Microcredit and Microsavings in Pakistan:
Pre-Analysis Plan

Uzma Afzal, Giovanna d'Adda, Marcel Fafchamps, Simon Quinn and Farah Said

Fieldwork locations:	Jhelum, Khushab, Mandi Bahauddin and Rawalpindi
Fieldwork dates:	October 2015 to September 2016
Date of Pre-Analysis Plan:	January 15, 2017

1 Introduction

This document outlines our pre-analysis plan for a microfinance field project. The document summarises (i) our experiment and resulting data and (ii) our plan of regressions.

This experiment was run in Jhelum, Khushab, Mandi Bahaudin and Rawalpindi from September 2015; we finished the endline questionnaire in September 2016. At the time of writing this plan, we have access to the administrative data on contract offers and take-up, which we have checked for general patterns and missing values; we have not yet run any substantive analysis on it. We are in the process of cleaning endline questionnaire data, and we anticipate beginning substantive analysis on this data in February 2017.

We intend to submit this Pre-Analysis Plan to the AEA RCT Registry.

2 Sample and treatment

2.1 Description of the sample

Our sample consists of National Rural Support Programme (NRSP) female members who are currently, or have in the past, been clients of some microfinance products being offered by the NRSP. The sample was drawn from the NRSP client database. We conducted a baseline face-to-face interview with a sample of approximately 2500 women at either their home or their business (as they preferred).

2.2 Structure of the treatment

We used a '3 $\times$ 3' factorial design plus controls; the following table summarises the assignment to different product types, including the probability of assignment to treatment. (The details of these treatments are described in the protocol document, and will be described in detail in the academic paper in due course.)

Basic treatment with **no reminders** 1/12	**Basic treatment** with **respondent reminders** 1/12	**Basic treatment** with **peer reminders** 1/12
Sunk treatment with **no reminders** 1/12	**Sunk treatment** with **respondent reminders** 1/12	**Sunk treatment** with **peer reminders** 1/12
Flex treatment with **no reminders** 1/12	**Flex treatment** with **respondent reminders** 1/12	**Flex treatment** with **peer reminders** 1/12

Control group 3/12

Assignment was done by computer, after forming matched blocks of either 12 or 13 respondents. We created these blocks by first interacting 'dummy for running a business' with 'whether the respondent makes the final decision on spending' with 'respondent would use a loan for investing in assets'. We then sort by household income and assign blocks. We denote the randomisation block dummies by ϕ_s, and record them as the variables '`RandomisationBlockDummy*`'.

Once assigned to a product type, each respondent keeps that assignment throughout the experiment. Treated clients (that is, all clients not assigned to 'control') were each visited on three separate occasions by NRSP staff members. On each occasion, they were offered a microfinance contract with a randomly-drawn interest rate ($r \in \{-0.086, 0, 0.086\}$) and week of lump-sum payment ($p \in \{1, 8\}$).

3 Data

3.1 Construction of variables: Administrative data

We have three data sources: (i) administrative data, recording whether each respondent was treated, what interest rate and repayment time were offered, and whether the respondent agreed to the contract, (ii) baseline and endline face-to-face interviews, and (iii) short phone interviews, conducted after each experiment wave.

We will construct variables from the administrative data as follows:

VARIABLE	DEFINITION	SOURCE
$\texttt{ID}_i$	The ID code for individual i.	Research team data
s_i	The randomisation strata code for individual i.	Research team data
$\texttt{Treated}_i$	A dummy variable for whether individual i was assigned to receive microfinance offers (*i.e.* treated).	Research team data
FLEX_i	A dummy for whether individual i was assigned to one of the three 'flex' treatments.	Research team data
SUNK_i	A dummy for whether individual i was assigned to one of the three 'sunk' treatments.	Research team data
RESP_REMINDERS_i	A dummy for whether individual i was assigned to one of the three 'respondent reminder' treatments.	Research team data

PEER_REMINDERS$_i$	A dummy for whether individual i was assigned to one of the three 'peer reminder' treatments.	Research team data
a_{iw}	A dummy variable for whether individual i accepted the contract in experiment wave w.	NRSP data
r_{iw}	The interest rate offered in period t, such that $r = 8.6\%$, $r = 0\%$ or $r = -8.6\%$.	Individual contract offers.
p_{iw}	The week payment is received by individual i in wave w, such that $p = 1$ or $p = 8$.	Individual contract offers.
$rneg_{iw}$	A dummy variable equal to 1 when the interest rate in wave w is -0.086; 0 otherwise.	Individual contract offers.
$rpos_{iw}$	A dummy variable equal to 1 when the interest rate in period w is 0.086; 0 otherwise.	Individual contract offers.
$p1_{iw}$	A dummy variable equal to 1 when payment is received in the first week of the cycle in wave w; 0 otherwise.	Individual contract offers.
$p8_{iw}$	A dummy variable equal to 1 when payment is received in the eighth week of the cycle in wave w; 0 otherwise.	Individual contract offers.

3.2 Construction of variables: Interview data

We will construct *outcome variables* from interview data in the following way:

VARIABLE	DEFINITION	SOURCE (QUESTION NUMBER)
	OUTCOME FAMILY 1: FINANCE AND CONSUMPTION	
$assets\_5$	A dummy variable for whether individual i participates in a committee.	5.5
$asset\_totalowed$	The total amount owed by individual i.	Sum of 5.16, 5.22, 5.30 and 5.38

$consumption\_total$	The total amount of household consumption in the last month.	Total sum of (6.6 × 4), plus 6.7, plus (6.8 / 3)
$income\_2$	The total monthly income	6.2
$household\_asset\_value$	Total value of assets owned by the household	Sum of assets_2c1 to assets_2c19

OUTCOME FAMILY 2: ENTERPRISES

$business\_1$	A dummy variable for whether individual i runs a business.	3.1
$business\_2$	The number of businesses owned by individual i or her household.	3.2
$business\_totalcapital$	The total capital invested in the businesses owned by individual i or her household.	Sum of 3.6 and 3.7.
$business\_12a$	The total monthly sales of the business.	3.12a
$business\_14a$	The total monthly expenses of the business.	3.14a
$business\_netprofit1$	The total monthly profit of the business.	3.12a minus 3.14a
$business\_netprofit2$	The total monthly profit of the business.	Sum of 3.15 and 3.17

OUTCOME FAMILY 3: RESPONDENT ATTITUDES

$preferences\_2$	A dummy variable for whether individual i finds it hard to save	When 7.2 is greater than 3.
$empowerment\_2$	A variable for whether individual i has her opinions taken into consideration when making decisions.	Following Anderson (2008), an index created of $consumption\_8a$ to $consumption\_8q$ for which answer is ≤ 2 (dummy), using the inverse of covariance matrix at baseline.
$preferences\_1$	A dummy variable for whether individual i faces pressure to share cash on hand.	When 7.1 is greater than 3.

$empowerment\_1$	A variable for whether individual i needs to ask permission for making decisions.	Following Anderson (2008), an index created of $consumption\_7a$ to $consumption\_7i$ for which answer is greater than 1 (dummy), using the inverse of covariance matrix at baseline.

If a respondent does not own a business, we will code any business-related outcome as zero (rather than as missing); in this way, we will estimate the average unconditional effect of treatment.

3.3 Describing balance

Throughout our analysis of balance, take-up and product impact, we will cluster our errors at the level of the individual respondent.[1]

Before the main estimations are run, we will describe balance for each of the outcome variables listed from the interview data (that is, for all candidate outcome variables). Denote the value for any given covariate in the baseline survey as y_{i0}. Then, for each covariate separately, we will estimate the following (where we list Stata code beneath the estimating equation) two specifications.

First, we will describe whether the variable is balanced across the 10 cells in our factorial design:

$$\begin{aligned} y_{i0} = \beta_0 + \beta_1 \cdot \text{BASIC}_i \times \text{NO\_REMINDERS} + \beta_2 \cdot \text{BASIC}_i \times \text{RESP\_REMINDERS}_i + \beta_3 \cdot \text{BASIC}_i \times \text{PEER\_REMINDERS}_i + \\ \beta_4 \cdot \text{SUNK}_i \times \text{NO\_REMINDERS} + \beta_5 \cdot \text{SUNK}_i \times \text{RESP\_REMINDERS}_i + \beta_6 \cdot \text{SUNK}_i \times \text{PEER\_REMINDERS}_i + \\ \beta_7 \cdot \text{FLEX}_i \times \text{NO\_REMINDERS} + \beta_8 \cdot \text{FLEX}_i \times \text{RESP\_REMINDERS}_i + \beta_9 \cdot \text{FLEX}_i \times \text{PEER\_REMINDERS}_i + \varepsilon. \end{aligned} \tag{1}$$

[1] In any estimations in which we have a single outcome observation per respondent, this is equivalent to simply using 'robust' standard errors.

We will estimate while clustering by respondent. We will then report a p-value from a joint test of the following null hypothesis:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0. \tag{2}$$

Second, we will test balance among contract offers (among those individuals who received offers). To do this, we will pool contract offers across all three waves, and estimate the following (again, clustering by respondent):

$$\begin{aligned} y_{i0} = \gamma_0 \cdot p1_{iw} + \gamma_1 \cdot p8_{iw} + \gamma_2 \cdot rneg_{iw} \times p1_{iw} + \gamma_3 \cdot rneg_{iw} \times p8_{iw} \\ + \gamma_4 \cdot rpos_{iw} \times p1_{iw} + \gamma_5 \cdot rpos_{iw} \times p8_{iw} + \mu_{iw}. \end{aligned} \tag{3}$$

We will then report a p-value from a joint test of the following null hypothesis :

$$H_0 : \gamma_0 = \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = \gamma_5. \tag{4}$$

We will then obtain a vector of p-values for covariate balance (one p-value for each covariate), and a vector of p-values for contract offers (again, one p-value for each covvariate). For each of these two vectors separately, we will calculate a vector of sharpened q values.

We view this exercise as essentially being descriptive: as showing how covariates differ across treatments. We view this as distinct from the question of whether any of these covariates should enter our regressions as controls; we discuss that issue in section 8.

4 Analysis: Determinants of take-up and compliance

4.1 Take-up

Simulated assignment in cases of immediate refusal: We plan to address six primary research questions on the determinants of take-up. In each case, we will test how take-up varies with the contractual terms offered. In some cases, respondents refused immediately, without even waiting to see the contractual terms that were randomly drawn. In such cases, we will randomly simulate contractual terms, *as if* such terms had been randomly drawn in the field. This is necessary in order to consistently estimate the effect of contractual terms on take-up; if we were simply to code such observations as missing, we would overstate the demand for our product. We will run this simulated assignment once at the start of our analysis, saving the simulated assignment for use in all subsequent estimation.

Research Question 1 *How does take-up vary with contract flexibility?*

To answer this question, we will compare average take-up between the 'basic treatment', the 'sunk treatment' and the 'flex treatment'. We will control for strata dummies (ϕ_s), and also for 'district-wave' effects (η_{dw}):

$$a_{iw} = \alpha_0 + \alpha_1 \cdot \text{SUNK}_i + \alpha_2 \cdot \text{FLEX}_i + \phi_s + \eta_{dw} + \varepsilon_{iw} \tag{5}$$

```
ivreg2 a Sunk Flex RandomisationBlockDummy* DistrictWaveDummy*,
partial(RandomisationBlockD* DistrictWaveDummy*) cluster(ID)
```

Research Question 2 *How does take-up vary with reminders?*

To answer this question, we will compare average take-up between the 'no reminders' treatment, the 'respondent reminders' treatment and the 'peer reminders' treatment. We will again control for strata dummies (ϕ_s), and also for 'district-wave' effects (η_{dw}):

$$a_{iw} = \beta_0 + \beta_1 \cdot \texttt{RESP\_REMINDERS}_i + \beta_2 \cdot \texttt{PEER\_REMINDERS}_i + \phi_s + \eta_{dw} + \varepsilon_{iw} \tag{6}$$

```
ivreg2 a RespondentRem PeerRem RandomisationBlockDummy* DistrictWaveDummy*,
 partial(RandomisationBlockD* DistrictWaveDummy*) cluster(ID)
```

Research Question 3 *Are there interaction effects between contract flexibility and payment reminders?*

The previous specifications estimate the average treatment effects of contract flexibility and reminders respectively. To understand further the potential mechanisms that might underpin these estimates, we will test whether there are interaction effects between contract flexibility and payment reminders. To do this, we will estimate the following:

$$\begin{aligned} a_{iw} &= \gamma_0 + \gamma_1 \cdot \text{SUNK}_i + \gamma_2 \cdot \text{FLEX}_i + \gamma_3 \cdot \text{RESP\_REMINDERS}_i + \gamma_4 \cdot \text{PEER\_REMINDERS}_i \\ &+ \gamma_5 \cdot \text{SUNK}_i \times \text{RESP\_REMINDERS}_i + \gamma_6 \cdot \text{SUNK}_i \times \text{PEER\_REMINDERS}_i \\ &+ \gamma_7 \cdot \text{FLEX}_i \times \text{RESP\_REMINDERS}_i + \gamma_8 \cdot \text{FLEX}_i \times \text{PEER\_REMINDERS}_i + \phi_s + \eta_{dw} + \varepsilon_{iw} \end{aligned} \tag{7}$$

```
ivreg2 a Sunk Flex RespondentRem PeerRem Sunk_RespondentRem Sunk_PeerRem
 Flex_RespondentRem Flex_PeerRem RandomisationBlockDummy* DistrictWaveDummy*,
 partial(RandomisationBlockD* DistrictWaveDummy*) cluster(ID)
```

We will run a joint test of the null hypothesis that there are no interaction effects:

$$H_0 : \gamma_5 = \gamma_6 = \gamma_7 = \gamma_8 = 0.$$

Research Question 4 *How does take-up vary with interest rate?*

To answer this question, we will compare average take-up between the 'zero interest rate' offers (*i.e.* 3500 rupees paid, 3500 rupees recovered), 'positive interest rate' offers (*i.e.* 3500 rupees paid, 3800 rupees recovered) and 'negative interest rate' offers (*i.e.* 3500 rupees paid, 3200 rupees recovered). We will again control for strata dummies (ϕ_s), and also for 'district-wave' effects (η_{dw}):

$$a_{iw} = \gamma_0 + \gamma_1 \cdot \texttt{rneg}_{iw} + \gamma_2 \cdot \texttt{rpos}_{iw} + \phi_s + \eta_{dw} + \varepsilon_{iw} \tag{8}$$

```
ivreg2 a NegativeRate PositiveRate RandomisationBlockDummy* DistrictWaveDummy*,
 partial(RandomisationBlockD* DistrictWaveDummy*) cluster(ID)
```

Research Question 5 *How does take-up vary with time of lump-sum payment?*

To answer this question, we will compare average take-up between contract offers in which the lump-sum is offered in week 1 and contract offers in which the lump-sum is offered in week 8 (which we denote with a dummy $p8_{iw}$):

$$a_{iw} = \delta_0 + \delta_1 \cdot p8_{iw} + \phi_s + \eta_{dw} + \varepsilon_{iw} \tag{9}$$

```
ivreg2 a PaymentWeek8 RandomisationBlockDummy* DistrictWaveDummy*,
 partial(RandomisationBlockD* DistrictWaveDummy*) cluster(ID)
```

Research Question 6 *Does the demand for flexibility and the demand for reminders differ between savings and credit contracts?*

To answer this question, we will repeat the estimations in equations 5, 6 and 7, disaggregating by whether $p8_{iw} = 0$ (that is, lump sum payment in week 1) or $p8_{iw} = 1$ (that is, lump sum payment in week 8). We will run cross-equation tests to test whether the demand for flexibility and the demand for reminders differs between savings and credit contracts.

Research Question 7 *Do the same clients demand both credit and savings products?*

To answer this question, we will repeat the descriptive stylised facts that we provide in our working paper from Sargodha (Afzal, d'Adda, Fafchamps, Quinn, and Said, 2016). Specifically, we plan to report the following (calculated, in each case, for the set of respondents who completed all three experimental waves):

(i). The proportion who accepted all contracts offered and the proportion who accepted none of the contracts offered;

(ii). The proportion who accepted both a credit contract and a savings contract;

(iii). The proportion who accepted both a savings contract with $r > 0$ and a credit contract with $r \leq 0$;

(iv). The proportion who accepted at least one savings contract with $r < 0$ (conditional on having been offered at least once such contract);

(v). The proportion who rejected at least on credit contract with $r > 0$ (conditional on having been offered at least one such contract).

4.2 Non-compliance

The preceding discussion relates to the outcome variable a_{iw}: a dummy for whether, in wave w, individual i agreed to the offered contract. Having agreed to a contract, there are two potential forms of non-compliance:

(i). A respondent can indicate that she no longer wants to proceed, having neither made or received any payments. We term this as 'immediate non-compliance'; in effect, the respondent is reconsidering her decision to agree, before proceeding with the contract.

(ii). A respondent can, after having made or received a payment, refuse to make any further payments. We term this as 'default'.

We will repeat the analysis in Research Questions 1 to 5, replacing the outcome variable a_{iw} with two alternative outcome variables:

(i). A dummy variable for immediate non-compliance; and

(ii). A dummy variable for default.

In this way, we will test separately whether immediate non-compliance and default are systematically related to the treatments offered.

5 Analysis: Effects on endline outcomes

Approach to multiple hypothesis testing: In this section, we set out a series of tests for the set of outcome variables described in section 3.2. For each estimation, we will report (i) the p-value for the estimated treatment effect (*i.e.* whether ITT or LATE), and (ii) a sharpened q-value, calculated within each listed outcome family (see Benjamini, Krieger, and Yekutieli (2006) and Anderson (2008)). (That is, we will calculate sharpened q-values separately for Outcome Family 1, Outcome Family 2 and Outcome Family 3.) In the following section, we set out a series of tests for two separate outcome families (which we describe as Outcome Family 4 and Outcome Family 5); for each of those estimations, we will also report (i) the p-value for the estimated treatment effect (*i.e.* whether ITT or LATE), and (ii) a sharpened q-value, calculated within each listed outcome family.

Research Question 8 *[ITT] What is the impact of being offered the product?*

Denote y_{i1} as the endline value for individual i for some variable; denote y_{i0} as the baseline value. Denote ϕ_s as a common parameter for stratum s. Then, for each variable in the previous table, we will estimate the following ANCOVA specification (where `y` denotes the outcome variable, and `y_pre` the baseline value of that outcome variable). We will again control for strata dummies (ϕ_s), and also for district effects (η_d):

$$y_{i1} = \beta_0 + \beta_1 \cdot \texttt{Treated}_i + \beta_2 \cdot y_{i0} + \phi_s + \eta_d + \varepsilon_i \tag{10}$$

```
ivreg2 y Treated y_pre RandomisationBlockD* DistrictD*,
partial(RandomisationBlockD* DistrictD*) cluster(ID)
```

Research Question 9 (LATE) *What is the impact of accepting the product?*

To estimate the impact of ever accepting the product, we will instrument adoption by treatment. Denote `EverAdopted` as a dummy for whether individual i ever accepted a product. Then we will estimate the following (where ϕ_{1s} and η_{1d} are strata dummies and district dummies in the first stage and ϕ_{2s} and η_{2d} are strata dummies and district dummies in the second stage) and :

$$y_{i1} = \beta_0 + \beta_1 \cdot \texttt{EverAdopted}_i + \beta_2 \cdot y_{i0} + \phi_{2s} + \eta_{2d} + \varepsilon_i \tag{11}$$

$$\texttt{EverAdopted}_i = \gamma_0 + \gamma_1 \cdot \texttt{Treated}_i + \gamma_2 \cdot y_{i0} + \phi_{1s} + \eta_{1d} + \mu_i \tag{12}$$

```
ivreg2 y (EverAdopted = Treated) y_pre RandomisationBlockD* DistrictD*,
partial(RandomisationBlockD* DistrictD*) cluster(ID)
```

Research Question 10 (ITT) *What is the impact of being offered more or less flexible products?*

The previous questions relate to the *average* effect of being offered (and then accepting) across different kinds of products. To understand further the mechanisms for any estimated effects, we will test for heterogeneity by the specific types of products offered. Again, we exploit the factorial design; again, we start by testing the effects of product flexibility.

We begin by testing the effects of being offered the ‘sunk’ or ‘flex’ products. As before, we denote $\texttt{SUNK}_i$ and $\texttt{FLEX}_i$ as dummy variables for whether respondent i was offered the sunk and flex options. We will estimate:

$$y_{i1} = \beta_0 + \beta_1 \cdot \texttt{Treated}_i + \beta_2 \cdot \texttt{SUNK}_i + \beta_3 \cdot \texttt{FLEX}_i + \beta_4 \cdot y_{i0} + \phi_s + \eta_d + \varepsilon_i \tag{13}$$

```
ivreg2 y Treated Sunk Flex y_pre RandomisationBlockD* DistrictD*,
 partial(RandomisationBlockD* DistrictD*) cluster(ID)
```

In this estimation, the variables `SUNK` and `FLEX` capture interaction effects — *i.e.* the additional effect on y_{i1} relative to already having $\texttt{Treated} = 1$. We will therefore test for interaction effects by testing the null hypotheses $H_0 : \beta_2 = 0$, $H_0 : \beta_3 = 0$ and $H_0 : \beta_2 = \beta_3 = 0$.

Research Question 11 (LATE) *What is the impact of accepting more or less flexible products?*

To understand the effect of accepting more or less flexible products, we will again estimate the LATE. We will do this by instrumenting adoption of each type of product by whether each type of product was offered. Specifically, denote $\texttt{EverAdopted\_Flex}_i$ as a dummy for whether individual i ever adopted the 'flex' product, and denote $\texttt{EverAdopted\_Sunk}_i$ as a dummy for having ever adopted the 'sunk' product. Then we will estimate:

```
ivreg2 y (EverAdopted EverAdopted_Flex EverAdopted_Sunk = Treated Flex Sunk)
 y_pre RandomisationBlockD* DistrictD*,
 partial(RandomisationBlockD* DistrictD*) cluster(ID)
```

We will then test for heterogeneous effects by testing the significance of the coefficients on `EverAdopted_Flex` and `EverAdopted_Sunk`, and testing jointly whether both are zero.

Research Question 12 (ITT) *What is the impact of being offered products with reminders?*

Symmetrically, we will then proceed by testing the effect of the different reminder treatments. We will estimate the ITT by estimating:

$$y_{i1} = \beta_0 + \beta_1 \cdot \texttt{Treated}_i + \beta_2 \cdot \texttt{RESP\_REMINDERS}_i + \beta_3 \cdot \texttt{PEER\_REMINDERS}_i + \beta_4 \cdot y_{i0} + \phi_s + \eta_d + \varepsilon_i \tag{14}$$

```
ivreg2 y Treated RespondentRem PeerRem y_pre RandomisationBlockD* DistrictD*,
 partial(RandomisationBlockD* DistrictD*) cluster(ID)
```

As before, we then test for interaction effects by testing the null hypotheses $H_0 : \beta_2 = 0$, $H_0 : \beta_3 = 0$ and $H_0 : \beta_2 = \beta_3 = 0$.

Research Question 13 (LATE) *What is the impact of accepting products with reminders?*

Finally, we will again estimate the LATE — this time, interacting with dummies for the reminder treatments. Denote $\texttt{EverAdopted\_RespondentRem}_i$ as a dummy for whether individual i ever adopted a product with respondent reminders; $\texttt{EverAdopted\_PeerRem}_i$ records having ever adopted a product with peer reminders. Then we will estimate:

```
ivreg2 y (EverAdopted EverAdopted_RespondentRem EverAdopted_PeerRem
  = Treated RespondentRem PeerRem) y_pre RandomisationBlockD* DistrictD*,
 partial(RandomisationBlockD* DistrictD*) cluster(ID)
```

We will then test for heterogeneous effects by testing the significance of the coefficients on `EverAdopted_RespondentRem` and `EverAdopted_PeerRem`, and testing jointly whether both are zero.

6 Analysis: Using the phone interviews

We conducted a phone interview at the end of the each experiment wave. We will construct *outcome variables* from the phone interview data in the following way:

VARIABLE	DEFINITION	SOURCE (QUESTION NUMBER)
OUTCOME FAMILY 4: FINANCE AND CONSUMPTION (PHONE DATA)		
$assets\_P2$	The change in total value of assets owned by individual i or her household in the last one month.	2.2
$consumption\_P2$	The total amount of household consumption in the last month.	Total sum of (2.3 $\times$ 4), plus 2.4
$borrow\_P2a$	The loan amount borrowed by the individual i or her household in the last one month.	Sum of 4.3
$borrow\_P3a$	The loan amount repaid by the individual i or her household in the last one month.	Sum of 4.5
OUTCOME FAMILY 5: ENTERPRISES AND EMPLOYMENT (PHONE DATA)		
$business\_P1$	A dummy variable for whether individual i runs a business.	3.1
$business\_P2$	The number of businesses owned by individual i or her household.	3.2
$business\_P18$	The total value of good and services re-invested in the businesses owned by individual i or her household in the last month.	3.3.
$wage\_P1$	A dummy variable if individual i has a regular wage job.	3.4.

We will estimate with the phone data using the same identification strategy as in the previous section (where, again, y_{i0} shall refer to the baseline value). We will estimate in two ways:

(i). By pooling all phone data, across all waves, and

(ii). By estimating separately by wave.

7 Analysis: Effects of treatment on women's empowerment

At endline, we ran an 'empowerment game', in which we allow the respondent and her spouse to make separate decisions about the form of a gift for having participated in the experiment. We have two primary outcomes from this game:

(i). The man's willingness to delegate to his wife:

  (a) Whether the man would delegate for no additional reward (`choice_2`);

  (b) Whether the man would delegate for an additional reward of 50 rupees (`choice_3a`); and

  (c) Whether the man would delegate for an additional reward of 200 rupees (`choice_3b`).

(ii). The women's willingness to insist upon her choice:

  (a) Whether the woman would forfeit a voucher of 50 rupees to ensure she gets her choice (`choice_2_4`); and

  (b) Whether the woman would forfeit a voucher of 200 rupees to ensure she gets her choice (`choice_2_5`).

Research Question 14 *[ITT] What is the impact on empowerment of being offered the product?*

We will test the effect of treatment on each of these measures, both using ITT and LATE specifications. That is, we will run the following regression, replacing y_{i1} in each case with the measures of willingness to delegate (for the men) and to guarantee own choice (women). We will again control for strata dummies

(ϕ_s), and also for district effects (η_w):

$$y_{i1} = \beta_0 + \beta_1 \cdot \texttt{Treated}_i + \phi_s + \eta_d + \varepsilon_i \tag{15}$$

```
ivreg2 y Treated RandomisationBlockD* DistrictD*,
 partial(RandomisationBlockD* DistrictD*) cluster(ID)
```

Research Question 15 (LATE) *What is the impact on empowerment of accepting the product?*

To estimate the impact of ever accepting the product, we will again instrument adoption by treatment, again repeating our estimation for the five separate empowerment measures. We will estimate:

$$y_{i1} = \beta_0 + \beta_1 \cdot \texttt{EverAdopted}_i + \phi_{1s} + \eta_{1d} + \varepsilon_i \tag{16}$$

$$\texttt{EverAdopted}_i = \gamma_0 + \gamma_1 \cdot \texttt{Treated}_i + \phi_{2s} + \eta_{2d} + \mu_i \tag{17}$$

```
ivreg2 y (EverAdopted = Treated) RandomisationBlockD* DistrictD*,
 partial(RandomisationBlockD* DistrictD*) cluster(ID)
```

8 Robustness: Inclusion of other controls

As discussed, we implemented our experiment as a Randomised Controlled Trial; therefore, our estimates are unbiased. However, as with every experiment, we may be concerned that, in a finite sample, imbalance on some covariates may be problematic. Therefore, as a robustness check, we will repeat our main estimation using 'post-double-selection' with LASSO, where we use all of the baseline variables in Outcome Family 1, Outcome Family 2 and Outcome Family 3 as potential controls (see Belloni, Chernozhukov, and Hansen (2014a), Belloni, Chernozhukov, and Hansen (2014b) and see Bubb, Kaur, and Mullainathan (2016) for an example of this approach in the context of a randomised field experiment).

9 Robustness: Attrition

Finally, we will test for differential attrition between treatments. We will do this by coding a dummy variable for whether individual i attrited for the endline survey, and we will repeat the analysis of Section 5 using this dummy variable as an outcome.[2] We will then repeat the exercise for each of the separate rounds in the phone survey, using the identification strategy summarised in Section 3.2. If we find significant differences in attrition between treatments, we will use Lee (2009) bounds to check the robustness of our analysis on Research Questions 8 to 14.

References

AFZAL, U., G. D'ADDA, M. FAFCHAMPS, S. QUINN, AND F. SAID (2016): "Two Sides of the Same Rupee? Comparing Demand for Microcredit and Microsaving in a Framed Field Experiment in Rural Pakistan," *Working Paper: http://www.simonrquinn.com/research/TwoSidesOfTheSameRupee.pdf*.

ANDERSON, M. L. (2008): "Multiple Inference and Gender Differences in the Effects of Early ntervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 103(484).

BELLONI, A., V. CHERNOZHUKOV, AND C. HANSEN (2014a): "High-Dimensional Methods and Inference on Structural and Treatment Effects," *The Journal of Economic Perspectives*, 28(2), 29–50.

——— (2014b): "Inference on Treatment Effects After Selection Among High-dimensional Controls," *The Review of Economic Studies*, 81(2), 608–650.

BENJAMINI, Y., A. M. KRIEGER, AND D. YEKUTIELI (2006): "Adaptive Linear Step-up Procedures that Control the False Discovery Rate," *Biometrika*, 93(3), 491–507.

BUBB, R., S. KAUR, AND S. MULLAINATHAN (2016): "Barriers to Contracting in Village Economies: A Test for Enforcement Constraints," *Working paper*.

LEE, D. S. (2009): "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *The Review of Economic Studies*, 76(3), 1071–1102.

[2] Of course, by definition, such a dummy is zero at baseline; therefore, the previous specifications each collapses to a 'difference' estimation rather than an ANCOVA estimation.