporary migration and microenterprises at the same time. The selection effect should be shown by the difference between F1 and TF1 (or F2 and TF2). The behavioral effect should be shown by the difference between TF1 and TT (or TF2 and TT). If the selection effect is statistically insignificant, the total effect can be interpreted as the causal effect of flexibility on borrowers. Since our sample size is not large,[3] combining surprise and non-surprise treatment groups will give us advantage in statistical power. For this reason as well, examining the existence of the selection effect is important.

4.3 Implementation of Surveys and RCT Interventions

4.3.1 Counterpart NGO and Study Area

GUK operates in the greater Gaibandha areas. The grearter Gaibandha areas comprise five districts in northern Bangladesh: Gaibandha, Kurigram, Rangpur, Lalmonirhat, and Nilphamari. It has offices in all 32 Upazillas (sub-districts) in Gaibandha District and 5 offices in Kurigram District.

The NGO has a focus on the ultrapoor in terms of economics well-beings, females in terms of gender, and *Char* areas geographically. *Char* literally means a river island, which is an area of land that regularly forms sediments of river bed that is eroded by major rivers of Bangladesh. *Char* islands are generally sandy and infertile in northern Bangladesh. Seasonal as well as occasional major floods cause regular sufferings for the people living there. These islands are just a few inches above normal river water level and are extremely vulnerable to flooding during the wet season as monsoon precipitation usually swells the river together with glacier melt of the Himalayas. Districts of Gaibandha and Kurigram contain many of Char areas of Bangladesh, due to Jamuna River. In other words, GUK targets toward the most vulnerable section in the greater Gaibandha areas.

GUK has limited experience in running traditional microfinance. It has been a promoter of flexible microfinance combined with their reportedly successful “asset transfer” program, which was financed by international donors. However, since their asset transfer program contains a high subsidy component, it is not clear how much of its success (the reach out to the ultrapoor) could be attributable to the flexibility in their

[3] Remember that our effective sample size to identify the treatment impact is the number of villages rather than the number of households, because our unit of intervention is village, not household.

repayment design *per se*. For instance, under one of GUK's program, the ultrapoor beneficiaries were provided with a livestock animal and were required to return the livestock with its offspring or equivalent monetary value only. This design also implies a much longer grace period than traditional microcredit.

GUK has also implemented several programs targeted at improving *Char* livelihood. Although not included in *Char* areas in a strict sense, mainland areas on Jamuna River and its distributaries are also vulnerable to similar flood- and erosion-related disasters. We call such areas "river-basin" and distinguish them from *Char* areas as well as mainland areas reasonably away from rivers (called "inland"). In terms of exposure to floods and erosions, river-basin people are as vulnerable as *Char* people. In terms of coping infrastructure with respect to flood and erosion risk, however, river-basin people are less vulnerable than *Char* people because river-basin areas are directly connected with inland areas by road, while *Char* areas are isolated from inland areas by the river.

For these reasons, in the randomization, we stratified villages based on the distance from the nearby station and village location types of *Char*, river-basin, and inland. The distribution of our final sample villages is shown in Table 4.1. Forty-five out of 72 sample villages (or 62.5% of the sample) are from Gaibandha District while the rest (37.5%) are from Kurigram District. Eighteen out of 72 sample villages (or 25.0% of the sample) are from *Char* areas, 42 villages (58.3%) are from inland areas, and the rest of 12 villages (16.7%) are from river-basin areas.

4.3.2 Survey and experiment schedule in the field

In the first half of 2011, we visited Gaibandha and GUK for preparatory investigations and logistical arrangements. After the agreement with GUK on the research design, village-level randomization was implemented, followed by the formation of *Samity*. The benchmark survey of 1,440 households was conducted in July-August 2011. It covered detailed information on household roster, education, health including weight of children, occupation, asset, income, migration experiences, agricultural production, non-farm enterprise, saving, credit, debt, *Monga* coping, and so on.

In the first three weeks of September 2011, microcredit of 3,000 Takas was issued to three fourths of our sample households. Our initial plan was to issue the microcredit earlier. However, due to the holy month of Ramazan and the following festival of Eid-ul-Fitr, the issue was delayed than initially planned. As a result, those households who were given flexible microcredit entered the designated *Monga* period

before their first installment's due date. Nevertheless, GUK was able to collect monthly installments (Flexible 2) and enlarged weekly installments in the post-*Monga* period (Flexible 1) without serious delay or non-repayment problem. Another small deviation from our initial design was that 15 out of 20 *Samity* members were issued with credit in all villages except for two. In one village, which was assigned the treatment arm of TF2 (surprise flexible 2), 16 members received credit, while in another village, which was assigned the Control status, 14 members received credit. The initial design of giving credit to three fourth of the sample households was thus achieved.

In late October 2011, during the peak period of *Monga*, a short survey was conducted on the same 1,440 households. The focus of this *Monga* survey was to collect information on how the household was coping with the on-going *Monga* difficulty. Consumption regularity, migration decisions, and the use of credit were major questions in the survey. We also implemented a short test on cognitive ability of the *Samity* member of each sample household.

4.4 Validity of Randomization

4.4.1 Village-level variable

If our randomization was implemented properly, we should observe no systematic difference in village-level characteristics across different treatment arms. Unfortunately, village-level variables have not yet been compiled. Table 4.2 shows the comparison of means of the distance from the nearby station to the village. The average distance is lower in TT villages than those in other treatment villages but the difference was statistically insignificant.[4] While this result is comfortable, it is as expected since we used the distance as one of stratification variables (see Subsection 4.2.3).

4.4.2 Household-level variables

If our randomization was implemented properly and the selection effect (due to the difference in announcement during the *Samity* formation period with respect to flexible microcredit experiment) was absent, we should observe no systematic difference in household-level characteristics across the six treatment arms. On the other hand, if our

[4] In the table, we report bivariate comparison using the Control (TT) households as the reference group. None of the other 10 combinations choosing two treatment arms out of six had statistically significant difference at the 10% level, either (not reported to save space).

randomization was implemented properly but the selection effect was not ignorable, we should observe no systematic difference in household-level characteristics across TT, TF1, and TF2 on the one hand and across F1, F2, and F3 on the other hand. The difference between the group of households belonging to TT, TF1, and TF2 and the group of households belonging to F1, F2, and F3 indicates the difference in observable characteristics due to the selection.

From the benchmark survey data, 13 variables have been compiled for the analysis in this chapter. Because data cleaning has not been completed, the results reported in this chapter are very preliminary, subject to further revisions once data cleaning will be completed. Furthermore, the 13 variables cover only demographic and education characteristics. Rich information on household asset and credit positions has been collected in the benchmark survey. The distribution of these variables should also be examined.

Definition and summary statistics of these variables are shown in Table 4.3. Out of the 13 variables, the first five correspond to the household head's characteristics, while the rest eight aggregates household members' demographic and education characteristics. As shown in the table, on average, the sample households and their heads are younger than the average Bangladeshi households: average household heads' age is 39 years old and household members' age is 25 years old. Reflecting this younger age probably, the average household size is not large either: the average is 3.99 persons. The ratio of female-headed households is as high as 19.90%, substantially higher than the Bangladesh's average, indicating the nature of GUK targeting. Regarding the literacy rate, there is a huge gap between the stricter definition and the less strict definition of literacy. Based on the former, 74% of the sample household heads are literate while their literacy rate was only 21% based on the latter.

Table 4.4 and Table 4.5 compare the mean of these variables across different treatment arms. To save space, we only report the mean values and *p*-values for two types of tests. In the first type of tests, pair-wise mean comparison tests using TT households as reference are conducted. For each variable, 5 tests are conducted and reported in a row just below the row reporting the mean values. If our randomization went fine, we expect insignificant results for TF1 and TF2 while we may find significant results for F1, F2, and F3. In the second type of tests, ANOVA tests for the null hypothesis of independence between treatment and the mean of the variable are conducted. For each variable, three ANOVA tests are conducted, where treatment types are TF1, TF2, and TT for the first entry, all six for the second entry, and F1, F2, and F3 for the last entry. If our

randomization went fine, we expect insignificant results for the first entry (among TF1, TF2, and TT) and last entry (among F1, F2, and F3). The second entry shows insignificant results only when our randomization went fine and the selection effects were ignorable.

In general, means reported in Tables 4.4-4.5 are fairly similar, suggesting the success of randomization. However, looking at the results of statistical tests of the significance of the mean difference, we come across a few cases where the difference was statistically significant at the 5% or lower. For each variable, out of eight tests reported in the tables, four should show insignificance under well-implemented randomization. There are thirteen variables in Tables 4.4 and 4.5. This means that there are 52 entries where we should expect insignificance (or at most 2 or 3 cases with a difference statistically significant at the 5% or lower. Out of these 52, 18 entries, or approximately 35% show significant results. The fact that such cases occur more frequently than 5% of tests suggests a possibility that randomization was not achieved. Out of the remaining 52 test statistic entries (i.e., those for which insignificant results only when there exists no selection effects), 16 entries, or approximately 31% show significant results. Since the occurrence rate is similar to the first 52 entries, the difference appears more like a failure of randomization rather than a reflection of the selection effects. However, this is only a conjecture, for which more detailed investigation is called for.

Looking at individual variables, the female-headed household status (*fem_dum*), non-Muslim dummy (*nonmuslim*), and three literacy variables based on a less restrict definition (*educ_lit1*, *lit_rate1m*, and *lit_rate1f*) show significant differences across treatment arms. Female headed households' shares are higher among TT, F1, and F2 households than among TF1, TF2, and F3 households. Non-Muslim households concentrate into F3 households. TT households' literacy rates (less strict definition) are lower than those of TF1 (in the case of household heads and household male adults) or TF2 (in the case of household female adults). However, the absolute size of the difference is not very large in all of these cases.

4.5 Summary and Conclusion

This chapter explained the experimental design of our randomized control trials (RCTs) in northern Bangladesh to examine the impact of flexible microcredit targeted towards the ultrapoor. In partnership with GUK, a local NGO with rich experience in poverty

reduction in northern Bangladesh, we began RCT interventions to test the effectiveness of introducing moratorium or monthly repayment during a lean season called *Monga*, instead of regular and inflexible weekly repayment throughout the year.

After describing our experimental design, this chapter also compared means of sample households' characteristics across different treatment arms. It was found that our direction with respect to RCT designs was implemented fairly well in the field, although we came across several instances of statistically significant difference in observable characteristics of sample households across different treatment arms. Unfortunately, the occurrence rate of such significant differences was much higher than explained by a pure chance, suggesting a possibility that randomization was not achieved in the strict sense.

It should be emphasized that the empirical investigation of the exogeneity of randomization in this chapter is highly preliminary. The benchmark data need to be cleaned completely. Similar comparison of means across treatment arms should be conducted for variables characterizing before-intervention household asset and credit positions. If these tests show that there still remains some doubt about the exogeneity, we should examine carefully in our final impact evaluation whether the failure to randomize in an ideal way was an indication of non-randomized elements that were not orthogonal to welfare indicators to be used in the evaluation. This is because even if treatments were assigned in a non-random way, the non-randomness would not give us any harm if the non-randomized elements were nearly orthogonal to such welfare indicators. For such examination, we need to understand more deeply the field-level implementation of how GUK mobilizes the ultrapoor households to form a *Samity*.

References:

Armendariz, Beatriz and Jonathan Morduch (2010) *The Economics of Microfinance*. Second edition. Cambridge, Mass.: MIT Press.

Berg, Claudia and M. Shahe Emran (2011) "Does Microfinance Help the Ultrapoor Cope with Seasonal Shocks? Evidence from Seasonal Famine (Monga) in Bangladesh," Mimeo, George Washington University.

Czura, Kristina, Dean Karlan, and Sendhi Mullainathan (2011) "Does Flexibility in Microfinance Pay Off? Evidence from a Randomized Evaluation in Rural India," Paper presented at NEUDC 2011 Conference.

Field, Erica and Rohini Pande (2008) "Repayment Frequency and Default in Micro-Finance: Evidence from India," *Journal of the European Economic Association*, 6(2-3) April-May: 501-509.

Gine, Xavier, and D.S. Karlan (2011) "Group versus Individual Liability: Short and Long Term Evidence from Philippine Microcredit Lending Groups," June 2011, Mimeo, Yale University.

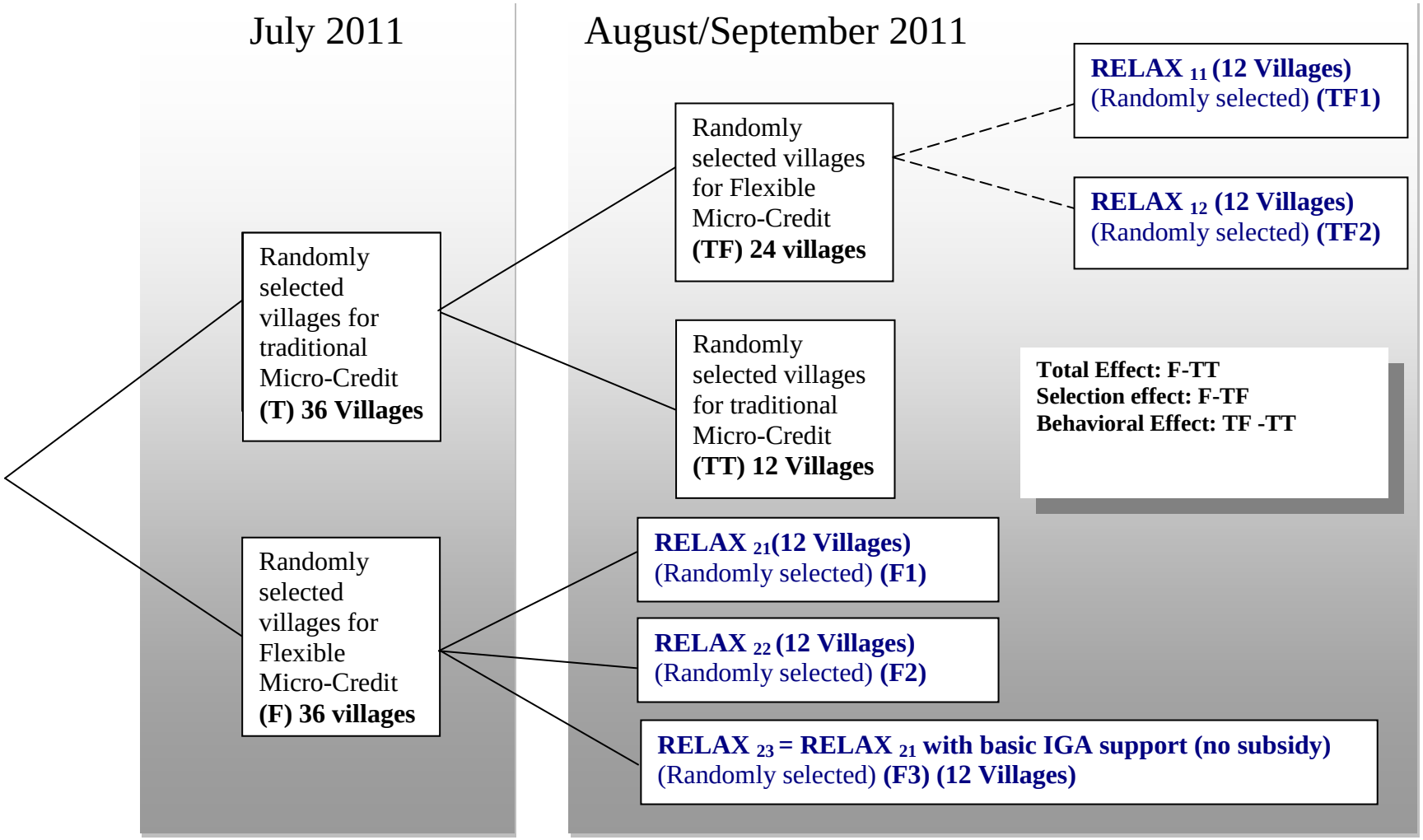
Karlan, Dean and Jonathan Zinman (2009) "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment," *Econometrica* 77(6) November 2009: 1993-2008.

Khandker, Shahidur R. (2012) "Seasonality of Income and Poverty in Bangladesh," *Journal of Development Economics* 97: 244-256.

Kurosaki, Takashi and Hidayat Ullah Khan (2012) "Vulnerability of Microfinance to Strategic Default and Covariate Shocks: Evidence from Pakistan," *The Developing Economies*, 50(2) June (in printing).

Shoji, Masahiro (2010) "Does Contingent Repayment in Microfinance Help the Poor During Natural Disasters?" *Journal of Development Studies*, 46(2): 191-210.

Shonchoy, Abu (2011) "Seasonal Migration and Micro-Credit in the Lean Period: Evidence from Northwest Bangladesh," Mimeo, Institute of Developing Economies, May 2011.



N = Borrowers (1080) + Non-borrowers (360) = 1440

Table 4.1: Distribution of sample villages by their treatment type, northern Bangladesh, 2011

	Treatment type						
	Surprise			Pre-announced			Grand total
	TF1 (Surprise flexible 1)	TF2 (Surprise flexible 2)	TT (Control)	F1 (Flexible 1)	F2 (Flexible 2)	F3 (Flexible 1+IGA)	
Total	12	12	12	12	12	12	72
By district and upazilla							
Gaibandha District	9	6	9	7	6	8	45
Fulchari	3	2	4	1	0	1	11
Gaibandha Sadar	4	3	1	2	4	7	21
Sundarganj	2	1	4	4	2	0	13
Kurigram District	3	6	3	5	6	4	27
Rajibpur	0	4	1	3	2	2	12
Rawmari	3	2	2	2	4	2	15
By location type							
Char	3	3	3	3	3	3	18
Inland	7	7	7	7	7	7	42
River-basin	2	2	2	2	2	2	12

Source: Compiled by the author using the benchmark survey described in the text.

Table 4.2: Mean of the distance from the sample village to a nearby station by their treatment type, northern Bangladesh, 2011

	Treatment type					
	Surprise			Pre-announced		
	TF1 (Surprise flexible 1)	TF2 (Surprise flexible 2)	TT (Control)	F1 (Flexible 1)	F2 (Flexible 2)	F3 (Flexible 1+IGA)
Number of observations (N)	12	12	12	12	12	12
Statistics of the distance (km)						
Mean	45.8	48.9	32.2	43.5	45.9	42.5
Std.Dev.	41.0	36.9	34.1	36.2	40.2	36.3
Minimum	3	9	4	4	5	8
Maximum	110	100	100	110	97	100
Test statistics of comparison of means vs. TT						
Abs.value of *t* -stat	0.882	1.155	(reference)	0.790	0.903	0.718
p -value	*0.388*	*0.261*	(reference)	*0.438*	*0.377*	*0.480*

Note: The *t* -statistic was calculated allowing for unequal variance. Pooling six treatment types together, ANOVA test was also conducted, yielding the acceptance of the null hypothesis of independence between treatment type and the distance (*p* -value 0.918).

Source: Compiled by the author using the benchmark survey described in the text.

Table 4.3: Definition and summary statistics of household-level variables compiled from the benchmark survey, northern Bangladesh, 2011

Variable	Definition	N	Mean	Std. Dev.	Min	Max
Characteristics of the household head						
age	Age of the head in years	1435	38.85	10.08	12	95
fem_dum	Dummy for a female head	1440	0.199	0.399	0	1
nonmuslim	Dummy for non-Muslim	1440	0.081	0.272	0	1
educ_lit1	Dummy for literacy defined as either "can write" or "can read"	1424	0.738	0.440	0	1
educ_lit2	Dummy for literacy defined as both "can write" and "can read"	1424	0.209	0.407	0	1
Aggregated characteristics of household members						
avg_age	Average age of all household members in years	1437	25.30	8.87	10.5	74
hhsize	Number of household members	1440	3.985	1.540	1	12
femrate	Ratio of females in "hhsize"	1440	0.534	0.209	0	1
adultrate	Ratio of members whose age is 15 or older in "hhsize"	1440	0.663	0.214	0.2	1
lit_rate1m	Adult literacy rate among male members, "either" definition for the literacy	1252	0.715	0.422	0	1
lit_rate2m	Adult literacy rate among male members, "both" definition for the literacy	1252	0.290	0.414	0	1
lit_rate1f	Adult literacy rate among female members, "either" definition for the literacy	1428	0.874	0.290	0	1
lit_rate2f	Adult literacy rate among female members, "both" definition for the literacy	1428	0.252	0.395	0	1

Note: N (the number of observations) is smaller than 1,440 when several missing observations existed. Mean and standard deviations are simple ones without weighting. In calculating adult literacy rates, those whose age is 15 or older are included.

Source: Compiled by the author using the benchmark survey described in the text.

Table 4.4: Mean comparison of household heads' characteristics by their treatment type, northern Bangladesh, 2011

	Treatment type					
	Surprise			Pre-announced		
	TF1 (Surprise flexible 1)	TF2 (Surprise flexible 2)	TT (Control)	F1 (Flexible 1)	F2 (Flexible 2)	F3 (Flexible 1+IGA)
age	37.79	39.10	39.28	39.08	39.25	38.62
p-value of mean-difference *t*-test vs. TT	*0.122*	*0.857*	(ref.)	*0.842*	*0.979*	*0.512*
p-value of three ANOVA tests	*0.232*		*0.570*		*0.763*	
fem_dum	0.183	0.138	0.250	0.250	0.250	0.121
p-value of mean-difference *t*-test vs. TT	*0.077*	*0.002*	(ref.)	*1.000*	*1.000*	*0.000*
p-value of three ANOVA tests	*0.007*		*0.000*		*0.000*	
nonmuslim	0.058	0.029	0.063	0.083	0.029	0.221
p-value of mean-difference *t*-test vs. TT	*0.848*	*0.081*	(ref.)	*0.381*	*0.081*	*0.000*
p-value of three ANOVA tests	*0.190*		*0.000*		*0.000*	
educ_lit1	0.806	0.684	0.685	0.738	0.749	0.767
p-value of mean-difference *t*-test vs. TT	*0.002*	*0.975*	(ref.)	*0.199*	*0.121*	*0.045*
p-value of three ANOVA tests	*0.003*		*0.016*		*0.769*	
educ_lit2	0.215	0.224	0.181	0.198	0.201	0.237
p-value of mean-difference *t*-test vs. TT	*0.346*	*0.245*		*0.625*	*0.576*	*0.130*
p-value of three ANOVA tests	*0.473*		*0.714*		*0.511*	

Note: For each variable, the first row shows the sample mean (unweighted). The number of observations is mostly 240 but sometimes less than 240 when missing values exist (see Table 4.3). The *t*-statistic was calculated allowing for unequal variance. The three ANOVA tests are for the null hypothesis of independence between treatment types and each variable, where treatment types are TF1, TF2, and TT for the first entry, all six for the second entry, and F1, F2, and F3 for the last entry.

Source: Compiled by the author using the benchmark survey described in the text.

Table 4.5: Mean comparison of household members' characteristics by their treatment type, northern Bangladesh, 2011

	Treatment type					
	Surprise			Pre-announced		
	TF1 (Surprise flexible 1)	TF2 (Surprise flexible 2)	TT (Control)	F1 (Flexible 1)	F2 (Flexible 2)	F3 (Flexible 1+IGA)
avg_age	25.28	24.73	26.27	25.89	24.85	24.76
p-value of mean-difference *t*-test vs. TT	*0.267*	*0.073*	(ref.)	*0.677*	*0.102*	*0.087*
p-value of three ANOVA tests	*0.176*		*0.271*		*0.281*	
hhsize	3.938	4.108	3.742	4.017	3.992	4.117
p-value of mean-difference *t*-test vs. TT	*0.141*	*0.007*	(ref.)	*0.062*	*0.066*	*0.007*
p-value of three ANOVA tests	*0.023*		*0.086*		*0.667*	
femrate	0.535	0.529	0.552	0.559	0.521	0.509
p-value of mean-difference *t*-test vs. TT	*0.387*	*0.228*	(ref.)	*0.746*	*0.106*	*0.022*
p-value of three ANOVA tests	*0.458*		*0.078*		*0.020*	
adultrate	0.658	0.639	0.703	0.675	0.624	0.677
p-value of mean-difference *t*-test vs. TT	*0.026*	*0.001*	(ref.)	*0.164*	*0.000*	*0.201*
p-value of three ANOVA tests	*0.004*		*0.001*		*0.007*	
lit_rate1m	0.778	0.662	0.663	0.682	0.737	0.763
p-value of mean-difference *t*-test vs. TT	*0.006*	*0.999*	(ref.)	*0.655*	*0.083*	*0.015*
p-value of three ANOVA tests	*0.007*		*0.007*		*0.126*	
lit_rate2m	0.295	0.290	0.263	0.284	0.311	0.295
p-value of mean-difference *t*-test vs. TT	*0.435*	*0.510*	(ref.)	*0.609*	*0.257*	*0.429*
p-value of three ANOVA tests	*0.704*		*0.919*		*0.815*	
lit_rate1f	0.866	0.896	0.830	0.881	0.887	0.887
p-value of mean-difference *t*-test vs. TT	*0.204*	*0.016*	(ref.)	*0.069*	*0.045*	*0.040*
p-value of three ANOVA tests	*0.053*		*0.149*		*0.961*	
lit_rate2f	0.204	0.293	0.219	0.237	0.260	0.298
p-value of mean-difference *t*-test vs. TT	*0.641*	*0.045*	(ref.)	*0.608*	*0.262*	*0.032*
p-value of three ANOVA tests	*0.027*		*0.042*		*0.242*	

Note: See Table 4.4.

Source: Compiled by the author using the benchmark survey described in the text.