Solar Electricity in Rural Sindh: Role of Flexibility and Planning in Repayment Discipline
Pre-Analysis Plan

Jacopo Bonan[*], Giovanna d'Adda[†], Mahreen Mahmud[‡], Farah Said[§], Massimo Tavoni[¶]

Fieldwork locations:	Pakistan
Fieldwork dates:	March 2017 to December 2017
Date of Pre-Analysis Plan:	October 18, 2017

1 Introduction

This study specifically involves a set of products that are relevant to the needs of the poor in rural-off grid areas. These solar products are comparable in cost to what local villagers currently spend on their lighting needs and allow access to lighting, fans, chargers and TV. The system is suitable to serve off-grid areas, as well as on-grid areas as a backup option. This document outlines our experiment and our plan of analyzing the data.

We collaborate with EcoEnergy (EE), a for-profit company supplying sustainable and efficient solar energy solutions (e.g. lights, fans, mobile chargers, TV) to small businesses and households in rural Pakistan, to evaluate an innovative market solution. The product relieves credit constraints to adoption and has strong enforcement features: customers access energy through a pay-as-you-go monthly scheme and are disconnected when their credit expires. Since there is at the moment no financial penalty for late payments, these represent a pure loss for EE. Therefore, timely and quality repayments are crucial for business sustainability. The solar system can be customized to match consumer needs and allows the monthly fee, which can range from USD 8 to 30, to match what they may be currently spending on energy alternatives.

[*]FEEM and CMCC
[†]Politecnico di Milano
[‡]University of Oxford
[§]Lahore School of Economics
[¶]Politecnico di Milano and FEEM

The experiment will run in Southern Sindh, Pakistan from March to December, 2017. At the time of writing this pre-analysis, we had access to administrative client data and baseline interviews of 350 clients. We have run an initial check for missing values and data entry errors but have not run substantive analysis on this data.

The research provides contributions both research and policy-wise. First, we investigate key determinants of the sustainability of the business model and of product take-up, by looking at the trade-off between discipline and flexibility in repayment schedule, which is a debated issue in the microfinance literature (see Labie et al., 2016 for a review). Second, we push the frontier of the behavioural and microfinance literatures, looking at individual constraints to repayment. The experimental design tests implementation intentions, that have been found to be effective soft ways to increase the salience of actions towards goals in other settings (Milkman et al., 2011, 2013; Nickerson and Rogers, 2010). Our study is the first to test this behavioural tool in the setting of product repayment, and could contribute not only to the issue of ensuring financial sustainability of off-grid solar solutions, but also to the wider debate on flexibility in microfinance, by providing cheap and scalable solutions to ensure quality repayments.

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

2 Experimental design

EE conducts product demonstrations at the village or bazar level. Interested individuals and businesses are met individually and applicants that fulfil the eligibility criteria are then offered the product.

The experimental design varies the terms of the product offered to treatment group clients along two dimensions: the flexibility of the repayment schedule, and the presence of tools to reduce inattention and commitment problems in repayment.

- **Flexibility of payment.** In order to compare the effect of flexibility in the monthly instalment payment, we compare two types of contracts:
  - A fixed contract, close to EE's existing one under which clients will be required to make their entire payments on a monthly basis.
  - A flexible contract, under which clients can decide when and at what frequency within each month they want to pay the instalment. Clients in this group will essentially be free to plan how they want to make payments within each month: at take-up, we will inform them of the daily rate and give examples of payments at different frequencies (e.g., weekly, bi-weekly, monthly, bi-monthly).

- **Implementation Intention Plan (IIP).** The specific screening protocol used by EE makes credit constraints an unlikely explanation for late or non-repayment. We thus focus on inattention and lack of salience as main factors behind default in this setting. To test their role, and assess the effectiveness of cheap and scalable solutions to contrast them, the proposed design randomises the offer of a planning tool to clients based on implementation plans. Drawing from literature in psychology on the use of implementation plans (Gollwitzer and Sheeran, 2006), we ask customers to formulate a plan for his next payments and circle the payment dates on a calendar, delivered by the enumerator, which can then be displayed at his work place or house. This process should help the subject anticipate possible issues in repayment and the strategies to overcome them.

The level of randomization for both treatments is at the individual level, which results in a 2x2 factorial design. A random generator number leading to either one or the other contract version has been incorporated in the software used by the salespersons to register new customers. The implementation intention plan intervention is delivered by the enumerator at the time of survey administration, some days after the contract is signed. The treatment is randomised by the research team via the survey software.
The study sample is expected to be formed by about 650 individuals who signed a contract with EE and installed a solar system. Customers are categorized as small business owners or households, depending on the place where the system is installed. All surveys are administered via tablets using SurveyCTO.

3 Research questions and identification strategy

The project aims to provide empirical evidence aimed at testing the following research questions:

- **RQ1:** *What is the average effect of flexibility on repayment performance?*

  There is not a clear prediction on the expected effect of a flexible contract on the quality of payment. We will estimate:

  $$y_i = \alpha + \beta_1 Flex_i + X\gamma + \varepsilon_i \tag{1}$$

  where β_1 is the effect of being assigned to the flexible contract with respect to a fixed one. One has to notice that the decision over the actual frequency of payment represents an endogenous decision, therefore Îš1 should be considered as ITT, i.e. the effect of the possibility to choose the schedule of payments, not of a particular frequency per se.

- **RQ2:** *What are the sources of heterogeneity of the effects of flexibility?*

  Allowing for flexibility of payments reduces poor quality payments for individuals for whom consumption smoothing is harder (seasonal income, irregular income sources, low assets); and for so-

phisticates; but worsens repayment outcomes for individuals who face self-control or commitment issues. We will investigate these predictions by estimating:

$$y_i = \alpha + \beta_1 Flex_i + \beta_2 Flex_i \times H_i + X\gamma + \varepsilon_i \tag{2}$$

where the vector H contains variables which are proxy for the ability to smooth consumption, for sophistication and for mental constraints. They are further detailed in the next session. We expect β_2 to be negative when interacted with the ability to smooth consumption and mental constraints; and positive when interacted with sophistication. We expect that effects may vary along the distance from the bazar where payments can be made (as a proxy for transaction costs), since the literature shows how even small practical barriers can have large effects on behavior when commitment problems are at work. We explore this heterogeneity by running 2 on the subsample of people living above and below the median distance from the bazar.

- **RQ3:** *What is the average effect of the IIP intervention on repayment performance?*

  On average the IIP intervention, if sufficiently effective and relevant for the target population, is expected to improve the quality of repayments. We will estimate:

$$y_i = \alpha + \beta_1 IIP_i + X\gamma + \varepsilon_i \tag{3}$$

- **RQ4:** *What are the sources of heterogeneity of the effects of IIP?*

  The effect of IIP is stronger for subjects who face higher mental constraints (self-control, cognitive capacity, self efficacy); and lower for sophisticates. This will be estimated with the following model:

$$y_i = \alpha + \beta_1 IIP_i + \beta_2 IIP_i \times H_i + X\gamma + \varepsilon_i \tag{4}$$

  The effect of IIP is higher for individuals who are assigned to the flexibility contract, because the latter represent a simple way to operationalize oneâĂŹs intentions to pay in a way closer to their needs. We then estimate:

$$y_i = \alpha + \beta_1 Flex_i + \beta_2 IIP_i + \beta_3 IIP_i \times Flex_i + X\gamma + \varepsilon_i \tag{5}$$

  and we expect β_3 to be positive and significant.

- **RQ5:** *Are IIPs effective in mitigating the negative effects of the flexible payment schedule on people with higher mental constraints?*

We estimate:

$$y_i = \alpha + (\beta_1 Flex_i + \beta_2 IIP_i + \beta_3 IIP_i \times Flex_i) \times (1 + \mu H_i) + X\gamma + \varepsilon_i \tag{6}$$

in RQ2 we hypothesized that flexible contracts would yield negative effects on individuals with higher mental constraints. By introducing IIP, we argue that such negative effect could be mitigated. We therefore expect that the coefficient of the triple interaction ($\beta_3 \times \mu$ in equation 6) is non-negative.

- **RQ6:** *What are the determinants of repayment frequency? What is the effect of the actual repayment schedule on repayment performance?*

  We explore the determinants of repayment frequency by running 6 on the average number of payments in a month over the study period.

  We estimate the effect of the actual repayment schedule using Local Average Treatment Effects (LATE). We will use the contractual feature treatment as an instrument for the actual repayment frequency (equation 7) and we expect a positive relationship between the flexible treatment and higher repayment frequencies (first step in equation 8). We estimate:

$$y_i = \alpha_1 + \beta_1 Freq_i + X\gamma + \varepsilon_i \tag{7}$$

$$Freq_i = \alpha_2 + \beta_2 Flex_i + X\gamma + \epsilon_i \tag{8}$$

- **RQ7:** *Are there differential treatment effects for customers who installed the system for their household vs for their business activity?*

  The research question is addressed by running the previous analysis on the different sub-samples of customer types. However, this will be done only conditional on the presence of sufficient power in the sample of business customers, which are expected to be a smaller share of the overall sample.

- **RQ8:** *Are the repayment contractual features (fix vs flex) affecting customers' dropout?*

  The research question is addressed by estimating regression 1 with dropout over the evaluation period as a dependent variable. We anticipate two types of dropout. The first type pertains to people who change their mind before making any experience of the system. The second type pertains to people who decide to drop out after actually trying the product.

4 Data and variables

Two sources provide data for the analysis. The first is EE administrative records on customers' subscription, type of system installed, all dues, deadlines and flows of payments made by customers. This allows to timely monitor late payments, non-payments and defaults. The second source is surveys. The baseline survey is administered few days after the contract with EE is signed and is conducted by an independent NGO. Baseline data collection takes place between March and November 2017, following EE's commercial expansion in new areas.

4.1 Outcomes

The main outcome of the analysis pertain to the sphere of quality of payments to EE and defaults by customers. For each dimension, we will assess both the extensive and intensive margin. The following family of outcomes will be considered for the time windows between the installation date and the end of the monitoring period, expected by June 2018:

Table 1: Outcome variables

Variable	**Name; Family**	**Description**	**Source**	**Hypothesis**
Y_{i1}	Probability of delayed payments; extensive margin	A dummy variable for whether individual i experienced at least one delayed payment	EE admin data	RQ1-7
Y_{i2}	Probability of switch-off; extensive margi	A dummy variable for whether individual i has been switched-off because of missed payments	EE admin data	RQ1-7
Y_{i3}	Probability of default; extensive margin	A dummy variable for whether individual i is considered as defaulter. Default occurs after a long non-payment period and implies that,the system is pull back	EE admin data	RQ1-7
Y_{i4}	Share of delayed payments; intensive margin	Number of delayed payments episodes over the total number of months of the contract	EE admin data	RQ1-7
Y_{i5}	N. of switch-offs; intensive margin	Number of switching-off episodes over the monitoring period	EE admin data	RQ1-7
Y_{i6}	Share of days of delay; extensive margin	Total number of days of delay in payment over the monitoring period	EE admin data	RQ1-7
Y_{i7}	Actual frequency of payments	The average number of payments in a month over the study period	EE admin data	RQ6
Y_{i8}	Dropout before installation	A dummy variable for whether individual i drops out, before being installed the system	EE admin data	RQ8
Y_{i9}	Dropout after installation	A dummy variable for whether individual i drops out, after being installed the system	EE admin data	RQ8

4.2 Treatments

Table 2: Treatment variables

Variable	Name; Family	Description	Source	Hypothesis
$Flex_i$	Flexible repayment schedule	A dummy variable which is equal to one if the individual is assigned a flexible repayment schedule and to 0 for fixed schedule	EE admin data	RQ1-2; RQ5-8
IIP_i	Intention Implementation Plan	A dummy variable which is equal to one if the individual received the IIP treatment and 0 otherwise	Question naire	RQ3-8

4.3 Dimensions of heterogeneity

Table 3: Variables for analysis of heterogeneity (1)

Variable	Name; Family	Description	Source	Hypothesis
H_{i1}	Seasonality of income sources	Seasonality of income sources/ ability to smooth consumption is given by an index constructed through PCA, including the following variables. It is calculated on the sample of household customers. i. The share of household active members who earn on an irregular basis. ii. Dummy variables for availability of savings (both formal and informal) and access to credit (at least one,formal credit in the past) iii. Index for assets owned (using PCA)	Q2.5; Q2.6; Q8.9; Q8.15; Q2.9	RQ2
H_{i2}	Cognitive capacity; mental constraints	Index aggregating answers to three calculus questions, each coded as dummy variable for correct answer. The index is the sum of correct answers	Q7.4; Q7.5; Q7.6	RQ 2, 4, 5
H_{i3}	Cognitive capacity; mental constraints	Performance index in a memory task consisting in repeating a series of numbers which increase as the task gets more difficult. Each series is coded as a dummy equal to 1 when the individual correctly repeats the series. The index is the sum of correct answers	Q7.7a - Q7.7g	RQ 2, 4, 5
H_{i4}	Ability to pay bill on time	Index given by the mean of answers to self-reported ability to perform all the steps required to pay a bill on time.	Q7.8.1 - Q7.8.5	RQ 2, 4, 5
H_{i5}	Main constraint to paying bill on time	Variable indicating the behavior corresponding to the lowest mean of answers to the questions on self-reported ability to perform the steps required to pay a bill on time. Specifically, this is a set of dummy variables indicating whether each specific constraint to paying on time is the main one (the one with the lowest mean).	Q7.8.1 - Q7.8.5	RQ 2, 4, 5

Table 4: Variables for analysis of heterogeneity (2)

Variable	Name; Family	Description	Source	Hypothesis
H_{i6}	Resist temptation; mental constraints	Index created from the 'Implicit Theory about the Willpower to Resist Temptations scale' by Job et al. (2010). Specifically, the index is constructed by averaging the answers to the 6 questions (items 1,2 and 4 are reverse-coded).	Q7.9	RQ 2, 4, 5
H_{i7}	Self-control; mental constraints	Index created from the 10 item self-control scale from Tangney et al. (2004). Specifically, the index is constructed by adding up all the points for the checked boxes and dividing by 10. The maximum score on this scale is 5 (extremely self-controlled), and the lowest scale on this scale is 1 (not at all self-controlled).	Q7.10	RQ 2, 4, 5
H_{i8}	Locus control; mental constraints	Index given,the mean of seven items of the locus of control scale REF (ranges from 1 to 5).	Q7.11	RQ 2, 4, 5
H_{i9}	Grit; mental constraints	Index constructed from the GRIT Scale (Duckworth et al., 2007). Specifically, the index is constructed by summing 8 items, each scored on a 1 to 5 point scale (items a,b,c,d,e,f,g; b,c,d, and g are reverse-coded). Higher score means lower grit.	Q7.12	RQ 2, 4, 5
H_{i10}	Discipline with previous loans; mental constraints	Two dummy variables for whether the individual ever failed to pay back a loan or missed at least one installment due in the past.	Q8.16-Q8.17	
H_{i11}	General mental constraint index	General index of mental constraints, constructed from variables H_{i2} to H_{i10} by: 1. Reversing individual scales, so that higher values of each scale correspond to higher levels of mental constraints 2. Making an index using Anderson (2012)		RQ 2, 4, 5

Table 5: Variables for analysis of heterogeneity (3)

Variable	Name; Family	Description	Source	Hypothesis
H_{i12}	Time inconsistency/ present biased	A dummy variable equal to one when the individual switches to the (higher) future amount later in the short-term frame (tomorrow vs one month), than in the long-term frame (5 vs 6 months)	Q7.2	RQ 2, 4, 5
H_{i13}	Management of financial issues; sophistication	Number of 'All the time' or 'Often' answers to questions 7.15-7.19.	7.13-7.19	RQ 2, 4, 5
H_{i14}	Financial literacy	The sum of correct answers to 10 true-false questions on the consequences of missing installments on loan	8.18	RQ 2, 4, 5
H_{i15}	Customer type	A dummy variable for whether individual i installed the solar in his private house or in the business place	Q0.1	RQ7

5 Analysis

Before running regressions, sample balance tests are conducted on control variables and dimensions of heterogeneity considered in the analysis. First, we will describe whether the variables in the vectors of controls and dimensions of heterogeneity are balanced across the 4 cells in our factorial design by running:

$$X_{i0} = \alpha_1 + \beta_1 Flex_i + \beta_2 IIP_i + \beta_3 IIP_i \times Flex_i + \varepsilon_i \tag{9}$$

$$H_{i0} = \alpha_1 + \beta_1 Flex_i + \beta_2 IIP_i + \beta_3 IIP_i \times Flex_i + \varepsilon_i \tag{10}$$

we will report F-statistics from a joint test of the null hypothesis that $\beta_1 = \beta_2 = \beta_3$.

If, for a given variable, we do not reject H0 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.

As far as regression analysis is concerned, when outcomes are measured with binary variables, linear probability models (LPM) are calculated. Probit and logit models are estimated as a robustness check. When outcomes are measured with continuous variables, tobit models are run. Robust White standard errors are calculated.

In order to test the hypothesis related to the quality of repayment, six variables have been identified. In order to correct for multiple hypothesis testing, sharpened q-values will be calculated, based on two families of outcomes for the intensive and extensive margin, as proposed by Anderson (2012). Similarly, the analysis of heterogeneous effects will present sharpened q-values for the families of dimensions of

heterogeneity depicted in the table.

5.1 Robustness checks

(i). Due to random assignment, our estimates of treatment and heterogeneous effects are expected to be unbiased. In order to account for possible imbalance that might occur in small sample, as a robustness check, we will repeat our main estimation first using all regression included in the table of controls, then using 'post-double-selection' with LASSO (Belloni et al., 2014a,b). It is possible that some customers dropout the study sample, as they abandon the service provided by EE. In such case, they would disappear from administrative data. The determinants of dropout will be analyzed, as part of research question h. We will run robustness check by excluding dropouts from the study sample.

(ii). We will estimate equation (1) with enumerator fixed effects and examine the distribution of enumerator fixed effects to see how much the treatment effects vary with people conducting the intervention. Enumerators are not randomly assigned to respondents, so this is not an experimental comparison.

(iii). RQ8 investigates the role of treatments on the probability to dropout before and after the installation of the system. Dropout represents a form of attrition, as the customers leave the sample of analysis. This may affect the analysis of RQ1 to RQ7. We therefore estimate Lee bounds (Lee and Lee, 2009).

References

Anderson, M. L. (2012). 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.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014a). High-Dimensional Methods and Inference on Structural and Treatment Effects. Journal of Economic Perspectives, 28(2):29–50.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014b). Inference on Treatment Effects after Selection among High-Dimensional Controls. The Review of Economic Studies, 81(2):608–650.

Duckworth, A. L., Peterson, C., Matthews, M. D., and Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. Journal of Personality and Social Psychology, 92(6):1087–1101.

Gollwitzer, P. M. and Sheeran, P. (2006). Implementation Intentions and Goal Achievement: A MetaâĂŘ-analysis of Effects and Processes. Advances in Experimental Social Psychology, 38:69–119.

Job, V., Dweck, C. S., and Walton, G. M. (2010). Ego depletion–is it all in your head? implicit theories about willpower affect self-regulation. Psychological Science, 21(11):1686–1693.

Labie, M., Laureti, C., and Szafarz, A. (2016). Discipline and flexibility: a behavioural perspective on microfinance product design. Oxford Development Studies.

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

Milkman, K. L., Beshears, J., Choi, J. J., Laibson, D., and Madrian, B. C. (2011). Using implementation intentions prompts to enhance influenza vaccination rates. Proceedings of the National Academy of Sciences, 108(26):10415–10420.

Milkman, K. L., Beshears, J., Choi, J. J., Laibson, D., and Madrian, B. C. (2013). Planning prompts as a means of increasing preventive screening rates. Preventive Medicine, 56(1).

Nickerson and Rogers, T. (2010). <p>Do You Have a Voting Plan?: Implementation Intentions, Voter Turnout, and Organic Plan Making</p>. Psychological Science, 21(2):194–199.

Tangney, J. P., Baumeister, R. F., and Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. Journal of Personality, 72(2):271–324.