# Microfinance field experiment in Bhakkar and Chakwal: Pre-Analysis Plan

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

**Fieldwork location:** Bhakkar and Chakwal (Pakistan)
**Fieldwork dates:** 25 August 2014 to 15 April 2015 (expected)
**Date of Pre-Analysis Plan:** 7 April 2015

**Abstract**

We are currently running a field experiment of a novel microfinance product, inspired by the rotating structure of a ROSCA. This is a scaled-up version of the product structure presented in CSAE Working Paper WPS/2014-32 (Afzal, d'Adda, Fafchamps, Quinn, and Said (2014), '*Two Sides of the Same Rupee? Comparing Demand for Microcredit and Microsaving in a Framed Field Experiment in Rural Pakistan*'). We are currently approaching the completion of fieldwork.

## 1 Introduction

This document outlines our pre-analysis plan for a microfinance field project; it is based on the recommendations of McKenzie (28 October 2012, Development Impact Blog). Specifically, the document summarises (i) our experiment and resulting data and (ii) our plan of regressions.

This experiment was run in Bhakkar and Chakwal from 25 August 2014; we anticipate finishing the endline questionnaire by approximately 15 April 2015. 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 missing values;

we have not yet run any substantive analysis on it. We have access to the baseline questionnaire data, but not to the endline questionnaire data; we anticipate receiving this data by mid-May.

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

# 2 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 files at the NRSP field offices in Bhakkar and in Chakwal. We conducted a baseline face-to-face interview with a sample of approximately 800 women at either their home or their business (as they preferred). Approximately half of the sample were assigned to treatment and half to control; this was done by blocked randomisation in Stata. Treated clients 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 ($r \in \{-0.1, 0, 0.1\}$) and week of lump-sum payment ($p \in \{1, 6\}$).[1]

Our estimations will include all individuals who agreed to participate in the trial (i.e. not just those who participated in all three experiment waves). In addition, we will report the default rates for individuals who agree to participate in the product offered in a particular round but fail to make the installments or drop out between the round has completed. We will also report their reasons for dropping out to ascertain their preferences for the product offered.

[1] This follows closely the contractual design in Afzal, d'Adda, Fafchamps, Quinn, and Said (2014).

# 3 Data

## 3.1 Construction of variables: Administrative data

We have two 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, and (ii) baseline and endline face-to-face interviews.

We will construct variables from the administrative data as follows:

| VARIABLE | DEFINITION | SOURCE |
|---|---|---|
| $\texttt{mohallah}_i$ | The *mohallah* recorded for individual $i$. | NRSP data |
| $s_i$ | The randomisation strata code for individual $i$. | Research team data |
| $m_i$ | A dummy variable for whether individual $i$ was assigned to receive microfinance offers (*i.e.* treated). | Research team data |
| $a_{iw}$ | A dummy variable for whether individual $i$ accepts the contract in experiment wave $w$. | NRSP data |
| $r_{iw}$ | The interest rate offered in period $t$, such that $r = 10\%$, $r = 0\%$ or $r = -10\%$. | Individual contract offers. |
| $p_{iw}$ | The week payment is received by individual $i$ in wave $w$, such that $p = 1$ or $p = 6$. | Individual contract offers. |
| $rneg_{iw}$ | A dummy variable equal to 1 when the interest rate in wave $w$ is -0.1; 0 otherwise. | Individual contract offers. |
| $rpos_{iw}$ | A dummy variable equal to 1 when the interest rate in period $w$ is 0.1; 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. |
| $p6_{iw}$ | A dummy variable equal to 1 when payment is received in the sixth 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: BUSINESS OWNERSHIP** | |
| $business\_1$ | A dummy variable for whether individual $i$ runs a business. | 3.1 |
| $business\_2$ | A variable for the number of businesses owned by individual $i$ or her household. | 3.2 |
| $business\_6$ | A variable for the total value of the assets invested in the business owned by individual $i$ or her household. | 3.6. |
| $business\_7$ | A variable for the total value of the working capital for the business owned by individual $i$ or her household. | 3.7. |
| $business\_share$ | A variable for the capital share invested in the business jointly owned by individual $i$ or her household | Ratio of 3.9 and (3.9 + 3.10) for those who jointly own a business as specified in 3.8. |
| $business\_totalcapital$ | A variable for the total capital invested in the businesses owned by individual $i$ or her household. | Sum of 3.6 and 3.7. |
| | **OUTCOME FAMILY 2: BUSINESS PERFORMANCE** | |
| $business\_11$ | A dummy variable for whether written accounts are kept for the business. | 3.11 |
| $business\_12$ | A dummy variable for whether services of an accountant are used for the business. | 3.12 |
| $business\_13$ | A variable for the total monthly sales of the business. | 3.13a |
| $business\_15$ | A variable for the total monthly expenses of the business. | 3.15a |
| $business\_netprofit1$ | A variable for the total monthly profit of the business. | 3.13a minus 3.15a |
| $business\_netprofit2$ | A variable for the total monthly profit of the business. | Sum of 3.16 and 3.18 |
| | **OUTCOME FAMILY 3: FINANCE** | |

| | | |
|---|---|---|
| $asset\_5$ | A dummy variable for whether individual $i$ participates in a committee. | 4.5 |
| $asset\_totalowed$ | A variable for the total amount owed by individual $i$. | Sum of 4.17, 4.24, 4.33 and 4.42 |
| $asset\_loancount$ | A variable for the total number of outstanding loans owed by individual $i$. | Sum of 4.16, 4.23, 4.32 and 4.41 |

**OUTCOME FAMILY 4: HOUSEHOLD ASSETS AND CONSUMPTION**

| | | |
|---|---|---|
| $assets\_purchased$ | A variable for the value of assets purchased by the household in the last 5 months. | Sum of 4.2c for cases in which 4.2d $\leq$ 5. |
| $consumption\_total$ | A variable for the total amount of household consumption in the last month. | Total sum of (5.1 $\times$ 4), plus 5.2, plus (5.3 / 3) |
| $household\_asset\_count$ | Number of different assets owned by the household | Sum of assets_2a1 to assets_2a16 |
| $household\_asset\_value$ | Total value of assets owned by the household | Sum of assets_2c1 to assets_2c16 |
| $individual\_asset\_count$ | Number of different assets owned by the individual | Sum of assets_2a1 to assets_2a16, interacted in each case with dummies for whether assets_2b1 to assets_2b16 are 1. |
| $individual\_asset\_value$ | Total value of assets owned by the individual | Sum of assets_2c1 to assets_2c16, interacted in each case with dummies for whether assets_2b1 to assets_2b16 are 1. |

**OUTCOME FAMILY 5: RESPONDENT ATTITUDES**

| | | |
|---|---|---|
| $doing\_2$ | A dummy variable for whether individual $i$ finds it hard to save | When 6.2 is greater than 3. |

| | | |
|---|---|---|
| $time\_preference$ | A measure of time preference. | Questions 6.12 to 6.23. Take the value for the first switch, subtract 5000 and divide by 5000. |
| $risk$ | A measure of risk aversion. | Question 6.24. Take the value of the first switch; calculate the smallest CRRA parameter that would justify such a switch. |
| $altruism$ | A index of altruism. | Constructed from from 6.25 to 6.27. |
| $posrec$ | A index of positive reciprocity. | Constructed from 6.28 to 6.30. |
| $trust$ | A index of trust. | An index constructed from 6.31 to 6.38, using the inverse of covariance matrix at baseline. |

**OUTCOME FAMILY 6: RELATIONSHIP WITH FAMILY MEMBERS**

| | | |
|---|---|---|
| $business\_26$ | A variable for the attitude of other family members towards the business of individual $i$. | Dummy for whether 3.26 is 1 ('They are supportive.') |
| $business\_27$ | A variable for the attitude of spouse towards the business of individual $i$. | Dummy for whether 3.27 is 1 ('He is supportive.') |
| $empowerment\_2$ | A variable for whether individual $i$ has her opinions taken into consideration when making decisions. | Following Anderson (2008), an index created of 5.8 for which answer is $\leq 2$ (dummy), using the inverse of covariance matrix at baseline. |

| | | |
|---|---|---|
| $empowerment\_3$ | A variable for whether individual $i$ prefers to choose by herself. | Following Anderson (2008), an index created of 5.23, 5.24, 5.25 and 5.26 for which answer is 1 (dummy) using the inverse of covariance matrix at baseline |
| $doing\_1$ | A dummy variable for whether individual $i$ faces pressure to share cash on hand. | When 6.1 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 5.7 for which answer is greater than 1 (dummy), using the inverse of covariance matrix at baseline. |
| $agency\_1$ | A variable for the preference for agency exhibited by individual $i$. | When 6.40 is greater than 3 |
| $agency\_2$ | A variable for the preference for agency exhibited by individual $i$. | When 6.41 is greater than 3 |

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.

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

| VARIABLE | DEFINITION | SOURCE (QUESTION NUMBER) |
|---|---|---|
| $intro\_11$ | The age of individual $i$. | 1.11 |
| $ever\_married$ | Dummy: Whether individual $i$ was ever married. | Dummy for whether 1.12 is greater than 1 |
| $currently\_married$ | Dummy: Whether individual $i$ is currently married. | Dummy for whether 1.12 is 2 |
| $education\_years$ | Number of years of education | Constructed from 1.13 |

| | | |
|---|---|---|
| $literate$ | Dummy: Individual $i$ can read and write | Dummy for whether 1.14 is 1 |
| $children\_count$ | Number of children the respondent has | Sum of household_4a and household_4b |
| $household\_5$ | Total number of people in the household | 2.5 |
| $household\_head$ | Dummy: Individual $i$ is the household head | Dummy: 2.7 is 1 |
| $home\_ownership\_self$ | Dummy: Individual $i$ owns the household home | Dummy: 4.1 is 1 |
| $home\_ownership\_husband$ | Dummy: Individual $i$'s husband owns the household home | Dummy: 4.1 is 2 |
| $home\_ownership\_joint$ | Dummy: Individual $i$ owns the household home jointly with her husband | Dummy: 4.1 is 3 |

## 3.3 Testing balance

Throughout our analysis of balance, take-up and product impact, we will cluster our errors at the level of the *mohallah* – that is, the *mohallah* in which the respondent is regularly contacted. (The concept of the *mohallah* is well understood in Pakistan; there is no ambiguity in the records that NRSP provided us about which respondent is in which *mohallah*. We will use the assignment in these records.)

Before the main estimations are run, we will check for balance for each of the outcome variables listed from the interview data (that is, both for outcome variables and for control 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):

$$y_{i0} = \beta_0 + \beta_1 \cdot m_i + \varepsilon_i \tag{1}$$

```
reg y treated, cluster(mohallah)
```
(2)

$$y_{i0} = \gamma_0 \cdot p1_{iw} + \gamma_1 \cdot p6_{iw} + \gamma_2 \cdot rneg_{iw} \times p1_{iw} + \gamma_3 \cdot rneg_{iw} \times p6_{iw} + \gamma_4 \cdot rpos_{iw} \times p1_{iw} + \gamma_5 \cdot rpos_{iw} \times p6_{iw} + \mu_{iw}. \tag{3}$$

```
reg y p1 p6 rnegp1 rnegp6 rposp1 rposp6, nocons cluster(mohallah)
```
(4)

For equation 3, we will estimate by pooling contract offers across all three waves. For equation 1, we will test balance by testing $H_0 : \beta_1 = 0$. For equation 3, we will test balance by a joint test of $H_0 : \gamma_0 = \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = \gamma_5$. We will report a table of summary statistics for each variable, showing (i) the number of observations, (ii) the mean, (iii) the standard deviation, (iv) the first quartile, (v) the median, (vi) the third quartile, (vii) the minimum, (viii) the maximum, (ix) the $p$-value for testing on equation 1 and (x) the $p$-value for testing equation 3.

If, for a given variable, we do not reject $H_0$ at the 90% confidence level, we will conclude that this variable is 'balanced across treatments'. If we *do* reject at the 90% confidence level for a given variable, we will conclude that this variable is 'unbalanced across treatments'. We will then include that variable as a control, in the robustness section (discussed shortly). We note that this is a very conservative balance strategy, because we are running each balance test separately, and without correcting for multiple inference. We anticipate rejecting balance at the 90% confidence level on approximately 10% of variables; this should not then be interpreted as evidence of a failure of the experimental protocol.

# 4 Analysis: Determinants of take-up

We plan to address three primary research questions on the determinants of take-up:

**Research Question 1** *How does take-up vary with interest rate?*

We will estimate the following two equations to test for sensitivity to interest rates:

$$a_{iw} = \beta_0 + \beta_{neg} \cdot rneg_{iw} + \beta_{pos} \cdot rpos_{iw} + \varepsilon_{iw} \tag{5}$$

```
reg a rneg rpos, cluster(mohallah)
```
(6)

where zero interest rate is the omitted category.

**Research Question 2** *How does take-up vary with the week in which the NRSP payment is received?*

We will estimate the following two equations:

$$a_{iw} = \beta_0 + \beta_p \cdot p6_{iw} + \varepsilon_{iw} \tag{7}$$

```
reg a p6, cluster(mohallah)
```
(8)

making week 1 the omitted category.

**Research Question 3** *How does take-up vary with the joint effect of the interest rate and the day of NRSP payment?*

We will jointly test for difference in product take up by both interest rate and bank payment day using the following 'saturated' specification:

$$\begin{aligned} a_{iw} = {} & \beta_0 + \beta_1 \cdot rneg_{iw} + \beta_2 \cdot rpos_{iw} + \beta_3 p6_{iw} \\ & + \beta_4 \cdot rneg_{iw} \times p6_{iw} + \beta_5 \cdot rpos_{iw} \times p6_{iw} + \varepsilon_{iw} \end{aligned} \tag{9}$$

```
reg a rneg rpos p6 rnegp6 rposp6, cluster(mohallah)
```
(10)

where the omitted category are individuals offered a zero interest rate with payment in week 1.

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

We will answer this question in three ways. First, we will estimate whether the average take-up changes over time:

$$a_{iw} = \beta_0 + \beta_w \cdot w + \varepsilon_{iw} \tag{11}$$

```
reg y w, cluster(mohallah)
```
(12)

Second, we will estimate equation 9 separately for each experiment wave, $w \in \{1, 2, 3\}$. We will use Seemingly Unrelated Estimation, and will run a joint test for parameter stability across periods.

Third, we will estimate the effect of accepting in wave $(w-1)$ on the probability of accepting in wave $w$. To do this, we will instrument acceptance in wave $(w-1)$ by the contractual terms offered in that period (where $w \in \{2, 3\}$, and $\mathbf{1}$ denotes the indicator function):

$$a_{iw} = \beta_0 + \beta_1 \cdot a_{i,w-1} + \beta_2 \cdot \mathbf{1}(w = 3) + \varepsilon_{iw} \tag{13}$$

$$\begin{aligned} a_{i,w-1} = {} & \gamma_0 + \gamma_1 \cdot rneg_{i,w-1} + \gamma_2 \cdot rpos_{i,w-1} + \gamma_3 \cdot p6_{i,w-1} \\ & + \gamma_4 \cdot rneg_{i,w-1} \times p6_{i,w-1} + \gamma_5 \cdot rpos_{i,w-1} \times p6_{i,w-1} + \gamma_6 \cdot \mathbf{1}(w = 3) + \mu_{i,w-1} \end{aligned} \tag{14}$$

```
ivreg2 a (L.a = L.rneg L.rpos L.p6 L.rnegp6 L.rpos6) ///
  DummyWave3, cluster(mohallah)
```
(15)

# 5 Analysis: Effects of treatment on firm outcomes

## 5.1 Identification strategy

We plan to address three primary research questions on the effects of the product:

**Research Question 5 (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 $\phi_s$ as a common parameter for strata $s$. Then, for each variable in the previous table, we will estimate the following ANCOVA specification with strata dummies (denoted by the dummy variables `strataD*`):

$$y_{i1} = \beta_0 + \beta_1 \cdot m_i + \beta_2 \cdot y_{i0} + \phi_s + \varepsilon_i \tag{16}$$

```
ivreg2 y treated y_pre strataD*, partial(strataD*) cluster(mohallah)
```
(17)

**Research Question 6 (LATE)** *What is the impact of accepting the product?*

To estimate the impact of accepting the product, we will instrument adoption by treatment. Denote `EverAdopted` as a dummy for whether individual $i$ ever accepted a product. Denote `AdoptionCount` as the number of times (0, 1, 2 or 3) that individual $i$ accepted. Then we will estimate:

$$y_{i1} = \beta_0 + \beta_1 \cdot \texttt{EverAdopted}_i + \beta_2 \cdot y_{i0} + \phi_{1s} + \varepsilon_i \tag{18}$$

$$\texttt{EverAdopted}_i = \gamma_0 + \gamma_1 \cdot m_i + \gamma_2 \cdot y_{i0} + \phi_{2s} + \mu_i \tag{19}$$

```
ivreg2 y (EverAdopted = treated) y_pre strataD*,
 partial(strataD*) cluster(mohallah)
```
(20)

and

$$y_{i1} = \beta_0 + \beta_1 \cdot \texttt{AdoptionCount}_i + \beta_2 \cdot y_{i0} + \phi_{1s} + \varepsilon_i \tag{21}$$

$$\texttt{AdoptionCount}_i = \gamma_0 + \gamma_1 \cdot m_i + \gamma_2 \cdot y_{i0} + \phi_{2s} + \mu_i \tag{22}$$

```
ivreg2 y (AdoptionCount = treated) y_pre strataD*,
 partial(strataD*) cluster(mohallah)
```
(23)

## 5.2 Heterogeneous effects

We plan to run several heterogeneity tests. In each case, we will test heterogeneity in (i) adoption and (ii) the impact of ever accepting the product. Denote a covariate as a dummy variable $x_i$, measured at baseline. To estimate heterogeneity in take-up, we will run the following regression, clustering by mohallah:

$$\begin{aligned} a_{iw} \quad &= \beta_{00} \cdot \mathbf{1}(x_i = 0) + \beta_{01} \cdot rneg_{iw} \times \mathbf{1}(x_i = 0) + \beta_{02} \cdot rpos_{iw} \times \mathbf{1}(x_i = 0) + \beta_{03} \cdot p6_{iw} \times \mathbf{1}(x_i = 0) \\ &+ \beta_{04} \cdot rneg_{iw} \times p6_{iw} \times \mathbf{1}(x_i = 0) + \beta_{05} \cdot rpos_{iw} \times p6_{iw} \times \mathbf{1}(x_i = 0) \\ &+ \beta_{10} \cdot \mathbf{1}(x_i = 1) + \beta_{11} \times rneg_{iw} \times \mathbf{1}(x_i = 1) + \beta_{12} \cdot rpos_{iw} \times \mathbf{1}(x_i = 1) + \beta_{13} \cdot p6_{iw} \times \mathbf{1}(x_i = 1) \\ &+ \beta_{14} \cdot rneg_{iw} \times p6_{iw} \times \mathbf{1}(x_i = 1) + \beta_{15} \cdot rpos_{iw} \times p6_{iw} \times \mathbf{1}(x_i = 1) + \varepsilon_{iw} \end{aligned} \tag{24}$$

```
reg a x0 rneg_x0 rpos_x0 p6_x0 rnegp6_x0 rposp6_x0 ///
x1 rneg_x1 rpos_x1 p6_x1 rnegp6_x1 rposp6_x1, cluster(mohallah)
```
(25)

We will then run a joint test of whether respondents with $x_i = 0$ respond in the same way as respondents with $x_i = 1$:

$$H_0 : \beta_{00} = \beta_{10}; \beta_{01} = \beta_{11}; \beta_{02} = \beta_{12}; \beta_{03} = \beta_{13}; \beta_{04} = \beta_{14}; \beta_{05} = \beta_{15}. \tag{26}$$

To estimate heterogeneity in effects, we will estimate the following regression (again, clustering by mohallah):

$$y_{i1} = \beta_{00} \cdot \mathbf{1}(x_i = 0) + \beta_{10} \cdot \texttt{EverAdopted}_i \times \mathbf{1}(x_i = 0) + \beta_{20} \cdot y_{i0} \times \mathbf{1}(x_i = 0) + \beta_{01} \cdot \mathbf{1}(x_i = 1) + \beta_{11} \cdot \texttt{EverAdopted}_i \times \mathbf{1}(x_i = 1) + \beta_{21} \cdot y_{i0} \times \mathbf{1}(x_i = 1) + \phi_{1s} + \varepsilon_i \quad (27)$$

$$\texttt{EverAdopted}_i \times \mathbf{1}(x_i = 0) = \gamma_{00} \cdot \mathbf{1}(x_i = 0) + \gamma_{10} \cdot m_i \times \mathbf{1}(x_i = 0) + \gamma_{20} \cdot y_{i0} \times \mathbf{1}(x_i = 0) + \gamma_{01} \cdot \mathbf{1}(x_i = 1) + \gamma_{11} \cdot m_i \times \mathbf{1}(x_i = 1) + \gamma_{21} \cdot y_{i0} \times \mathbf{1}(x_i = 1) + \phi_{2s} + \mu_{1i} \quad (28)$$

$$\texttt{EverAdopted}_i \times \mathbf{1}(x_i = 1) = \delta_{00} \cdot \mathbf{1}(x_i = 0) + \delta_{10} \cdot m_i \times \mathbf{1}(x_i = 0) + \delta_{20} \cdot y_{i0} \times \mathbf{1}(x_i = 0) + \delta_{01} \cdot \mathbf{1}(x_i = 1) + \delta_{11} \cdot m_i \times \mathbf{1}(x_i = 1) + \delta_{21} \cdot y_{i0} \times \mathbf{1}(x_i = 1) + \phi_{3s} + \mu_{2i} \quad (29)$$

```
ivreg2 y (EverAdopted_x0 EverAdopted_x1 = treated_x0 treated_x1) ///   (30)
y_pre_x0 y_pre_x1 strataD*, partial(strataD*) cluster(mohallah)       (31)
```

We will test for a common LATE between subgroups:

$$H_0 : \beta_{10} = \beta_{11}, \quad (32)$$

We will interact with the following baseline covariates (where, in the case of covariates that are not already dummy variables, we will run a 'median split', denoting $x_i = 1$ if a respondent is at or above the baseline sample median):

(i). Whether individual $i$ ran a business ($business\_1$);

(ii). Whether individual $i$'s spouse is supportive ($business\_27$);

(iii). A combined index of the the variables used to create $empowerment\_1$, $empowerment\_2$ and $empowerment\_3$;

(iv). Whether individual $i$ faces pressure to share cash on hand ($doing\_1$);

(v). Whether individual $i$ finds it hard to save ($doing\_2$);

(vi). Time preference ($time\_preference$);

(vii). Risk preference ($risk$);

(viii). Literacy ($literate$).

## 5.3 Multiple testing

For each treatment effect tested in section 5.1 and 5.2, we will report two statistics for inference:

(i). The standard $p$-value (calculated using clustering at the level of the mohallah, as discussed); and

(ii). The False Discovery Rate $q$-value (calculated using the method described in Anderson (2008)).

For the False Discovery Rate, we will analyse outcomes in families. To do this, we will use the families delineated earlier in the table of outcome variables ('Outcome Family 1', 'Outcome Family 2', and so on). In each case, we will apply the False Discovery Rate correction separately for (a) the basic estimations in section 5.1 and (b) all of the heterogeneity estimations in section 5.2. (For example, we will define one family as 'Outcome Family 1: Basic Estimations' and another family as 'Outcome Family 1: All Heterogeneity Estimations'.)

# 6 Robustness checks, attrition and breach of protocol

**Robustness:** As a robustness check, we will rerun our estimations including any covariate that was found to be 'unbalanced across treatments' (in the way outlined earlier). We will do this in two distinct ways:

(i). If a variable is unbalanced on contractual offers (equation 3), we will include it as a control in both (i) the analysis on 'determinants of take-up' (ii) the analysis of 'effects of treatment on firm outcomes'. This is because balance between different contractual offers is potentially relevant both to take-up and to interpreting effects of treatment.

(ii). If a variable is unbalanced on treatment against control (equation 1), we will include it as a control only in the estimation of 'effects of treatment on firm outcomes'. (This is because balance between treatment and control is not directly relevant to the determinants of take-up.)

**Attrition:** We do not expect high levels of attrition since we used monetary incentives for participation (namely, a participation fee of 1000 Pakistani rupees),[2] and expect fixed costs to the participant will be low. We will report the attrition rates and will summarise responses from participants who chose to exit the trial. We will check for the consistency of our results by running the above regressions separately for each round, with only participants of that round as a sample. We will then test for the stability of treatment effects using Seemingly Unrelated Estimations. This will allow us to see if the attrition is selective and correlative with the willingness to participate in the products offered.

**Breach of protocol:** We are not aware of any cases of breach of experimental protocol. If we become aware of any serious breaches of protocol, we will omit the relevant observations from our analysis.

# References

AFZAL, U., G. D'ADDA, M. FAFCHAMPS, S. QUINN, AND F. SAID (2014): "Two Sides of the Same Rupee? Comparing Demand for Microcredit and Microsaving in a Framed Field Experiment in Rural Pakistan," *CSAE Working Paper WPS/2014-32*.

ANDERSON, M. L. (2008): "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American statistical Association*, 103(484).

[2] This is divided as 500 rupees for completing the endline survey and 500 rupees for participating in all three waves of the experiment.