

NBER WORKING PAPER SERIES

IMPLICIT AND EXPLICIT COMMITMENT IN CREDIT AND SAVING CONTRACTS:
A FIELD EXPERIMENT

Uzma Afzal
Giovanna D'Adda
Marcel Fafchamps
Simon R. Quinn
Farah Said

Working Paper 25802
<http://www.nber.org/papers/w25802>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2019

We thank Josh Dean, Pascaline Dupas, Pam Jakiela, Anett John, Supreet Kaur, Gautam Rao, Chris Roth and Jack Willis for their very useful comments, as well as participants at the ASSA Annual Meeting, the BREAD Conference on Behavioral Economics and Development (Stanford), the CSAE Annual Conference, the Lahore School of Economics, the University of Milan and the Paris School of Economics. We are grateful to RCons for their help collecting the survey data and to NRSP for their invaluable assistance in running the experiment. Funding for the experiment and data collection was provided by the International Growth Centre and by the Economic and Social Research Council (UK). We thank Mahreen Mahmud and Muhammed Meki, both for excellent field assistance and for valuable comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2019 by Uzma Afzal, Giovanna D'Adda, Marcel Fafchamps, Simon R. Quinn, and Farah Said. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Implicit and Explicit Commitment in Credit and Saving Contracts: A Field Experiment
Uzma Afzal, Giovanna D'Adda, Marcel Fafchamps, Simon R. Quinn, and Farah Said
NBER Working Paper No. 25802
May 2019
JEL No. G02,O16

ABSTRACT

We conduct a field experiment to test the demand for flexibility and for soft and hard commitment among clients of a microfinance institution. We offer a commitment contract inspired by the rotating structure of a ROSCA. Additional treatments test ex ante demand for soft commitment (in the form of reminders), hard commitment (in the form of a penalty for missing an installment), and flexibility (an option to postpone an installment). Our design is unique in the literature for allowing us to test— using the same respondent population — how demand for explicit commitment features differs between loan and savings contracts. We find substantial demand for both credit and savings contracts but no demand for additional commitment features — either in isolation or in combination — in spite of their effectiveness in improving repayment. In particular, demand for savings is insensitive to explicit commitment features. Individuals offered loans actively dislike commitment and flexibility, unless the latter is combined with reminders. These findings complement a literature showing that commitment devices induce financial discipline. They show that demand for commitment depends on whether commitment features are implicit or explicit.

Uzma Afzal
University of Nottingham
Nottingham
United Kingdom
uzmaafzal1@gmail.com

Simon R. Quinn
Department of Economics
Oxford University
United Kingdom
simon.quinn@economics.ox.ac.uk

Giovanna D'Adda
Department of Economics, Management
and Quantitative Methods (DEMM)
University of Milan
Via Conservatorio 7
20122 Milan
Italy
giovanna.dadda@unimi.it

Farah Said
Lahore School of Economics
Lahore
Pakistan
farahz3@gmail.com

Marcel Fafchamps
Freeman Spogli Institute
Stanford University
616 Serra Street
Stanford, CA 94305
and NBER
fafchamp@stanford.edu

1 Commitment devices, explicit and implicit

Commitment problems — whether due to intra-personal factors such as time-inconsistent preferences, or to inter-personal ones like the inability to resist demands from others — are often cited as important barriers to saving and impediments to the repayment of loans (Casaburi and Macchiavello, 2018; Dupas and Robinson, 2013b; Duflo et al., 2011). Features of formal and informal financial products testify to the importance of commitment issues in financial decisions. The high frequency of instalments, the rigidity of the repayment schedules and an emphasis on group lending — all typical of many microfinance contracts — are believed to provide financial discipline and commitment devices to borrowers (Field and Pande, 2008; Field et al., 2013). Similar elements characterize rotating credit and savings associations, one of the oldest and most prevalent informal financial product in the developing world (Gugerty, 2007). Consistent with this, commitment devices — relying either on psychological or economic motivations — have proven effective in encouraging savings and reducing loan defaults (Ashraf et al., 2006; Brune et al., 2016; Dupas and Robinson, 2013b; Karlan et al., 2016; Stango and Zinman, 2014).

The evidence on commitment problems in financial decisions comes from two largely distinct streams of research, which often treat saving and borrowing as two separate behavioral realms, both conceptually and practically. However, when individuals struggle to hold savings over time and wish to incur lumpy expenditures, saving and borrowing may be substitutes (Afzal et al., 2018). While the idea that individuals may ‘borrow to save’ is not new (Rutherford, 2000; Morduch, 2010; Collins et al., 2009; Armendáriz and Morduch, 2010; Kast and Pomeranz, 2018), its implications for behavioral innovations in the design of financial instruments have not been explored.

In this paper, we aim to fill this gap — by presenting evidence on a large field experiment in Pakistan, conducted with clients of a prominent microfinance institution. Participants are offered financial commitment products that differ along several dimensions. Some take the form of a standard credit contract, with a lumpsum disbursed at the outset followed by a sequence of regular instalments to be repaid. Others take the form of a commitment saving contract, with a sequence of regular instalments followed by a lumpsum disbursed at the end. Both contracts offer the same commitment device — that is, a regular instalment schedule — but differ in the timing of the lumpsum disbursement. We then augment this standard product with a set of behavioral contractual variations, designed to be representative of the major tools tested in the literature on savings and borrowing: we offer soft commitment in the form of reminders to the clients or her peer; and hard commitment in the form of a penalty for missing an instalment. We introduce higher flexibility by allowing clients to defer one instalment.

This design allows us to test directly how demand for commitment varies depending on whether it is implicitly embodied in a saving or credit contract, and on whether it is increased by explicit behavioral variations — through hard commitment, through assistance in keeping track, and through exogenous variation in exposure to intrahousehold pressure. Crucially, we can test whether demand for the same commitment features, which have been added by previous studies either to saving or to credit contracts only (Bryan et al., 2010), varies when these features are part of a credit or to a saving product. Since we offer multiple cycles of the product to the same subjects, and randomly vary contract terms in each cycle, we can test how demand changes not only between subjects in the same sample, but also within subjects. These tests can help further our understanding of the behavioral

foundations of microfinance ([Bauer et al., 2012](#)), and for guiding future development of next-generation microfinance products incorporating behavioral features.

We find substantial demand for credit contracts and, to a lower extent, saving contracts. When offered both a credit and a saving contract, 46 percent of the participants who take up at least one offer actually accept both types of contracts. We also find that some subjects take up savings contracts with a negative return, while others do not take credit contracts with an interest subsidy. We interpret these results as demand for commitment: saving and credit are substitutes for some individuals, as both represent means to finance expenditure by committing to periodic instalments. However, we find no demand for the additional features of the product — like flexibility and reminders. In particular, take-up of the saving product is insensitive to the presence of any additional feature. Indeed, clients offered credit contracts appear actively to dislike the explicit commitment features, and find no value in reminders.

To characterize heterogeneity in demand for our product, we adapt the machine learning method recently proposed by [Chernozhukov et al. \(2018\)](#). We find significant demand heterogeneity across different kinds of respondents, and show that those most likely to adopt the product are, on a variety of measures, more likely to have reported difficulties in saving, more likely to face pressure to share their wealth, and more likely to believe in women’s empowerment in the household. Finally, we find no effect of our microfinance products on either business growth or household welfare — a result consistent with a growing body of literature in both the microcredit and the microsaving domains.

Together, we see our results as making three distinct contributions. First, by presenting the

same set of clients with both debt and credit products, we show that, in developing countries, many microfinance clients ‘borrow to save’. In previous pilot work, we showed this pattern of behaviour for small financial products with daily repayments (Afzal et al., 2018). This paper substantially extends that result, by showing that the same general behaviour holds for a product with larger payments, over a longer period — in particular, a product with payment sizes and repayment periods similar to those of standard financial products on the market. To the best of our knowledge, this is the first paper that randomly offers the same client pool both credit and savings products of a size comparable to standard microfinance products — and, therefore, the first paper to confirm that, for such products, many of the same clients will accept both credit and savings products. This complements recent work by Kast and Pomeranz (2018) showing that, for many microfinance clients, provision of savings accounts reduces levels of debt. It also supports the already cited literature on borrowing to save among microfinance clients (Morduch, 2010; Collins et al., 2009; Armendáriz and Morduch, 2010).¹

Second, this is the first time that demand for multiple types of commitment devices has been tested within the same field experiment. This is important, because it allows us to quantify the relative effect of different kinds of nudge, in the context of a common underlying contract and a common sample.² Further, since we cross-cut behavioral features with both savings and credit products, we are uniquely positioned to test their relevance for both

¹ The paper also relates to a recent literature on formalisation of informal savings products (Dupas and Robinson, 2013a,b). Similarly, Brune and Kerwin (2019) find a positive effect of deferred income streams designed as lumpsum payments.

² Given the constant enrichment of the behavioral policy toolbox, evaluating the relative effectiveness of multiple behavioral and economic treatments within one setting is a valuable exercise. In doing this, we join a small set of papers that test multiple treatments in the realm of employment contracts (DellaVigna and Pope, 2017), investments in energy efficiency (Allcott and Greenstone, 2017) and credit marketing (Bertrand et al., 2010).

savings and credit products. Our results on demand for commitment are consistent with the existing microfinance and microsaving literature. In particular, the decrease in take up of our product, when augmented with the explicit commitment devices, and the high sensitivity of take-up rates to contractual terms mirror the wide range of take-up figures found in the literature, especially for credit contracts (Karlan et al., 2010). It also conforms with the mixed evidence on demand for commitment (Ashraf et al., 2006; Allcott and Kessler, 2015; Damgaard and Gravert, 2018) and on its sensitivity to cost (Laibson, 2015).

The contrast between the high demand for implicit commitment (built into the repayment structure of the financial product we offer) and the low demand for explicit commitment (embodied by the additional contract features), complements existing evidence on individuals' demand for commitment devices. Such evidence is mixed, with some authors finding high demand for commitment and welfare-improving impacts of commitment contracts (Kaur et al., 2015; Schilbach, 2019; Augsburg et al., 2018), while others showing the opposite (Bai et al., 2017).³ Our results suggest that demand for commitment is not just driven by the level of commitment offered, but also by whether commitment is implicit or explicit. This distinction is consistent with the observed tendency of private institutions to shroud commitment mechanisms (Laibson, 2018) and to exploit individuals' partial naivete in the contracts they offer (DellaVigna and Malmendier, 2004). Shrouded paternalism is motivated by the tendency of naive consumers to demand sub-optimal levels of commitment (John, 2019; Bai et al., 2017), which drives firms to make commitment features implicit.

³ The evidence on the impact of flexibility on repayment quality in microfinance contracts is similarly mixed, with recent contributions showing that flexibility is valued, leads to improved business outcomes, and does not increase default (Battaglia et al., 2018; Barboni and Agarwal, 2018), and other papers providing evidence to contrary (Field and Pande, 2008; Czura, 2015).

The result — that the appeal of credit and commitment saving contracts is high, but there is no demand for flexibility or for explicit commitment devices, either soft or hard — also complements a recent literature documenting the hidden welfare cost of nudges. Recent empirical research shows that, while nudges can encourage intended behaviour, they also increase avoidance behaviour (Allcott and Kessler, 2015; Damgaard and Gravert, 2018).⁴ We speculate that many kinds of commitment devices, including the rather ostentatious features that we add in this experiment, are viewed by respondents as patronising and infantilising, rather than supportive or helpful.

Finally, we join a growing set of papers in microfinance by measuring the impact of our financial product on a wide range of household and business outcomes. Consistent with previous studies in the literature, we do not find transformative effects of microcredit on either business outcomes or household material welfare (Meager, 2018b; Angelucci et al., 2015; Attanasio et al., 2015; Augsburg et al., 2015; Banerjee et al., 2015; Crépon et al., 2015; Tarozzi et al., 2015; Karlan and Zinman, 2011; Liu and Roth, 2019). We also contribute to a smaller and more recent literature showing similarly limited effects of microsaving (Dupas et al., 2018; De Mel et al., 2018; Castellanos et al., 2019).

Our paper proceeds as follows. In Section 2, we present the experimental design and the different treatments, together with implementation details. We analyse demand for the product in Section 3, and characterize heterogeneity in that demand in Section 4. We test the effect of access to the product in Section 5, and conclude in Section 6.

⁴ In a related literature, experimental studies on ‘avoiding the ask’ and ‘moral wiggle room’ demonstrate how individuals avoid information or requests that make them feel morally obliged to act in a certain way, when such actions are costly (Andreoni et al., 2017; d’Adda et al., 2018; Dana et al., 2007). Another relevant phenomenon, ‘control aversion’, causes incentives and regulations to backfire when they are perceived as overbearing (Falk and Kosfeld, 2006; Fehr and List, 2004).

2 Experimental design

Our experiment is designed to answer two questions. First, *why do people commit to microfinance contracts with regular payment schedules?* One view — represented by standard intertemporal optimization models, without behavioral features — suggests that demand is driven by the net present value of the cash flow provided by the contract. Under such models, individuals always accept credit if the net cash return is positive, always refuse a savings contract if the net cash return is negative, and — depending on heterogeneity in individual circumstances — the same individual either demands a credit contract or a savings contract. An alternative view — representing the behavioral approach — posits both that (i) individuals struggle to hold cash balances over time (whether due, for example, to problems of self-control or problems of resisting social pressure (Karlan et al., 2019)) and (ii) that those individuals also hold a preference for lumpy expenditures.⁵ Under this alternative model, individuals may accept a savings contract having a negative net cash return, may refuse a credit contract having a positive net cash return, and a given individual may simultaneously exhibit both demand for savings and for credit (Afzal et al., 2018). To answer the first question, we offer an individual-liability microfinance product requiring regular repayments. By varying both the time of payment and size of the lumpsum, we are able to test, in several ways, for a behavioral motivation for product demand. As we report shortly, our results under that ‘basic contract’ are consistent with a behavioral motivation, rather than with a standard model.

⁵ This could arise, for example, through a preference for lumpy consumption goods (Besley et al., 1993) or for indivisible investment opportunities (Field et al., 2013). It could also arise through a ‘concentration bias’ (Kőszegi and Szeidl, 2012; Dertwinkel-Kalt et al., 2017) — though our results on use of the lumpsum, presented shortly, suggest that this mechanism, even if relevant in this context, would be operating in addition to the mechanisms of lumpy consumption and lumpy investment.

This result then begs the second question that our experiment addresses: *if microfinance demand is driven by behavioral foundations, can we increase demand by incorporating explicit behavioral features?* One might be tempted to assume that the answer to the first question implies the answer to this second question; however, as we explain in detail shortly, there are several reasons that this may not be the case. To answer this question, we augment our basic contract with combinations of four different behavioral features — two relating to repayment flexibility, and two relating to reminders.

2.1 The basic contract

The financial product offered in the experiment is inspired by the repayment structure of rotating savings and credit associations: ‘ROSCAs’. Such associations have many different names in different parts of the world; in Pakistan, they are generally known as ‘committees’. In a committee, a group of individuals come together with the goal of facilitating saving. They agree to meet at regular intervals — for example, each week — for a set number of meetings at which they each make a fixed monetary contribution, the amount of which is agreed at the beginning of the contract. At each meeting, the contributions of all members are put into a common pool, which is then allocated to a group member. Participants take turns receiving the content of the pot, until everyone has received the pot once, at which time the contract ends. The order in which members receive the pot is either determined randomly or assigned by bidding, depending on committee rules.

The contract we offer in our experiment has the same general profile of payments: fixed instalments at regular intervals over a set number of periods, plus one lumpsum payment mimicking receiving the pot. But the contract does not require the formation of a group —

and thus sidesteps the selection and enforcement issues inherent to the formation of saving committees. Instead, the contract is designed as an individual financial product offered by our partner institution, the National Rural Support Programme (NRSP), a microfinance institution with extensive experience offering credit to women across Pakistan. The timing of the lumpsum disbursement is known to participants at the time of take-up.

Subjects are offered the opportunity to take up a contract in each of three product cycles, typically with different contract terms. At the beginning of a cycle, each participant is offered an individual contract with known terms. If they accept the contract, payment starts the following Monday (Week 1). Participants pay a weekly instalment of size M in $N - 1$ of the N weeks, and receive a lump-sum payment of size L in the remaining week. A missed payment is considered a default and results in cancellation of the contract. In case of default, the participant has to return any payment owed to NRSP as soon as possible and, at the latest, by the end of Week N . If not, the participant is not offered any contract in the following cycle. In case the participant has a positive balance with NRSP at the time of default, this balance is returned to the subject at the end of the cycle — that is, after Week N . Within this basic design we experimentally vary the contracts offered along several dimensions: the number of weeks N ; the size of each instalment M ; the week in which the lumpsum payment is made (either Week 1 or Week N); and the amount of the lumpsum payment L .

Recent literature has emphasized the value of replicating similar experiments in different variations and across different contexts; this is valuable for providing a breadth of contexts, for understanding the generalizability of results, and for understanding whether results are sensitive to specific aspects of design (see, for example, [Dupas et al. \(2018\)](#) and [Banerjee](#)

et al. (2015)). With this principle in mind, we implemented our experiment in two distinct phases. These phases used different sampling frames (one focusing on microenterprises; the other focusing on households), with contractual terms adapted to the respective respondent population.

In the first phase, we restricted participation to female NRSP clients — past and current — whose household owns a business. For this group, we set $N = 6$ and $M = \text{PKR } 1000$ and we let the lumpsum payment take three possible values: $\text{PKR } 5000$, $\text{PKR } 4500$ or $\text{PKR } 5500$. Since participants pay $N - 1 = 5$ instalments of $\text{PKR } 1000$ each, a lumpsum of $\text{PKR } 5000$ simply returns the five instalments to the subject. A lumpsum of $\text{PKR } 4500$ is equivalent to deducting 10% from the lumpsum, while a lumpsum of 5500 means adding 10% to the sum of instalments received. Table 1 illustrates the payment schedule for a basic contract with a lumpsum payment on Week 1 and an interest of -10%.

< Table 1 here. >

Since there are three possible lumpsum values and two possible disbursement weeks, there are six possible contracts. Three of these contracts have a lumpsum paid in Week 1: they are a form of commitment credit contract. Three have a lumpsum paid in Week N : they are a form of commitment savings contract. Note that some credit contracts charge a negative interest: credit is subsidized. Similarly, some saving contracts yield a negative interest: subjects pay to save. This latter feature seeks to mimic the fact that savings instruments made available to the poor often yield a negative return, either because of fees and charges (e.g., Dupas and Robinson (2013a)), or because of inflation. More generally, the variation in

total remuneration allows us to understand subjects' willingness to pay for such products.

In the second phase, we drew our sample from past and current female NRSP clients, whether or not their household owned a business. Following guidance from local partners, we decided for this broader sample to use more payments, with smaller amounts: specifically, we set $N = 8$ and $M = PKR\ 500$. In these sessions, the lumpsum takes three values: $PKR\ 3500$, 3200 , or 3800 . As in the first phase, the middle value is equivalent to setting a zero interest rate, as in a standard ROSCA contract. The other two values are equivalent to adding or subtracting 8.6% to the total instalments paid by the participant.

2.2 Behavioral features: Flexibility and reminders

The experimental literature on microfinance has examined how contractual flexibility affects repayment. With some exceptions (e.g. [John \(2019\)](#)), there has been almost no research on the demand for either flexibility or hard commitment among clients of MFIs in developing countries. With the aim of filling this gap, we introduced four additional treatments to the second experimental phase — one treatment to reduce flexibility, one to increase it and two treatments to remind respondents to repay. These treatments are added to the basic contract of some subjects; other subjects receive just the basic contract described in the previous sub-section. Our main research focus is on the *demand* for reminders and for contractual rigidity/flexibility. Consequently, these treatments are always introduced to subjects before take-up. In other words, when a subject decides to accept a contract offer, she knows all the details of the contract that is being offered.

2.2.1 Reminders

Reminders are the most common form of soft commitment studied in the behavioral literature on savings. Their purpose is to help participants follow their regular schedule of payments. In our experiment, we send reminders one day before an instalment is due. Reminders are transmitted through phone messages. In the ‘respondent reminder’ treatment, the message is sent directly to the participant; in the ‘peer reminder’ treatment, the message is sent to a family member of the participant. Subjects are told that the financial product offered to them includes reminders. For instance, if a subject is assigned to a respondent reminder treatment in the first product cycle, she is told that she will receive a reminder before each instalment is due.⁶ This is different from other experiments that have externally introduced reminders and observed how these reminders affect payment patterns (see, for example, [Karlan et al. \(2016\)](#) who introduce reminders via letters and text messages). Here we investigate whether subjects are more willing to accept a financial contract that includes reminders.

The reason behind the inclusion in the experimental design of both ‘respondent reminders’ and ‘peer reminders’ is the desire to evaluate the relative role of two different sources of commitment problems discussed in the literature. Inability to meet financial obligations may derive from their lack of salience, or from individual time-inconsistency and self-control problems: in such instances, personal reminders increase commitment attain-

⁶ The experimental protocol stipulates that subjects are told: ‘To help you commit to a regular schedule of payments, we will call you on the day before an instalment is due... This call will be directed to you personally, on a phone number that you will provide to us if you take up the product.’ For peer reminders, the text is: ‘To help you commit to a regular schedule of payments, we will call a member of your family on the day before an instalment is due.’ Staff we instructed that, for reminder calls to respondents in the ‘peer reminder’ treatment, it was not permissible to leave reminder messages with any other person who might answer the call.

ment (Karlan et al., 2016). On the other hand, peer pressure and demands from household members or peers to share available resources may also limit individual ability to save or repay a loan (Ashraf, 2009; Jakiela and Ozier, 2015), and may also encourage household members to act as ‘saving monitors’, to help the respondent to stick to the savings commitment (Breza and Chandrasekhar, 2019) given their access to local community information (Hussam et al., 2017). By comparing the impact of reminders on take-up and repayment when sent to the respondent or to a peer we are able to assess the relative weight of these constraints.

2.2.2 Commitment features

Our commitment arm involves either adding a cancellation fee (we term this the ‘sunk’ treatment), or adding additional contractual flexibility (we term this the ‘flex’ treatment).

The ‘sunk’ treatment adds a cancellation fee of PKR 500 for defaulting on a contract. This penalty is added to the total amount owed by the participant to the bank. If subjects demand harder commitment contracts, we expect more take-up in this treatment. How this penalty operates depends on whether the contract is a credit contract (i.e., lumpsum paid in Week 1) or a savings contract (lumpsum paid in Week N). To recall, in case of default in a savings contract, our implementing partner NRSP returns to the subject in Week N all the instalments paid before defaulting. For instance, if a subject has paid three instalments totalling PKR 1500 then defaults, this subject receives PKR 1500 in Week N in the standard savings contract, but only PKR 1500 minus the cancellation fee, that is, PKR 1000 in the ‘sunk’ treatment. This is equivalent to making the first instalment ‘sunk’ (e.g., John (2019)). In case of default in a credit contract, the remainder of the debt becomes immediately due. For instance, if a subject had repaid PKR 1500 on a PKR 3500 loan granted in Week 1 but

stops paying in week 5, the unpaid portion of the loan becomes due in that week, i.e., PKR 2000. In the ‘sunk’ treatment, the cancellation fee of PKR 500 is added to this amount.

In the ‘flex’ treatment, in contrast, more repayment flexibility is added to the contract. In this treatment, we give the participant the flexibility of delaying *one* instalment by *one* week only.⁷ To illustrate, the subject may decide not to pay the instalment PKR 500 in Week 3. In this case, the subject will have to pay the regular instalment of PKR 500 in Week 4 plus the delayed instalment of PKR 500 from Week 3 — i.e., a total of PKR 1000 in Week 4. Other instalments remain unchanged. Note that the subject in the ‘flex’ treatment decides when to use the option to delay an instalment. It can be applied to any instalment between the first instalment and the last — or to none at all. All other rules regulating default continue to apply.

The design of the ‘sunk’ and ‘flex’ treatments draws from existing studies introducing hard commitment features to saving products (John, 2019) or adding flexibility to the rigid repayment schedule typical of microfinance products (Field et al., 2013; Czura, 2015; Barboni and Agarwal, 2018; Battaglia et al., 2018).⁸ We combine these features in a unique experimental design and test their impact on take-up and default both in the saving and credit domains.

⁷ Subjects are told that “We understand that it is not always possible to pay instalments every week. Therefore, over the course of eight weeks, we will allow you on one occasion only to delay a payment by one week.”

⁸ Our ‘flex’ treatment is closest to the flexibility option in Barboni and Agarwal (2018), with the key difference that customers have to pay in full the instalment that they decide to skip with the following instalment, rather than spreading out the outstanding balance over the remaining loan instalments.

2.3 Implementation

The first phase was conducted from 25 August 2014 to 1 March 2015 in two districts of Pakistan Punjab: Bhakkar and Chakwal. The endline survey was completed by 30 March 2015. The second phase was implemented from October 2015 to May 2016 in four districts of Punjab: Jhelum, Rawalpindi, Khushab and Mandi Bahuddin. The endline survey was carried out after Ramadan, in July-August 2016. Funding for the first phase was provided by the International Growth Centre (IGC); funding for the second phase was provided by the Economic Social Research Council (ESRC).

Participants are drawn from clients of microfinance products offered by NRSP. All subjects are women. In the first phase, participation is restricted to past and current clients whose household own at least one business. In the second phase, this restriction does not apply. The implementation of the experiment was carried out by NRSP field staff. Table 2 shows the districts, offices and sample size that was included in the two phases of the experiment.

< Table 2 here. >

In Table 3, we summarize the experimental design, and report the share of participants assigned to each treatment. In Phase 1, we used a simple treatment/control division (with 50% of our sample in each). In Phase 2, we assigned 25% of participants to the control group; the remaining 75% were then assigned in a 3×3 factorial design, covering all combinations of (i) sunk, flex or no commitment feature, and (ii) respondent reminders, peer reminders and no reminders. Subjects were told whether they are in a ‘sunk’, ‘flex’ or ‘reminder’ treatment before drawing the cards determining interest rate and lumpsum date.

Hence individuals who reject the contract before drawing cards already know whether the contract would include a ‘sunk’ or ‘flex’ component or a ‘reminder’ component.

The combination of treatment dimensions results in a total of six different generalized ROSCA contracts in the first phase (that is, the six combinations of interest rate and time of lumpsum payment), and 54 different generalized ROSCA contracts in the second phase (the six combinations from the first phase interacted with the nine treatment cells). We do not have sufficient statistical power to study the effect of each possible combination of treatments separately, but this was never the intention of the experimental design: rather, in the empirical analysis, we use the orthogonality of treatment assignment to examine different treatment dimensions separately.

< Table 3 here. >

We assigned participants to these various treatments by stratified randomization.⁹ We fielded baseline and endline surveys on both treated and control subjects. In the second phase of sessions we also conducted phone surveys to gather higher frequency information about short list of indicators, with the view of more precisely estimating any potential impact of treatment on business outcomes.

In practice, assignment to treatment was implemented by inviting participants to draw a card at random at the beginning of a cycle in order to determine the interest charge (i.e.,

⁹ We stratify as follows. First we divide participants into blocks based on their answers to questions on ‘running a business’, ‘whether the respondent makes the final decision on spending’ and ‘whether the respondent would use a loan for investment’. We then sort within blocks by household income. Individuals with a similar income in the same block are then assigned to different treatments.

zero, negative, or positive) and the week of the lumpsum payment (Week 1 or Week N). The card is drawn in Week 0, at which time subjects are asked whether they take the contract or not. If they do, NRSP field officers return a week later to start the contract proper. Inserting a week between the acceptance of the contract and any initial payment is done both for logistical reasons, and to minimize the effect of pure present bias on take-up.

Tables A1 and A2 describe main characteristics of the sample in the two phases. Monthly household consumption averages PKR 25,000 (\$250) in Phase 1 and PKR 20,000 (\$200) in Phase 2. A large proportion (60%) of the sample in Phase 1 is self-employed but this proportion is much smaller in Phase 2. On average, respondents in the two samples report finding it difficult to save. Tables A1 and A2 also report p -values for randomisation balance across treatments. This is done by regressing each variable on the assigned treatment status in a saturated specification. We also test for randomisation balance across contract terms, using a similar saturated specification that regresses each variable on randomly assigned interest rate and week of payment. We find strong balance across treatment status and contract terms in Phase 1. We find four variables to be unbalanced at the 90% confidence level in Phase 2. All our main estimation results are unaffected if we include these four variables as additional controls.

3 Demand for the product

Our empirical analysis — both of product demand and consequences of adopting — follows two Pre-Analysis Plans.¹⁰

3.1 Average take-up rates

We start by documenting average take-up frequencies for the six combinations of interest charge and lumpsum week offered in the two phases of experimental sessions. To do this, we estimate only among treated respondents (that is, we omit respondents in the control groups, because they were not offered the contracts). Take-up frequencies are obtained by estimating a linear probability model of the form:

$$a_{it} = \sum_{w=1}^2 \sum_{r=1}^3 \beta_{wr} \cdot T_{it}^w \cdot T_{it}^r + \varepsilon_{it}, \quad (1)$$

where $a_{it} = 1$ if individual i accepts the contract in cycle t and 0 otherwise. Variables T_{it}^w and T_{it}^r are dummies equal to 1 if individual i in cycle t is offered a contract with payment in week 1 or N and with a negative, zero or positive interest added to the lumpsum. In phase 2 some subjects said they were not interested in any contract and consequently refused to select a card to determine T_{it}^w and T_{it}^r . These subjects thus refused all six possible contracts, each of which they would have been offered with probability $1/6$. We treat these cases as six different refusal observations each given a weight of $1/6$. Standard errors are clustered at the individual level.

¹⁰ Our pre-analysis plan for Phase 1 (filed on 10 May 2015) is available at <https://www.socialscisceregistry.org/trials/684>, and the extensive implementation of that analysis is available at http://www.simonrquinn.com/research/Microfinance_PreAnalysis_Phase1.pdf; our pre-analysis plan for Phase 2 (filed on 15 January 2017) is available at <https://www.socialscisceregistry.org/trials/1916>, with extensive implementation available at http://www.simonrquinn.com/research/Microfinance_PreAnalysis_Phase2.pdf.

< Table 4 here. >

Results for all subjects from both experimental phases appear in the top panel of Table 4; in the bottom panel, we exclude ‘automatic refusers’ — that is, respondents who refused the contract before learning the contractual terms.¹¹

Focusing on the top panel, we first note that, in both phases, take-up is positive for all six contracts. Similar qualitative results are obtained if we only consider subjects who were offered the basic contract. Take-up responds to contract terms: demand for the product is higher when payment is in week 1 instead of in week N ; and demand is higher when the lumpsum is larger. These differences are all statistically significant at the 1% level. Comparing the impact of interest rate on take-up between credit and saving contracts, we see a larger sensitivity of demand to the size of the lumpsum for the former, consistent with the existing evidence (Karlan et al., 2010).

We observe behaviors consistent with a pure demand for saving commitment. For instance, 2.7% of participants in phase 1 and 4.1% in phase 2 take up saving commitment contracts with a fee — i.e., for which the lumpsum is less than total instalments $M \cdot (N - 1)$. In both cases, participants could in principle accumulate the instalments themselves and end up

¹¹ The proportion of automatic refusers in each cycle of each phase is reported in Table A3. Across all three cycles, automatic refusers account from one third of phase 1 observations, and two fifths of phase 2 observations. In phase 1, the proportion of automatic refusers increases slowly across cycles; in phase 2, the proportion of automatic refusers is twice as high in the first cycle than in the other two. 58% and 29% of subjects are never automatic refusers in phase 1 and 2, respectively. Automatic refusers tend not to refuse in all product cycles: the proportion of those who always refuse automatically is 25% and 20% in phase 1 and 2, respectively.

with more money. Other behaviors are consistent with a difficulty to save independently: a large proportion of subjects, 53% and 62.8% in phase 1 and 2 respectively, refrain from taking a subsidized credit contract. In both cases, subjects could have taken the loan, paid back the instalments, and be left with the subsidy. The only reason for not doing so is having to hold onto the funds to pay the instalments. All these results are consistent with earlier findings obtained by Afzal et al. (2018) using a similar contract design but a much shorter contract duration.

It is possible to construct standard models of saving and borrowing — that is, models in which respondents do not have a preference for commitment savings — in which participants accept a saving contract with a fee, or refuse subsidized credit. However, the implied discount rates for such a model are so extreme as to be implausible. To illustrate, consider the week 1 frequencies, i.e., the take-up of loans. The change in take-up rates over experimental interest rates implies that 8.2% of phase 1 subjects have annual compound discount rates higher than 128%, and 53% of them have *negative* discount rates higher than 50%, with remaining subjects lying in between these two extremes. Similar patterns are observed in phase 2.

This is not all. If we turn to take-up of savings contracts, we get completely different implied discount rates if we believe our subjects follow a standard model of borrowing and saving. For instance, the fact that 89% of phase 1 subjects turn down an opportunity to save at an annual compound interest of 128% (10% over 6 weeks) would imply that these subjects have a discount rate higher than 128% – a number that is much larger than the 8.2% of subjects who take a loan at that same interest rate. A discrepancy of the same magnitude is found in phase 2. The same contradiction arises for those subjects exhibiting negative

discount rates: 2.7% (4.1%) of phase 1 (2) subjects appear to have a negative discount rate higher than 60%, compared to 53% (62.8%) based on loans.

The results excluding automatic refusers are similar in terms of sensitivity to contract terms and in terms of discrepancies in the implied distribution of discount rates across saving and credit contracts. It is reasonable to assume that automatic refusers include all the subjects who are uninterested in our contracts because of the implied transaction costs. It follows that the discrepancies in discount rate distribution observed across the saving and credit frames are not an artefact of transaction costs.

Even more damaging for a standard model with stable time preferences is the fact that the same subjects often accept — or reject — both a saving and a credit contract with comparable terms. This is summarized in Table 5. In the first panel of the table, we consider individuals who take a loan charging a positive interest — implying impatience — as well as a savings contract with a zero or negative return — implying a strong desire to postpone consumption. After dropping automatic refusers (for whom contract terms were not drawn), we have 107 individuals in phase 1 and 350 in phase 2 who were offered each type of contract at least once during one of the three product cycles. In both cases, the majority of subjects rejected both contracts — in line with the generally low take-up of low payout contracts documented in Table 4. Of those who take at least one loan contracts with zero or positive interest, 20 to 24% also take the savings contract. Similarly, of those who take at least one saving contract with a negative return, 67 to 68% also take a loan contract with a zero or positive interest. This shows that most individuals who take up a negative return savings contract also take a credit contract — suggesting that credit, for them, serves the same lumpsum accumulation role as credit.

< Table 5 here. >

In the second panel of Table 5 , we consider individuals who accept either of the two costly contracts discussed in Panel 1, while at the same time refusing a credit contract with a negative interest — i.e., a contract paying PKR 5500 (Phase 1) or 3800 (Phase 2) in week 1. We have already argued that refusing such a contract violates the standard model — except for the existence of transaction costs. We have 101 individuals in phase 1 and 399 individuals in phase 2 who are offered both types of contracts, i.e., a negative interest credit contract and a contract (credit or saving) with a low payout. As noted earlier, the majority of individuals who are offered the negative interest credit take it. There is, however, a non-negligible fraction of the subjects who refuse this contract. Of those, all take the lower payout contract. Similarly, among those who take the low payout credit or savings contract, two thirds refuse the attractive credit contract. In spite of the different sample sizes, the proportions are almost identical in the two subject populations.

Taken as a whole, the evidence is impossible to reconcile with a standard model of borrowing and saving in which the distribution of discount rates is stable, reasonable, and identical for borrowing and lending. The fact that take-up is higher for credit contracts with a low or negative interest suggests that subjects prefer more money at an earlier date. But many are also willing to take less advantageous contracts, while others are unwilling to accept even the most generous contract. As we show in Afzal et al. (2018), this behavior is consistent with a model in which individuals have a preference for lumpsum accumulation but are unable to save without contractual commitment. When their demand for lumpsum

accumulation is high enough, they are more likely to accept a commitment contract, provided the effort to save is not too high, i.e., provided that the interest charge is low enough and the timing of lumpsum disbursement early enough. When the need for lumpsum accumulation is low, subjects refuse any contract, even those with a credit subsidy.

Further support for this conclusion lies in respondents' description of how they used the lumpsum payments in Phase 2. We show the top categories for this use in Figure A4 (appendix). Of the top eight categories (which together cover about 80% of respondents), seven categories quite obviously describe lumpy purchases, whether in the form of consumption durables (home appliances and clothing), investment (home repairs and assets for a business), wedding and festival expenses, and medical expenses. The only category of the top eight that does not necessarily fit this categorization is 'food purchases' (representing 20% of respondents); we do not know more about the specific form of the food purchases, but we note that even lumpsum food purchases often attract quantity discounts (see, for example, [Brune and Kerwin \(2019\)](#) and [Attanasio and Pastorino \(2015\)](#)).

3.2 Demand for explicit commitment features

We now turn to the reminder and commitment variations. We have seen that the behavior of many participants is consistent with a demand for lumpsum accumulation and an inability to save at home, which leads them to accept contracts that commit them to the payment of a sequence of regular instalments. Since this behavior indicates a demand for commitment contracts, we are interested in finding out whether these same participants are also interested in contracts with additional explicit commitment features, such as receiving reminders or facing a cancellation fee. The literature has shown that such features reduce

the probability of default. What is less clear is whether subjects have an *ex ante* demand for these additional commitment features. At the same time, we wonder whether subjects might benefit from the option to postpone an instalment to deal with an emergency. If this is the case, we expect a higher take-up when this option is included in the contract. Further, we are able to test whether these contractual features are valued differently when participants have to pay to save up for to a lumpsum amount in a savings product or when they have to pay down to repay a lumpsum provided under a standard credit framework.

Figure 1 presents the results from this investigation. The figure shows (on the far left) take-up rates for the basic contract (that is, the product with neither the ‘flex’/‘sunk’ variation nor the ‘self reminder’ / ‘peer reminder’ variation); it then shows take-up rates for each of the eight possible behavioral variations. Error bars show 90% and 95% confidence intervals on the difference in take-up relative to the basic contract. These take-up rates — and confidence intervals — are obtained using an OLS regression of take-up on dummies for the nine combinations of reminder and commitment treatments, as well as dummies for the six combinations of payment week and interest rate. Standard errors are clustered at the individual level and observations from all three product cycles are combined. We show the original regression coefficients, and pairwise significance tests, in Tables A4 and A5 in the appendix.¹² Panel A of Figure 1 (at the top) shows these results for credit contracts (that is, contracts where the lumpsum is offered to be paid in the first period); Panel B (at the

¹² Panels 2B and 2C of these tables show respectively the marginal effects of flexibility and of reminders. Relative to a basic contract, the addition of flexibility increases demand only when coupled with the option of reminders to the respondent (Panel 2B). On their own, reminders reduce take-up, especially when they are sent to peers. Adding reminders to the respondent has a significant effect on demand only when coupled with the added contractual flexibility (Panel 2C). This seems to suggest that reminders are more valued *ex ante* when the contract is more flexible — perhaps because subjects feel that the lower level of commitment needs to be compensated by reminders. This interpretation finds some additional support in the fact that reminders have no effect on take-up in the ‘sunk’ treatment. When reminders are sent to peers, the positive effect on take-up in the ‘flex’ treatment is smaller in magnitude and no longer significant.

bottom) shows the results for savings contracts (where the lumpsum is to be paid in the final period).

< Figure 1 here. >

The implications of the figure are stark: contrary to what one might expect, clients do not value additional commitment features. This is particularly evident for credit contracts. Here, of the eight variations on the 'basic, no reminders' product, demand is lower in seven cases; in three of these cases ('flex, no reminders', 'sunk, no reminders', and 'sunk, reminder to peers'), the demand reduction exceeds 25% (i.e. 5 percentage points), and is significant. A joint test that take-up is equal across all nine contracts is rejected with $p = 0.011$ (see Table A4). We do not see the same pattern for saving contracts, for which take-up is generally lower — however, we cannot reject the null hypothesis that the relative take-up pattern that we observe for credit is the same pattern as for saving.¹³

In our view, this is a novel result for understanding the demand for microfinance in developing countries. It shows that microfinance products with a fixed repayment schedule — an extremely common form of contract across developing countries — may represent an important form of what Laibson (2018) refers to as 'shrouded paternalism'. As Laibson explains:

...lots of thriving institutions have bundled commitment features that appear to be specifically designed to help agents overcome their self-control problems.

¹³ When we conduct a joint test here of the null hypothesis that the take-up rate is equal across all nine contracts, we find $p = 0.321$ and we do not reject; see Table A5. When we conduct a joint test across Panel A and Panel B, of the null hypothesis that the estimates in Panel B simply scale down those in Panel A by a common ratio, we also do not reject: we obtain $p = 0.206$.

On the other hand, these institutions generally don't market these commitment features — i.e., the forcing mechanisms are shrouded.

Our results in this paper — namely, our finding that respondents value the commitment implicit in microfinance, but do not value additional explicit behavioral features — are consistent with precisely this narrative. Similarly, although studies from other domains show significant demand for contracts featuring explicit commitment (Kaur et al., 2015; Bai et al., 2017), existing evidence on soft and hard commitment devices in the saving domain confirms the greater success of the former (in the form of 'labeled savings accounts') over the latter (both in terms of demand and impact on outcomes) (Karlan and Linden, 2014; Benhassine et al., 2015). The lower demand for explicit commitment features is also consistent with a related behavioral literature on 'avoiding the ask' and control aversion (Andreoni et al., 2017; Falk and Kosfeld, 2006; Fehr and List, 2004). Our results on demand for flexibility also emphasize the importance of the specific details of flexibility, in the context of the specific contract being offered. In particular, both Barboni and Agarwal (2018) and Battaglia et al. (2018) find significantly higher demand for contracts with a more flexible repayment schedule — provided in the form of the possibility to take a three-month repayment holiday and spread the outstanding balance over the remaining monthly instalments (in the former); and of the option to delay up to two monthly instalments with a corresponding increase in the duration of the loan cycle (in the latter).

4 Heterogeneity in take-up: A machine learning analysis

The previous sections have analysed the average take-up rates across our sample. In this section, we use a rich set of covariates to characterize heterogeneity in these patterns. To do so, we use a modified version of the machine learning approach recently proposed by [Chernozhukov et al. \(2018\)](#). In this context, we see this method as serving two related purposes. First, the exercise allows us to track take-up for different product types across groups with different take-up rates. This serves as a robustness check to our earlier conclusions: one might be concerned that the average patterns that we have just documented might change substantially when we focus on heterogeneous sub-groups, but we show in this section that this is not the case. Second, and more fundamentally, the method allows us to test directly for heterogeneity across covariates. In doing so, it then allows us to describe the characteristics of those women who have high demand for the product, and those who have low demand. We argued earlier that take-up for our product is driven by a ‘borrowing to save’ motivation; if this is the case, this should be reflected in the descriptive characteristics of those groups having higher product demand.

We implement the [Chernozhukov et al. \(2018\)](#) method as follows. First, we randomly split treated respondents into auxiliary and main samples. In the auxiliary sample, we use a machine learning method to estimate the probability of product adoption, conditional on a vector of 58 baseline covariates. Specifically, we use an elastic net with a logistic link function; for each random split of the data, we rescale the covariate vector, then tune and train the model using two-fold cross validation — choosing α (the mixing percentage) and λ (the regularization parameter) to minimize deviance. We then use the estimated parame-

ters from this model to predict take-up in the main sample, for post-processing.¹⁴ We focus primarily on results for Phase 2 (both because Phase 2 incorporated commitment features, and because Phase 2 collected a more extensive set of baseline covariates); we show similar results from Phase 1 in the appendix.

Figure 2 shows Group Average Treatment Effects ('GATES'), sorted by take-up propensity. That is, we group our data into quintiles of the overall take-up propensity; for each quintile, we then characterize average take-up rates and 90% confidence intervals. Consider first the black bars; these show the estimated take-up rates for all contracts pooled. These bars, which are rescaled versions of the top and bottom panels of Figure 2, are a direct analog to Figure 4 in Chernozhukov et al. (2018). They show that we have substantial heterogeneity in take-up rates across individuals with different covariates: for the lowest quintile, the average take-up rate is approximately 10%, and for the highest quintile, the rate is above 25%.¹⁵

< Figure 2 here. >

We augment this analysis in two ways. First, in the top panel of Figure 2, we add take-up rates and confidence intervals for (i) products offering the 'flex' variation, (ii) products of-

¹⁴ This follows closely the approach in Chernozhukov et al. (2018). Note that, in our context, the outcome of interest is the take-up rate — which, for members of the control group, is zero by construction. Therefore, using the terminology of Chernozhukov et al. (2018), we are estimating $s_0(z)$, and imposing $b_0(z) \equiv 0$ by construction. We construct both point estimates and confidence intervals using the 'variational estimation and inference' method described in Chernozhukov et al. (2018) (for which we use 1000 random sample splits).

¹⁵ Chernozhukov et al. (2018) provide a method for testing whether this heterogeneity is significant, by testing whether the 'best linear predictor' of take-up varies with respect to predicted take-up. We find that it does: using the terminology of Chernozhukov et al. (2018), we estimate $\hat{\beta}_2 = 0.983$, with a 90% confidence interval of (0.691, 1.275) (where $\beta_2 = 0$ represents the null hypothesis of no heterogeneity across covariates).

fering the ‘sunk’ variation, (iii) products offering the ‘respondent reminder’ variation and (iv) products offering the ‘peer reminders’ variation.¹⁶ The patterns are remarkable for their stability across quintiles. In short, ‘a rising tide lifts all boats’: the covariate factors that correlated to an overall increase in take-up rates also correlate with increased demand for each of the various behavioral variations. Second, in the bottom panel of Figure 2, we repeat the analysis for the contract terms: that is, for variations in the lumpsum amount and in the time of payment. Again, the basic pattern — and all of the stylised facts noted in Table 4 — holds across all quintiles.

Who, then, are the respondents who fall into these quintiles? Following Chernozhukov et al. (2018), we answer this question by describing the characteristics of those respondents in the ‘most affected’ and ‘least affected’ groups — that is, the 20% with the highest adoption rate (which we term the ‘highest adopters’) and the 20% with the lowest adoption rate (the ‘lowest adopters’). In Table A7 (appendix), we perform this comparison for all 58 of the baseline covariates used for our analysis. In Table 6, we focus on those covariates with a specific behavioral interpretation — namely, variables relating to respondents’ baseline saving difficulties, respondents’ attitudes about women’s empowerment, and respondents’ ability to keep track of tasks and finances.

< Table 6 here. >

Table 6 shows large and highly significant differences in respondent characteristics for al-

¹⁶ To be clear: for each of these variations, we graph against the same quintiles calculated earlier — that is, quintiles in overall take-up rates, rather than quintiles calculated separately for each variation. This is important for comparability across graphs, and comparability to the cluster analysis that follows shortly.

most all of the ‘behavioral’ characteristics in Phase 2. It is particularly noteworthy that, of the highest adopters, 89% said at baseline that they find it hard to save, and 94% said they face pressure to share; the equivalent figures for the lowest adopters are just 54% and 55% respectively. Further — and consistent with our interpretation that the basic contract provides a useful commitment device — the highest adopters are significantly less likely to have described themselves at baseline as ‘good at keeping track of time’, ‘good at keeping track of finances’, to follow a strict schedule on finances, to follow a tight routine, and less likely to act early to avoid forgetting (either generally or with respect to finances). Finally, as one might expect, the highest adopters report significantly higher intra-household empowerment at baseline: they report a significantly higher share of household decisions in which the woman’s view is always considered, and are more than twice as likely to agree that it is appropriate for a woman to invest in her business without consulting her husband and to go shopping for a personal item (specifically, a scarf).

The Phase 1 counterpart to Table A7 appears in the appendix as Table A8. Here we see both important similarities and important differences to the patterns in Phase 2. Both similarities and differences can be explained by the different sampling strategies used — in particular, by the fact that Phase 1 deliberately includes many more self-employed respondents. For example, we find in both phases that the ‘highest adopters’ are more likely to be self-employed than the ‘lowest adopters’ (though the highest takers in phase 2 have essentially the same self-employment frequency (22%) as the lowest takers in phase 1 (19%)). Similarly, we find in both cases that the highest takers are have larger households (with higher household consumption), and are more likely to be a member of a savings committee. In contrast, we find no significant difference in Phase 1 between the highest adopters and the lowest adopters in terms of pressure to share — and, in Phase 1, the highest adopters are

significantly *less* likely to have declared, at baseline, that they find it hard to save. One possible interpretation is that the Phase 2 and 1 samples form a continuum, with the highest take-up respondents from phase 2 sharing many similarities with low take-up respondents from phase 1 — notably in self-employment, consumption expenditures, household size, membership in a savings committee, and pressure to share. Large households with more self-employment and a higher income are presumably more able to save — and thus to join a savings committee — while their daily income from self-employment exposes them more to the pressure to share. This interpretation would explain why, across the two samples, take-up increases with self-employment, income, family size, ability to save, and pressure to share.

Finally, we show the Phase 1 counterpart to Figure 2 in the appendix, as Figure A6. The general patterns are the same, though Phase 1 respondents appear to have a greater sensitivity to the size of lumpsum payment.

5 Consequences of adopting

5.1 Contract features and repayment

How do contract features affect repayment? This question is important for shedding light on our earlier take-up analysis. If, for example, late repayment problems are widespread, this would have implications for the practical viability of the products studied; similarly, if late repayment rates do not differ between the basic contract and the ‘sunk’ variation, this might suggest that respondents are naive about the value of the commitment device for their future behavior (DellaVigna and Malmendier, 2006; John, 2019).

In Figure 3, we show the rate of late repayment by behavioral variations. The structure of the figure mirrors that of the earlier figure showing take-up rates (Figure 1): we show the rate of late repayment both for the basic contract and for the behavioral variations, and we divide the analysis between credit contracts (Panel A, top) and saving contracts (Panel B, below). The figure shows late repayment that is not authorised by the contract — so, for example, a respondent under the ‘flex’ contract who is exercising her right to delay one payment by one week is not considered here to be late. Figure A7 (appendix) repeats the analysis, but recording those women as making a late payment.

< Figure 3 here. >

Several stylized facts deserve noting here. First, on average, the probability that at least one of a client’s payments is made late is about 12%. Second, this rate generally decreases with the various behavioral features — and is significantly lower in several of those cases.¹⁷ That is, on their own, reminders (whether reminders to the respondent or to a peer) reduce the rate of repayment problems, but not significantly; the larger decreases — which drive significant differences — operate through the contractual flexibility variations. That is, we do *not* find evidence, in this context, that clients are naively opting in to the additional penalty scheme (‘sunk’) without further consequence on their behavior.

¹⁷ Under credit, it is lower under the combination of ‘flex’ and the respondent reminders, under the combination of ‘sunk’ and the respondent reminders, under the combination of ‘flex’ and the peer reminders, and under the combination of ‘sunk’ and the peer reminders. Under savings, it is lower under ‘sunk’ (with no reminders), the combination of ‘sunk’ and respondent reminders, the combination of ‘flex’ and peer reminders, and the combination of ‘sunk’ and peer reminders.

Third, we note that — with the important exception of the ‘sunk’ contracts — the rate of late repayment is higher under saving contracts than under credit contracts. This makes intuitive sense, for two related reasons. On the one hand, subjects who renege on a commitment saving contract only face mild penalties: their paid-in instalments are kept until the end of the product cycle, at which point they are returned. Therefore — for clients not facing the ‘sunk’ contract — subjects essentially have the option to walk away from the contract. On the other hand, default in credit contracts is much lower because NRSP collection effort are much stronger. The logic is simple: the subject has already received the lumpsum, so renegeing is individually optimal for the borrower and thus has to be disincentivized by a concerted debt recovery effort. While these findings are not particularly surprising, they nonetheless bring to light the inherent difficulty of getting a third party to enforce a commitment savings contract, as opposed to a credit contract. This simple dichotomy may go a long way in explaining the predominance of credit contracts in microfinance, in spite of the fact that an important purpose of microfinance is to enable households to save.

5.2 Business and household outcomes

We now turn to the estimation of the impact of treatment on business and household outcomes. The main focus of this analysis is the comparison of control participants (who were not invited to take-up any of our commitment contracts) with treated participants (who were). We use the following ANCOVA specification:

$$y_{i1} = \beta_0 + \beta_1 \cdot T_i + \beta_2 \cdot y_{i0} + \phi_s + \eta_d + \varepsilon_i, \quad (2)$$

where y_{i1} denotes an outcome variable of interest measured at endline 1, y_{i0} is the baseline value of y_{i1} , ϕ_s are strata dummies, and η_d are district fixed effects. We cluster errors at the

household level.

The variable T_i takes two interpretations, depending on the specification. First, we denote T_i as assignment to treatment; in that case, we estimate equation 2 using OLS, and interpret $\hat{\beta}_1$ as the ITT. Second, we denote T_i as take-up. This takes four possible values, which depend on whether the subjects takes up the contract in 0, 1, 2 or 3 cycles. In this case, we calculate average take-up at the individual level, and instrument this using assignment to treatment and to contractual terms; we then interpret $\hat{\beta}_1$ as providing the LATE, normalized for a case where a respondent takes up in all three product cycles.¹⁸

Outcomes are divided into two broad categories: business outcomes and household finance and consumption. In Table 7, we report business outcomes; we collect ITT and LATE estimates for both phase 1 and phase 2 samples. In principle, our MFI partner lends for business purposes: it is therefore of primary interest whether our commitment saving contract is able to improve business investment and performance.

< Table 7 here. >

¹⁸ To instrument average take-up, we proceed as follows. First, for each cycle s , we estimate the predicted take-up of individual i based on the different types of treatments i was exposed to in that cycle – i.e., payment week, negative or positive interest, reminders, and ‘flex’ or ‘sunk’ treatment. This is achieved using the same regression that was used in generating Table 4 for the six combinations of payment week and interest rate — except that it is estimated separately for each cycle. This generates a predicted take-up for each product cycle. The sum over all three cycles is then used as instrument for T_i when estimating. For automatic refusers, we do not have a specific payment week or interest rate on which to base our prediction – since these subjects refused the contract before cards were drawn. To circumvent this issue, we ascribe to each of these observations the average predicted take-up associated with their commitment and reminder treatment, assuming an average interest rate and payment week. In practice this is achieved, as before, by generating six observations for each refuser, one for each combination of payment week and interest rate, and ascribing a weight of 1/6 to each of these observation when estimating the predicting equation.

We find almost no significant effect on business and household outcomes of having been offered our treatment; this is consistent with a growing body of evidence on the effects of microfinance (see, for example, [Meager \(2018b\)](#) and [Meager \(2018a\)](#)). In the phase 1 sample, 60% of respondents have a business. Among these subjects, we find generally positive point estimates on business performance, as measured by investment, sales, or profit. But these point estimates are in general not statistically significant. Two of the ITT coefficients are above the 10% significance level, but only one of the LATE coefficients is significant, and it is for another dependent variable. In contrast, among the phase 2 sample, estimated treatment effects are small in magnitude and never significant. This may be because a much smaller proportion (12.5%) of these households have a business at baseline.

Results for household material outcomes are presented in Table 8. We find no significant effect on household consumption or household income (the latter being measured only in the phase 2 sample). In the phase 1 sample, we find a large and significant LATE coefficient on total household assets and total individual assets. This encouraging result is, however, negated in the phase 2 sample where we find a large but negative LATE effect on total household assets.

< Table 8 here. >

The bottom part of Table 8 relates to household finances. We see that 75% of control subjects in the phase 1 sample save in a ‘committee’. The proportion is smaller in phase 2: 16.6%. We find a positive and significant LATE effect on participation in a committee, but given that the corresponding ITT coefficient is essentially 0, it is unclear how much faith to put

in this result. We also find a positive LATE for participation in a committee among phase 2 respondents, but the effect is not statistically significant. The last row of Table 8 reports results for the total debt of the respondent. Our commitment saving product should have helped participants reduce their stock of debt. We find little evidence of this. Among phase 1 subjects, ITT and LATE coefficients are positive but not significant, while among phase 2 subjects the ITT is negative and significant but the LATE coefficient is not.

Further, we measure the impact on a short list of indicators using higher frequency information from phone surveys conducted at the end of each experiment wave. Table A9 (appendix) summarizes the results for business and household outcomes. We find generally insignificant effects. There are no significant effect on the likelihood of running a business, the number of businesses or on the value of capital invested in the business in the last one month. Treated participants have higher consumption and lower debt but this difference is never significant.

Finally, we check for heterogeneity in these effects, by the quintiles of take-up rates estimated earlier. Specifically, we estimate equation 2 separately for each of those quintiles, for all of the outcomes in Table 8 and Table 8. We use the bootstrap method of Chernozhukov et al. (2018), both for obtaining point estimates and for inference. We do not find heterogeneous effects (it is not the case, for example, that some quintiles are benefiting from being offered the treatment while others are not). In Figures A8 and A9 (appendix), we show this for two outcomes of particular interest: business investment and household consumption (for Phase 1 and for Phase 2 respectively).

6 Conclusions

Recent years have produced a wealth of exciting research on commitment problems, including empirical work on the demand for commitment devices in developing countries. In such work, it is often assumed that, if people are aware of their commitment problems, they will welcome the opportunity to take commitment devices. But this need not be the case: depending upon how they are designed, commitment devices can either serve to be supportive and helpful, or can serve instead to patronise and infantilise. Many of us, for example, welcome the implicit commitment in a Pay-As-You-Earn taxation system, or appreciate that consumption taxes are deducted at the point of sale rather than at the end of the year. Yet most of us would likely be appalled if the government were to send monthly reminders to pay our taxes, or to offer to institute harsher penalties as a way to assist us to pay our taxes on time. This basic fact has been long understood in the design of many commitment devices. For instance, groups like ‘Alcoholics Anonymous’ or ‘Self-Management and Recovery Training’ seek to provide commitment while at the same time maintaining strong philosophies of respect and mutual support.

For this reason, the optimal design of commitment features remains an open question for empirical research. This paper makes progress on that issue by testing the role of commitment devices in microfinance in two distinct ways. We have done this in two distinct ways. First, we test directly whether the rotating structure of a committee can be implemented as an individual commitment-saving product. In previous pilot work, we established this fact for small product sizes with daily repayments (Afzal et al., 2018); in this paper, we show that the same structure can be used for a product with larger payments, over a longer period. We find substantial demand for such a product. Many microfinance clients ‘borrow

to save' (Collins et al., 2009; Armendáriz and Morduch, 2010; Bauer et al., 2012; Kast and Pomeranz, 2018; Afzal et al., 2018). But take-up is much higher for credit contracts than for commitment savings contract. In addition, we find a significantly higher incidence of repayment difficulties with commitment savings contract and a lower willingness of MFI staff to enforce such contracts.

Second, we then add additional 'behavioral' features in the form of reminders (both for respondents and for respondents' peers) and in variation of repayment flexibility. Our design allows to compare how demand for these features varies between the saving and credit domain. Our findings show that all these contract features are not valued by clients, on the contrary, they appear to be actively disliked. These results have important policy implications for thinking about the future design of microfinance products. Specifically, our results imply that microfinance institutions should *not* be seeking to build explicit commitment features into their products — not because their clients have no demand for commitment devices, but because that demand is already met through the regular repayment schedule implicit in standard microfinance products.

- Afzal, U., G. d'Adda, M. Fafchamps, S. Quinn, and F. Said (2018). Two Sides of the Same Rupee? Comparing Demand for Microcredit and Microsaving in a Framed Field Experiment in Rural Pakistan. *The Economic Journal* 128(614), 2161–2190.
- Allcott, H. and M. Greenstone (2017). Measuring the welfare effects of residential energy efficiency programs. Technical report, National Bureau of Economic Research.
- Allcott, H. and J. B. Kessler (2015). The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons. *NBER Working Paper No. 21671*.
- Andreoni, J., J. M. Rao, and H. Trachtman (2017). Avoiding the Ask: A Field Experiment on Altruism, Empathy, and Charitable Giving. *Journal of Political Economy* 125(3), 625–653.
- Angelucci, M., D. Karlan, and J. Zinman (2015). Microcredit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco. *American Economic Journal: Applied Economics* 7(1), 151–82.
- Armendáriz, B. and J. Morduch (2010). *The Economics of Microfinance*. MIT press.
- Ashraf, N. (2009). Spousal Control and Intra-household Decision Making: An Experimental Study in the Philippines. *American Economic Review* 99(4), 1245–77.
- Ashraf, N., D. Karlan, and W. Yin (2006). Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines. *The Quarterly Journal of Economics* 121(2), 635–672.
- Attanasio, O., B. Augsburg, R. De Haas, E. Fitzsimons, and H. Harmgart (2015). The Impacts of Microfinance: Evidence from Joint-liability Lending in Mongolia. *American Economic Journal: Applied Economics* 7(1), 90–122.
- Attanasio, O. and E. Pastorino (2015). Nonlinear Pricing in Village Economies. *NBER Working Paper No. 21718*.
- Augsburg, B., B. Caeyers, S. Giunti, and B. Malde (2018). Labelled loans, credit constraints and sanitation investments.
- Augsburg, B., R. De Haas, H. Harmgart, and C. Meghir (2015). The Impacts of Microcredit: Evidence from Bosnia and Herzegovina. *American Economic Journal: Applied Economics* 7(1), 183–203.
- Bai, L., B. Handel, E. Miguel, and G. Rao (2017). Self-control and demand for preventive health: Evidence from hypertension in india. Technical report, National Bureau of Economic Research.
- Banerjee, A., E. Duflo, R. Glennerster, and C. Kinnan (2015). The Miracle of Microfinance? Evidence from a Randomized Evaluation. *American Economic Journal: Applied Economics* 7(1), 22–53.
- Banerjee, A., D. Karlan, and J. Zinman (2015). Six Randomized Evaluations of Microcredit: Introduction and Further Steps. *American Economic Journal: Applied Economics* 7(1), 1–21.

- Barboni, G. and P. Agarwal (2018). Knowing what's good for you: Can a repayment flexibility option in microfinance contracts improve repayment rates and business outcomes? Technical report, Working paper.
- Battaglia, M., S. Gulesci, and A. Madestam (2018). Repayment flexibility and risk taking: Experimental evidence from credit contracts.
- Bauer, M., J. Chytilová, and J. Morduch (2012). Behavioral Foundations of Microcredit: Experimental and Survey Evidence from Rural India. *American Economic Review* 102(2), 1118–39.
- Benhassine, N., F. Devoto, E. Duflo, P. Dupas, and V. Pouliquen (2015, August). Turning a shove into a nudge? a "labeled cash transfer" for education. *American Economic Journal: Economic Policy* 7(3), 86–125.
- Bertrand, M., D. Karlan, S. Mullainathan, E. Shafir, and J. Zinman (2010). What's advertising content worth? evidence from a consumer credit marketing field experiment. *The quarterly journal of economics* 125(1), 263–306.
- Besley, T. J., S. Coate, and G. C. Loury (1993). The Economics of Rotating Savings and Credit Associations. *American Economic Review* 83(4), 792–810.
- Breza, E. and A. G. Chandrasekhar (2019). Social Networks, Reputation, and Commitment: Evidence From a Savings Monitors Experiment. *Econometrica* 87(1), 175–216.
- Brune, L., X. Giné, J. Goldberg, and D. Yang (2016). Facilitating savings for agriculture: Field experimental evidence from malawi. *Economic Development and Cultural Change* 64(2), 187–220.
- Brune, L. and J. T. Kerwin (2019). Income Timing and Liquidity Constraints: Evidence from a Randomized Field Experiment. *Journal of Development Economics* 138, 294–308.
- Bryan, G., D. Karlan, and S. Nelson (2010). Commitment devices. *Annual Review of Economics* 2(1), 671–698.
- Casaburi, L. and R. Macchiavello (2018). Firm and Market Response to Saving Constraints: Evidence from the Kenyan Dairy Industry. *Working Paper*.
- Castellanos, S. G., D. Jiménez-Hernández, A. Mahajan, and E. Seira (2019). The Difficulties of Financial Inclusion via Large Banks: Evidence from Mexico. *Working paper*.
- Chernozhukov, V., M. Demirer, E. Duflo, and I. Fernández-Val (2018). Generic Machine Learning Inference on Heterogenous Treatment Effects in Randomized Experiments. *NBER Working Paper* 24678.
- Collins, D., J. Morduch, S. Rutherford, and O. Ruthven (2009). *Portfolios of the Poor: How the World's Poor Live on \$2 a Day*. Princeton University Press.
- Crépon, B., F. Devoto, E. Duflo, and W. Parienté (2015). Estimating the Impact of Microcredit on Those who take it up: Evidence from a Randomized Experiment in Morocco. *American Economic Journal: Applied Economics* 7(1), 123–50.

- Czura, K. (2015). Do flexible repayment schedules improve the impact of microcredit? Discussion papers in economics, University of Munich, Department of Economics.
- d’Adda, G., Y. Gao, R. Golman, and M. Tavoni (2018). It’s so Hot in Here: Information Avoidance, Moral Wiggle Room, and High Air Conditioning Usage. *FEEM Working Paper No. 07.2018*.
- Damgaard, M. T. and C. Gravert (2018). The Hidden Costs of Nudging: Experimental Evidence from Reminders in Fundraising. *Journal of Public Economics* 157, 15–26.
- Dana, J., R. A. Weber, and J. X. Kuang (2007). Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness. *Economic Theory* 33(1), 67–80.
- De Mel, S., C. McIntosh, K. Sheth, and C. Woodruff (2018, December). Can Mobile-Linked Bank Accounts Bolster Savings? Evidence from a Randomized Controlled Trial in Sri Lanka. CEPR Discussion Papers 13378, C.E.P.R. Discussion Papers.
- DellaVigna, S. and U. Malmendier (2004). Contract design and self-control: Theory and evidence. *The Quarterly Journal of Economics* 119(2), 353–402.
- DellaVigna, S. and U. Malmendier (2006). Paying not to go to the gym. *American Economic Review* 96(3), 694–719.
- DellaVigna, S. and D. Pope (2017). What motivates effort? evidence and expert forecasts. *The Review of Economic Studies* 85(2), 1029–1069.
- Dertwinkel-Kalt, M., H. Gerhardt, G. Riener, F. Schwerter, and L. Strang (2017). Concentration Bias in Intertemporal Choice.
- Duflo, E., M. Kremer, and J. Robinson (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from kenya. *American Economic Review* 101(6), 2350–90.
- Dupas, P., D. Karlan, J. Robinson, and D. Ubfal (2018, April). Banking the unbanked? evidence from three countries. *American Economic Journal: Applied Economics* 10(2), 257–97.
- Dupas, P. and J. Robinson (2013a). Savings Constraints and Microenterprise Development: Evidence from a Field Experiment in Kenya. *American Economic Journal: Applied Economics* 5(1), 163–92.
- Dupas, P. and J. Robinson (2013b). Why Don’t the Poor Save More? Evidence from Health Savings Experiments. *American Economic Review* 103(4), 1138–71.
- Falk, A. and M. Kosfeld (2006). The Hidden Costs of Control. *American Economic Review* 96(5), 1611–1630.
- Fehr, E. and J. A. List (2004). The Hidden Costs and Returns of Incentives: Trust and Trustworthiness Among CEOs. *Journal of the European Economic Association* 2(5), 743–771.
- Field, E. and R. Pande (2008). Repayment frequency and default in microfinance: Evidence from india. *Journal of the European Economic Association* 6(2/3), 501–509.

- Field, E., R. Pande, J. Papp, and N. Rigol (2013). Does the classic microfinance model discourage entrepreneurship among the poor? experimental evidence from india. *American Economic Review* 103(6), 2196–2226.
- Gugerty, M. K. (2007). You Can’t Save Alone: Commitment in Rotating Savings and Credit Associations in Kenya. *Economic Development and cultural change* 55(2), 251–282.
- Hussam, R., N. Rigol, and B. Roth (2017). Targeting High Ability Entrepreneurs using Community Information: Mechanism Design in the Field. *Working paper*.
- Jakiela, P. and O. Ozier (2015, 09). Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies. *The Review of Economic Studies* 83(1), 231–268.
- John, A. (2019). When Commitment Fails — Evidence from a Field Experiment. *Management Science*.
- Karlan, D. and L. L. Linden (2014, January). Loose knots: Strong versus weak commitments to save for education in uganda. Working Paper 19863, National Bureau of Economic Research.
- Karlan, D., M. McConnell, S. Mullainathan, and J. Zinman (2016). Getting to the Top of Mind: How Reminders Increase Saving. *Management Science* 62(12), 3393–3411.
- Karlan, D., J. Morduch, and S. Mullainathan (2010). *Take-up: Why microfinance take-up rates are low and why it matters*. Financial Access Initiative research framing note.
- Karlan, D., S. Mullainathan, and B. N. Roth (2019). Debt Traps? Market Vendors and Moneylender Debt in India and the Philippines. *American Economic Review: Insights*.
- Karlan, D. and J. Zinman (2011). Microcredit in Theory and Practice: Using Randomized Credit Scoring for Impact Evaluation. *Science* 332(6035), 1278–1284.
- Kast, F. and D. Pomeranz (2018). Saving Accounts to Borrow Less: Experimental Evidence from Chile. *Working paper*.
- Kaur, S., M. Kremer, and S. Mullainathan (2015). Self-control at work. *Journal of Political Economy* 123(6), 1227–1277.
- Kőszegi, B. and A. Szeidl (2012). A model of focusing in economic choice. *The Quarterly journal of economics* 128(1), 53–104.
- Laibson, D. (2015). Why Don’t Present-biased Agents make Commitments? *American Economic Review* 105(5), 267–72.
- Laibson, D. (2018). Private paternalism, the commitment puzzle, and model-free equilibrium (richard t. ely lecture). *American Economic Review Papers and Proceedings* 108, 1–21.
- Liu, E. and B. N. Roth (2019). Contractual Restrictions and Debt Traps. *Working paper*.
- Meager, R. (2018a). Aggregating Distributional Treatment Effects: A Bayesian Hierarchical Analysis of the Microcredit Literature. *Working paper*.

- Meager, R. (2018b). Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments. *American Economic Journal: Applied Economics*.
- Morduch, J. (2010). Borrowing to Save. *Journal of Globalization and Development* 1(2).
- Rutherford, S. (2000). *The Poor and Their Money*. Oxford University Press.
- Schilbach, F. (2019). Alcohol and Self-Control: A Field Experiment in India. *American Economic Review*.
- Stango, V. and J. Zinman (2014). Limited and varying consumer attention: Evidence from shocks to the salience of bank overdraft fees. *The Review of Financial Studies* 27(4), 990–1030.
- Tarozzi, A., J. Desai, and K. Johnson (2015). The Impacts of Microcredit: Evidence from Ethiopia. *American Economic Journal: Applied Economics* 7(1), 54–89.

Tables and Figures

Table 1: An illustrative contract structure

	WEEK 0	WEEK 1	WEEK 2	WEEK 3	WEEK 4	WEEK 5	WEEK 6
Participant pays	<i>take up</i>		1000	1000	1000	1000	1000
Bank pays	<i>decision</i>	4500					

This table shows a payment schedule for a basic contract with lump-sum in Week 1 and interest of -10% .

Table 2: Sample structure across phases and locations

	DISTRICT	OFFICES	RESPONDENTS
<i>Phase 1</i>	Bhakkar	3	418
	Chakwal	5	372
<i>Total</i>		8	790
<i>Phase 2</i>	Khushab	5	725
	Mandi Bahauddin	4	674
	Jhelum	6	296
	Rawalpindi	2	721
<i>Total</i>		17	2416

This table shows the breakdown of our 3206 respondents, between Phase 1 (790 respondents) and Phase 2 (2416 respondents).

Table 3: **Structure of treatments**

<u>Phase 1</u>		
Basic treatment (1/2) n = 394		
Control group (1/2) n = 396		
<u>Phase 2</u>		
Basic treatment with no reminders (1/12) (n = 197)	Basic treatment with respondent reminders (1/12) (n = 204)	Basic treatment with peer reminders (1/12) (n = 199)
Sunk treatment with no reminders (1/12) (n = 201)	Sunk treatment with respondent reminders (1/12) (n = 202)	Sunk treatment with peer reminders (1/12) (n = 207)
Flex treatment with no reminders (1/12) (n = 202)	Flex treatment with respondent reminders (1/12) (n = 204)	Flex treatment with peer reminders (1/12) (n = 198)
Control group (1/4) n = 602		

This table shows the structure of treatments: a simple treatment/control division in Phase 1, and a 3×3 factorial design with controls in Phase 2. In each case, the fractions (1/2, 1/4 and 1/12) show the proportion of the respondents in the phase who were intended for assignment; in each case 'n' refers to the actual number assigned.

Table 4: Average take-up by contract terms

ALL SUBJECTS

		<i>Lumpsum amount</i>		
		4500	5000	5500
Phase 1	<i>Lumpsum paid in</i>			
	Week 1	8.2%	30.2%	47.0%
	Week 6	2.7%	4.3%	11.0%
		<i>Lumpsum amount</i>		
		3200	3500	3800
Phase 2	<i>Lumpsum paid in</i>			
	Week 1	11.0%	26.0%	37.2%
	Week 8	4.1%	8.9%	11.3%

EXCLUDING AUTOMATIC REFUSERS

		<i>Lumpsum amount</i>		
		4500	5000	5500
Phase 1	<i>Lumpsum paid in</i>			
	Week 1	12.8%	43.1%	64.2%
	Week 6	4.3%	6.8%	17.6%
		<i>Lumpsum amount</i>		
		3200	3500	3800
Phase 2	<i>Lumpsum paid in</i>			
	Week 1	20.0%	41.9%	57.0%
	Week 8	7.2%	14.6%	19.3%

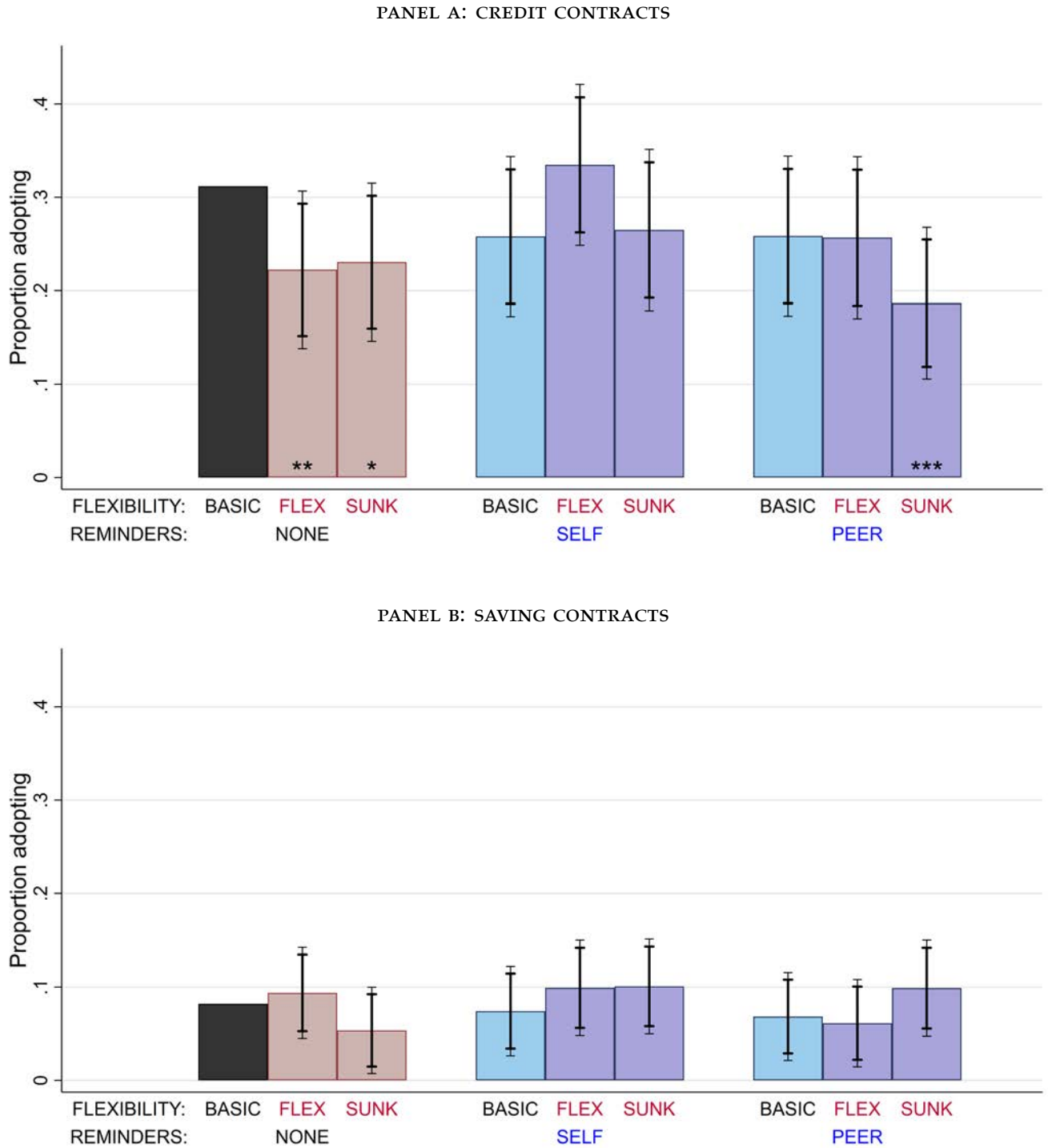
This table shows the average take-up rates by contractual terms (lumpsum value and timing). Weekly instalments were PKR 1000 in Phase 1 and PKR 500 in Phase 2. 'Automatic refusers' refers to respondents who declined the contract even before knowing the contractual terms on offer.

Table 5: **Proportion of individuals who contradict a standard model**

	Phase 1	Phase 2
<i>Subjects were offered at least one loan charging a zero or positive interest AND a savings contract with a zero or negative return:</i>		
took neither	76	225
took the loan	30	102
took the saving contract	7	47
took both	6	24
<i>Total:</i>	107	350
conditional on loan, take saving	20%	24%
conditional on saving, take loan	86%	51%
<i>Subject was offered at least one loan charging a negative interest AND at least one loan charging a zero or positive interest OR a savings contract with a zero or negative return:</i>		
took both	6	39
took the subsidized loan	89	315
took the loan or saving contract	18	123
took neither	0	0
<i>Total:</i>	101	399
conditional on taking a loan or saving contract, refused subsidized loan	67%	68%
conditional on refusing a subsidized loan, took a loan or saving contract	100%	100%

This table reports the proportion of individuals whose take-up behaviour would contradict a standard model, using two descriptive tests. Automatic refusers are omitted from these calculations.

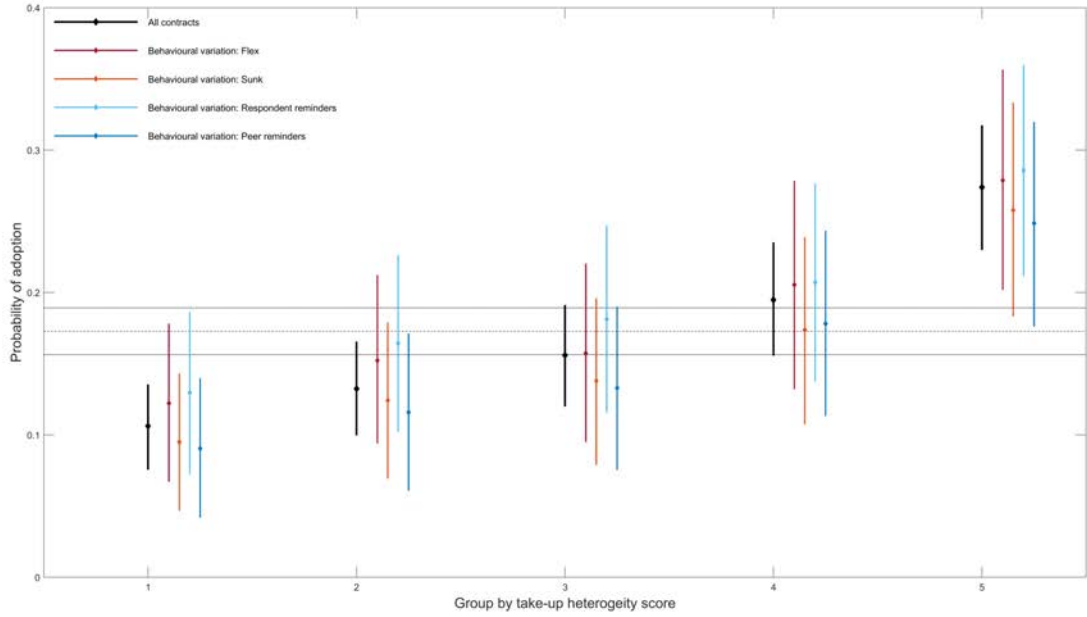
Figure 1: Average take-up by behavioral variations



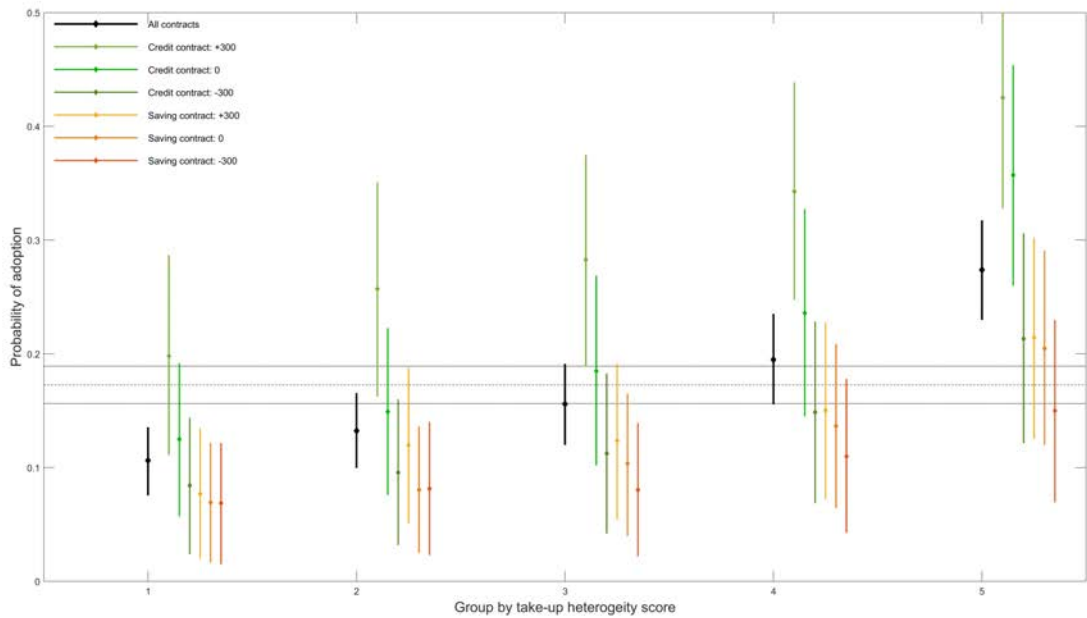
This figure shows the average take-up for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'peer' variation), and take-up for each of the eight possible variations. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic contract. Stars indicate a significant difference from take-up of the basic contract; that is, we reject a null hypothesis of equal take-up rates for the 'sunk' variation and for the 'sunk and peer' variation, each at the 5% significance level.

Figure 2: Group Average Treatment Effects (sorted by take-up propensity)

PANEL A: TAKE-UP BY BEHAVIORAL VARIATION



PANEL B: TAKE-UP BY CONTRACTUAL TERMS (PAYMENT AND TIMING)



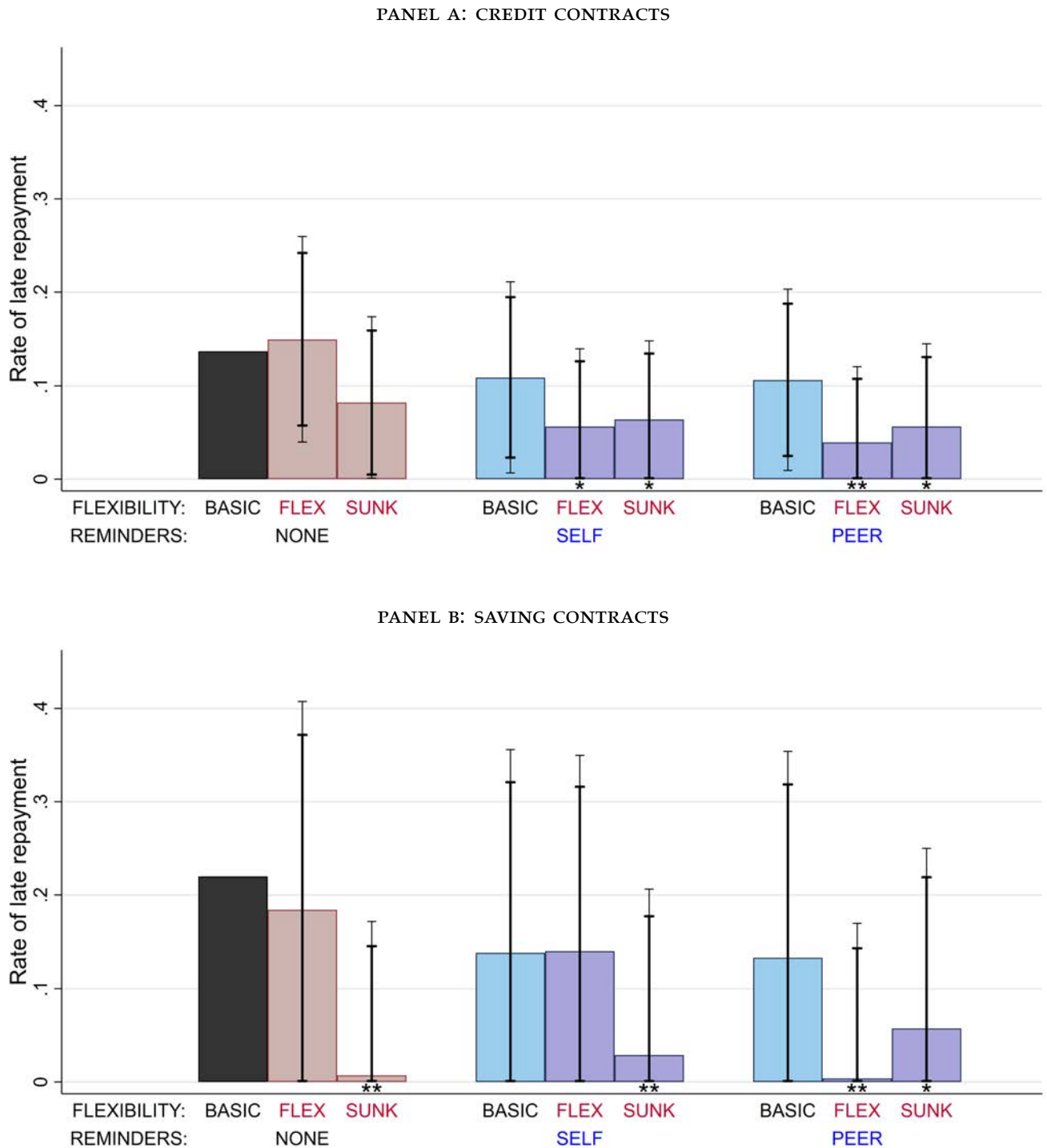
This figure shows the Group Average Treatment Effects, sorted by the take-up propensity estimated in the main text. In each figure, the leftmost (black) lines for each group show the average probability of take-up across all contract types; note that these leftmost lines are identical across figures (allowing for a different scaling of the vertical axis). In the top panel, the four subsequent lines in each group (in color) show the average take-up across 'flex', 'sunk', 'respondent reminders' and 'peer reminders' respectively. In the bottom panel, the six subsequent lines in each group (in color) show the average take-up across the six different variations on contract payment and timing. For each category, the graphs show point estimates and 90% confidence intervals (both formed using the bootstrap methodology proposed by [Chernozhukov et al. \(2018\)](#)).

Table 6: Cluster Analysis: Descriptive characteristics of extreme groups

	20% LEAST LIKELY TO ADOPT		20% MOST LIKELY TO ADOPT		DIFF. (p)		
	ESTIMATE	90% CONFIDENCE	ESTIMATE	90% CONFIDENCE			
<i>Saving challenges:</i>							
Dummy: Finds it hard to save	0.54	0.48	0.60	0.89	0.86	0.93	0.00***
Dummy: Faces pressure to share	0.55	0.49	0.61	0.94	0.91	0.97	0.00***
<i>Keeping track:</i>							
Dummy: Good at keeping track of time	0.85	0.81	0.90	0.58	0.52	0.64	0.00***
Dummy: Good at keeping track of finances	0.78	0.73	0.83	0.47	0.41	0.52	0.00***
Dummy: Follows a strict schedule on finances	0.77	0.72	0.82	0.50	0.44	0.56	0.00***
Dummy: Follows a tight routine	0.61	0.56	0.67	0.41	0.35	0.47	0.00***
Dummy: Acts early to avoid forgetting	0.58	0.52	0.64	0.48	0.42	0.54	0.02**
Dummy: Acts early to avoid forgetting finances	0.57	0.51	0.63	0.43	0.37	0.49	0.00***
Dummy: Keeps cash earmarked	0.64	0.59	0.70	0.47	0.41	0.53	0.00***
Dummy: Keeps funds earmarked in accounts	0.17	0.12	0.21	0.14	0.10	0.19	0.36
Dummy: Present bias	0.08	0.05	0.12	0.12	0.08	0.15	0.22
Dummy: Future bias	0.15	0.11	0.20	0.07	0.04	0.10	0.00***
<i>Women empowerment:</i>							
Share of examples where view always considered	0.61	0.57	0.66	0.73	0.69	0.77	0.00***
Appropriate for a woman to invest in her business	0.14	0.10	0.19	0.41	0.35	0.47	0.00***
Appropriate for a woman to buy a scarf	0.20	0.15	0.25	0.48	0.42	0.54	0.00***

This table provides a cluster analysis of baseline covariates with a specific behavioral interpretation. Specifically, we describe the characteristics of those respondents in the 'most affected' and 'least affected' groups, defined in terms of estimated probability of adopting. We provide average characteristics, confidence intervals and a p-value on a test of equality of means ('diff. (p)') using the methodology proposed by Chernozhukov et al. (2018).

Figure 3: Rate of late repayment by behavioral variations



This figure shows the rate of late repayment for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'peer' variation), and for each of the eight possible variations. Note that we are here studying the rate of late repayment; that is, we use a linear probability model for having delayed payment, for the subsample of observations where the respondent agreed to the contract. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic contract. Stars indicate a significant difference from the basic contract.

Table 7: Summary of ITT and LATE estimates of business outcomes

	Phase 1		Phase 2	
	Control mean	ITT	Control mean	ITT
		LATE		LATE
Runs a business	0.606	0.016 (0.019)	0.125	-0.009 (0.013)
Number of businesses	1.116	0.087* (0.047)	0.156	-0.002 (0.017)
Value of capital invested in business	7803	610 (607)	2023	-301 (371)
Value of monthly sales	8184	709 (519)	1237	-42 (188)
Value of monthly expenses	6228	152 (452)	502	-42 (82)
Monthly profit (sales - expenses)	1871	737** (353)	665	25 (112)
Monthly profit (self-reported)	2933	518 (329)	869	-23 (130)
Observations		789		1991

*This table reports regression estimates of equation 2. We report standard errors under each coefficient in parentheses. All values are in Pakistani rupees. Monthly self-reported profits include the imputed values of business goods consumed. Confidence: * \leftrightarrow $p < 0.1$; ** \leftrightarrow $p < 0.05$; *** \leftrightarrow $p < 0.01$.*

Appendix

Table A1: Description of the sample - Phase 1

	N	Mean	Treatment balance (p)	Terms balance (p)
Dummy: participates in a committee	790	0.7	0.176	0.957
Total amount owed by individual (PKR)	790	17695.1	0.281	0.345
Total household consumption in the last month (PKR)	780	25581.9	0.454	0.945
Total value of assets owned by household (PKR)	790	47662.6	0.052	0.357
Dummy: runs a business	790	0.6	0.783	0.341
Number of businesses owned by respondent or household	790	0.9	0.186	0.663
Value of total capital invested in business(es) (PKR)	790	9633.6	0.554	0.310
Total monthly sales of the business (PKR)	790	9602.9	0.591	0.827
Total monthly expense of the business (PKR)	790	6688.4	0.393	0.768
Total monthly profits(1) of the business (PKR)	790	2834.2	0.789	0.234
Total monthly profits(2) of the business (PKR)	789	4029.3	0.785	0.339
Dummy: finds it hard to save	790	0.6	0.144	0.297
Index: opinions taken into account in household decisions	790	-0.0	0.928	0.768
Index: needs to ask permission for making decisions	790	0.0	0.078	0.671
Dummy: faces pressure to share cash on hand	790	0.6	0.523	0.099
Age (years)	790	38.0	0.212	0.157
Dummy: is currently married	790	0.8	0.567	0.774
Number or years of education	790	4.7	0.098	0.220
Dummy: can read and write	790	0.5	0.151	0.717
Number of children	790	3.4	0.096	0.338
Dummy: is the household head	790	0.1	0.937	0.601

This table provides basic summary statistics for sample characteristics. For each variable, ‘Treatment balance’ reports a p-value from a joint test of the null hypothesis that the variance is balanced across the treatment status and ‘Terms balance’ reports a p-value from a joint test of the null hypothesis that the variance is balanced across the contract terms (interest and week of lumpsum payment)

Table A2: Description of the sample - Phase 2

	N	Mean	Treatment balance (p)	Terms balance (p)
Dummy: participates in a committee	2416	0.2	0.842	0.644
Total amount owed by individual (PKR)	2406	12061.2	0.851	0.060
Total household consumption last month (PKR)	2416	19312.2	0.143	0.169
Total household monthly income (PKR)	2407	19958.2	0.720	0.710
Total value of assets owned by household (PKR)	2416	35546.4	0.713	0.469
Dummy: runs a business	2416	0.1	0.785	0.964
Number of businesses owned by respondent or household	2416	0.2	0.907	0.994
Total capital invested in business(es)	2416	2182.0	0.881	0.318
Total monthly sales of the business (PKR)	2416	1218.8	0.730	0.978
Total monthly expense of the business (PKR)	2416	551.2	0.980	0.991
Total monthly profit(1) of the business (PKR)	2416	617.2	0.256	0.701
Total monthly profit(2) of the business (PKR)	2416	787.3	0.930	0.679
Dummy: finds it hard to save	2416	0.7	0.159	0.244
Index: opinions taken into account in household decisions	2416	0.0	0.042	0.101
Dummy: faces pressure to share cash on hand	2416	0.8	0.003	0.000
Index: needs to ask permission for making decisions	2416	0.0	0.005	0.003
Age (years)	2416	39.1	0.473	0.667
Dummy: is currently married	2416	0.8	0.398	0.346
Number of years of education	2416	4.3	0.098	0.000
Dummy: can read and write	2416	0.5	0.250	0.000
Number of children	2416	3.5	0.704	0.481
Dummy: is the household head	2416	0.2	0.357	0.177

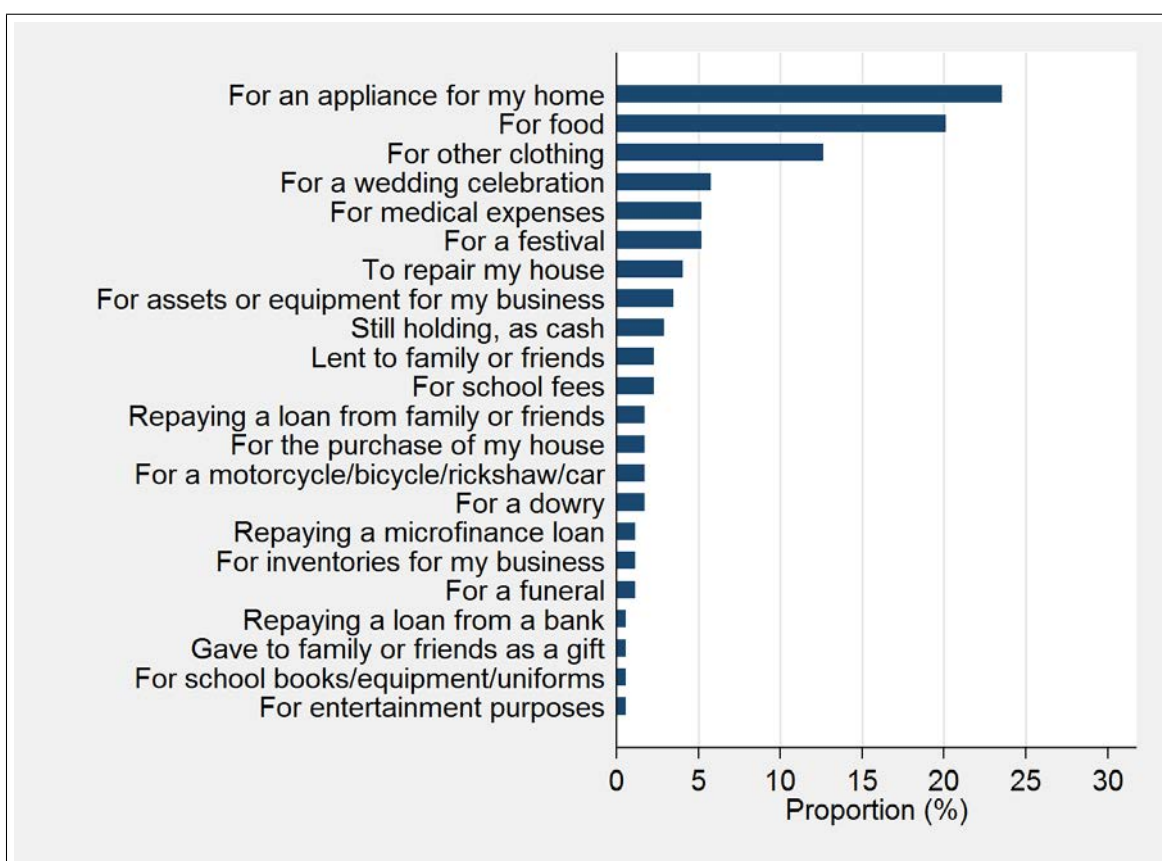
This table provides basic summary statistics for sample characteristics. For each variable, ‘Treatment balance’ reports a p-value from a joint test of the null hypothesis that the variance is balanced across the treatment status and ‘Terms balance’ reports a p-value from a joint test of the null hypothesis that the variance is balanced across the contract terms (interest and week of lumpsum payment)

Table A3: Proportion of treated who refused the product offered before the contract offer cards were drawn

	Phase 1	Phase 2
Percent of refusers in:		
Cycle 1	26.1%	63.0%
Cycle 2	33.5%	27.7%
Cycle 3	40.1%	29.6%
All cycles	33.3%	40.1%
Percent of subjects refusing:		
Never	58.6%	29.4%
Once	7.6%	40.6%
Twice	9.1%	10.5%
Three times	24.6%	19.6%
Number of treated subjects	394	1814

This table shows the proportion of the treated who refused the product offered before the cards determining the contract terms were drawn.

Figure A4: **Spending of the lumpsum: Most common categories**



This figure shows the most common categories listed by respondents in describing how they used the lumpsum payments from Phase 2.

Table A4: Average take-up by behavioral variations: Credit contracts

Panel 1: Average take-up in each treatment combination

	Flex	Basic	Sunk
No reminders	22.2%	31.2%	23.1%
Reminder to self	33.5%	25.8%	26.5%
Reminder to peers	25.7%	25.9%	18.7%

Joint equality test (*p*-value): 0.011**

Panel 2: Pairwise take-up comparisons

A. Difference from basic contract with no reminders

	Flex	Basic	Sunk
No reminders	-8.9%**	<i>reference</i>	-8.1%*
Reminder to self	2.3%	-5.4%	-4.7%
Reminder to peers	-5.5%	-5.3%	-12.5%***

B. Difference from basic contract

	Flex	Basic	Sunk
No reminders	-8.9%**	<i>reference</i>	-8.1%*
Reminder to self	7.7%*	<i>reference</i>	0.7%
Reminder to peers	-0.2%	<i>reference</i>	-7.2%*

C. Difference from no reminder contract

	Flex	Basic	Sunk
No reminders	<i>reference</i>	<i>reference</i>	<i>reference</i>
Reminder to self	11.2%***	-5.4%	3.4%
Reminder to peers	3.4%	-5.3%	-4.4%

All the calculations in this Table are based on an OLS regression of take-up on all interactions between reminder and commitment treatments. Interaction terms for payment week and interest rate are included as controls. Standard errors clustered at the household level. We use '*' to denote confidence at the 90% level. For Panel 2A, *p*-values for pairwise tests come from OLS coefficient estimates. For Panels 2B and 2C, *p*-values come from the relevant pairwise coefficient tests.

Table A5: Average take-up by behavioral variations: Savings contracts

Panel 1: Average take-up in each treatment combination			
	Flex	Basic	Sunk
No reminders	9.4%	8.2%	5.4%
Reminder to self	9.9%	7.4%	10.1%
Reminder to peers	6.1%	6.8%	9.9%
<i>Joint equality test (p-value): 0.321</i>			
Panel 2: Pairwise take-up comparisons			
A. Difference from basic contract with no reminders			
	Flex	Basic	Sunk
No reminders	1.2%	<i>reference</i>	-2.9%
Reminder to self	1.7%	-0.8%	1.9%
Reminder to peers	-2.1%	-1.4%	1.7%
B. Difference from basic contract			
	Flex	Basic	Sunk
No reminders	1.2%	<i>reference</i>	-2.9%
Reminder to self	2.5%	<i>reference</i>	2.7%
Reminder to peers	-0.7%	<i>reference</i>	3.1%
C. Difference from no reminder contract			
	Flex	Basic	Sunk
No reminders	<i>reference</i>	<i>reference</i>	<i>reference</i>
Reminder to self	0.5%	-0.8%	4.7%
Reminder to peers	-3.3%	-1.4%	4.5%

All the calculations in this Table are based on an OLS regression of take-up on all interactions between reminder and commitment treatments. Interaction terms for payment week and interest rate are included as controls. Standard errors clustered at the household level. We use '*' to denote confidence at the 90% level. For Panel 2A, p-values for pairwise tests come from OLS coefficient estimates. For Panels 2B and 2C, p-values come from the relevant pairwise coefficient tests.

Table A6: Average take-up by behavioral variations: Pooled across credit and saving

Panel 1: Average take-up in each treatment combination

	Flex	Basic	Sunk
No reminders	16.0%	20.1%	14.5%
Reminder to self	22.2%	16.9%	18.5%
Reminder to peers	16.2%	16.6%	14.6%

Joint equality test (*p*-value): 0.087*

Panel 2: Pairwise take-up comparisons

A. Difference from basic contract with no reminders

	Flex	Basic	Sunk
No reminders	-4.1%	<i>reference</i>	-5.6%**
Reminder to self	2.1%	-3.2%	-1.6%
Reminder to peers	-3.9%	-3.5%	-5.5%**

B. Difference from basic contract

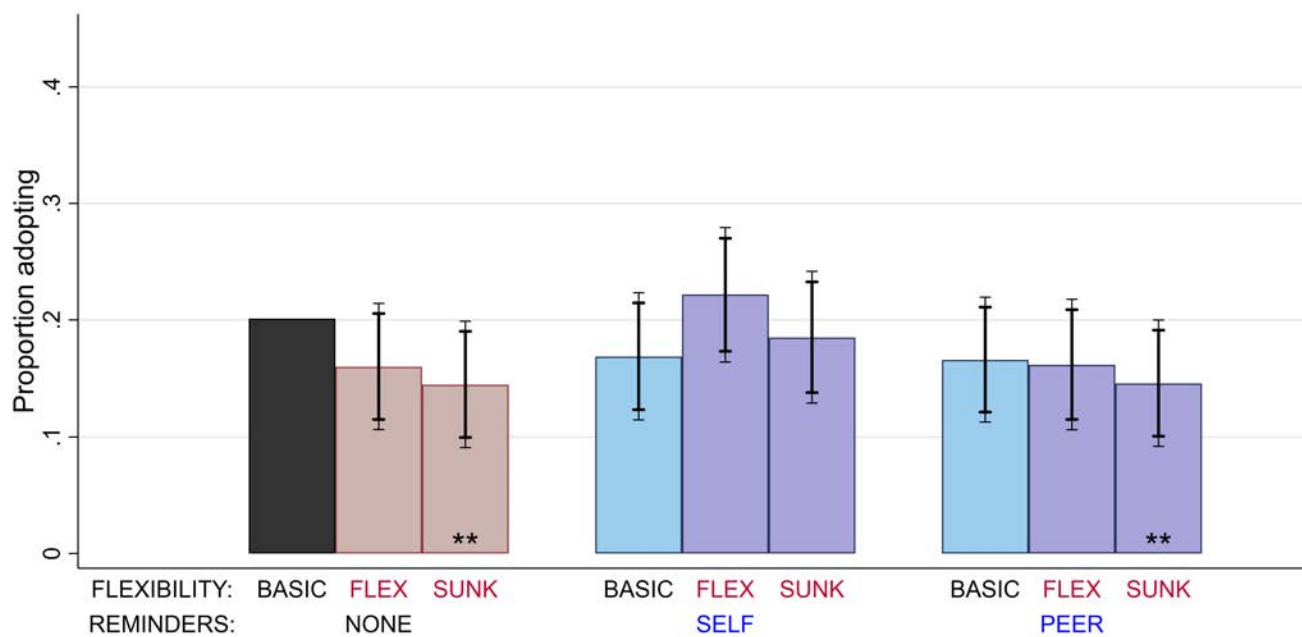
	Flex	Basic	Sunk
No reminders	-4.1%	<i>reference</i>	-5.6%**
Reminder to self	5.3%**	<i>reference</i>	1.6%
Reminder to peers	0.4%	<i>reference</i>	-2.0%

C. Difference from no reminder contract

	Flex	Basic	Sunk
No reminders	<i>reference</i>	<i>reference</i>	<i>reference</i>
Reminder to self	6.2%**	-3.2%	4.0%
Reminder to peers	0.2%	-3.5%	0.1%

All the calculations in this Table are based on an OLS regression of take-up on all interactions between reminder and commitment treatments. Interaction terms for payment week and interest rate are included as controls. Standard errors clustered at the household level. We use '*' to denote confidence at the 90% level. For Panel 2A, *p*-values for pairwise tests come from OLS coefficient estimates. For Panels 2B and 2C, *p*-values come from the relevant pairwise coefficient tests.

Figure A5: **Average take-up by behavioral variations: Pooled across credit and saving**



This figure shows the average take-up for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'peer' variation), and take-up for each of the eight possible variations. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic contract. Stars indicate a significant difference from take-up of the basic contract; that is, we reject a null hypothesis of equal take-up rates for the 'sunk' variation and for the 'sunk and peer' variation, each at the 5% significance level.

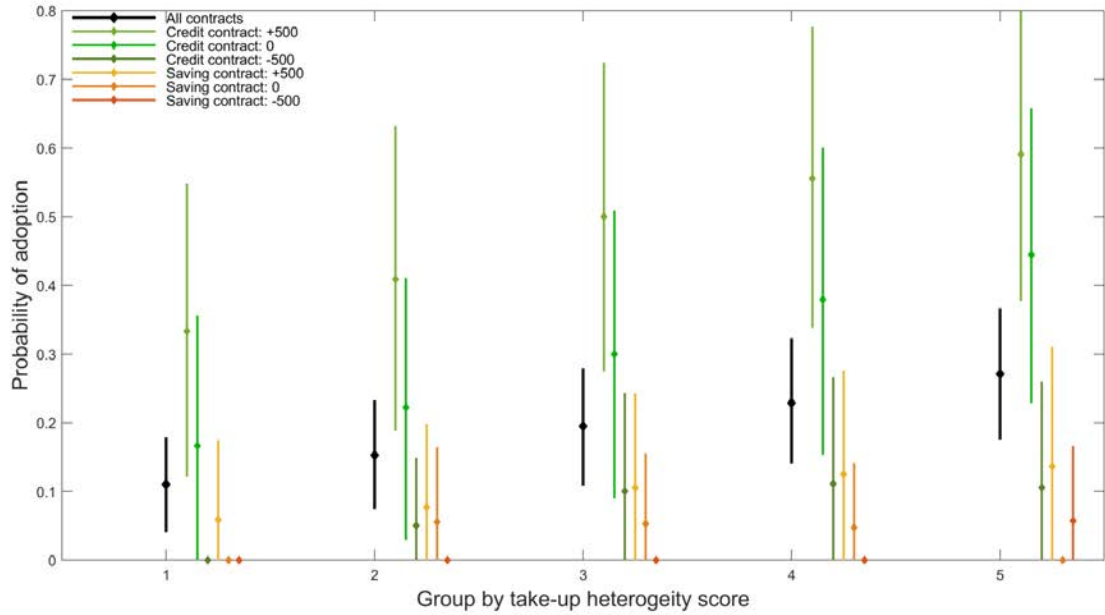
Table A7: Cluster Analysis: Description of extreme groups (all covariates: Phase 2)

	20% LEAST LIKELY TO ADOPT			20% MOST LIKELY TO ADOPT			DIFF. (<i>p</i>)
	ESTIMATE	90% CONFIDENCE		ESTIMATE	90% CONFIDENCE		
Dummy: Age ≤ 28	0.26	0.21	0.31	0.16	0.11	0.20	0.00***
Dummy: Age ≤ 34	0.39	0.34	0.45	0.31	0.25	0.36	0.05**
Dummy: Age ≤ 40	0.68	0.63	0.74	0.52	0.46	0.58	0.00***
Dummy: Age ≤ 48	0.82	0.78	0.87	0.79	0.74	0.84	0.21
Number of daughters	1.33	1.18	1.48	1.79	1.63	1.95	0.00***
Dummy: Missing the number of daughters	0.12	0.09	0.16	0.05	0.02	0.07	0.00***
Digit span test score	4.76	4.62	4.90	4.24	4.09	4.39	0.00***
Dummy: Education to class 5	0.18	0.13	0.22	0.14	0.10	0.19	0.30
Dummy: Education to class 8	0.10	0.07	0.14	0.07	0.04	0.10	0.21
Dummy: Education to degree	0.04	0.02	0.07	0.02	0.00	0.04	0.19
Dummy: Education to matric	0.39	0.33	0.44	0.05	0.02	0.07	0.00***
Household size	5.23	5.00	5.46	6.42	6.18	6.66	0.00***
Dummy: Household head	0.09	0.06	0.13	0.28	0.22	0.33	0.00***
Dummy: Literate	0.73	0.68	0.78	0.32	0.26	0.37	0.00***
Dummy: Married	0.80	0.75	0.85	0.78	0.73	0.83	0.27
Dummy: Correct on math question 1	0.76	0.71	0.81	0.34	0.29	0.40	0.00***
Dummy: Correct on math question 2	0.63	0.57	0.68	0.69	0.63	0.74	0.18
Dummy: Self-employed	0.06	0.03	0.09	0.22	0.17	0.26	0.00***
Number of sons	1.34	1.20	1.48	2.07	1.92	2.22	0.00***
Dummy: Missing the number of sons	0.12	0.09	0.16	0.05	0.02	0.07	0.00***
Dummy: Spouse of household head	0.71	0.66	0.76	0.59	0.53	0.65	0.01***
Dummy: Has a wage job	0.11	0.07	0.14	0.11	0.08	0.15	0.50
Dummy: Currently in a savings committee	0.10	0.06	0.13	0.25	0.20	0.30	0.00***
Dummy: Has experience in a savings committee	0.26	0.20	0.31	0.30	0.24	0.35	0.22
Monthly household consumption (z-score)	-0.28	-0.38	-0.18	0.36	0.23	0.49	0.00***
Dummy: Has a bank account	0.11	0.07	0.15	0.02	0.01	0.04	0.00***
Household income last week (z-score)	-0.02	-0.13	0.09	0.01	-0.13	0.14	0.53
Dummy: Missing household income	0.04	0.02	0.06	0.03	0.01	0.05	0.34
Dummy: Currently owes family or friends	0.16	0.12	0.21	0.06	0.03	0.08	0.00***
Dummy: Currently owes an MFI	0.01	-0.00	0.03	0.03	0.01	0.05	0.20
Dummy: Currently owes NRSP	0.14	0.10	0.18	0.64	0.58	0.69	0.00***
Number of minutes to walk to NRSP (z-score)	0.12	-0.00	0.25	-0.13	-0.24	-0.03	0.01***
Dummy: Acts early to avoid forgetting	0.58	0.52	0.64	0.48	0.42	0.54	0.03**
Dummy: Acts early to avoid forgetting finances	0.56	0.51	0.62	0.43	0.37	0.49	0.00***
Appropriate for a woman to buy a scarf	0.20	0.15	0.25	0.47	0.41	0.53	0.00***
Appropriate for a woman to invest in her business	0.15	0.11	0.19	0.41	0.35	0.47	0.00***
Dummy: Keeps cash earmarked	0.64	0.58	0.70	0.47	0.41	0.53	0.00***
Share of examples where view always considered	0.61	0.57	0.66	0.73	0.69	0.77	0.00***
Dummy: Usually makes final decision on spending	0.79	0.75	0.84	0.59	0.53	0.64	0.00***
Dummy: Keeps funds earmarked in accounts	0.17	0.12	0.21	0.14	0.10	0.18	0.38
Dummy: Future bias	0.15	0.11	0.20	0.07	0.04	0.10	0.00***
Would keep a gift for herself	0.84	0.79	0.88	0.42	0.36	0.47	0.00***
Dummy: Good at keeping track of time	0.85	0.81	0.89	0.58	0.52	0.64	0.00***
Dummy: Good at keeping track of finances	0.78	0.73	0.83	0.47	0.41	0.52	0.00***
Dummy: Finds it hard to save	0.54	0.48	0.60	0.90	0.86	0.93	0.00***
Patience measure: Maximum measured patience	0.28	0.22	0.33	0.65	0.60	0.71	0.00***
Maximum measured patience in future frame	0.26	0.21	0.31	0.68	0.62	0.74	0.00***
Dummy: Present bias	0.08	0.05	0.12	0.12	0.08	0.16	0.18
Dummy: Faces pressure to share	0.55	0.49	0.61	0.94	0.91	0.97	0.00***
Risk aversion measure 1 (higher is more risk-tolerant)	0.27	0.24	0.29	0.29	0.27	0.31	0.14
Risk aversion measure 2 (higher is more risk-tolerant)	0.18	0.16	0.20	0.35	0.32	0.37	0.00***
Dummy: Others remind of appointments	0.57	0.51	0.63	0.43	0.37	0.49	0.00***
Dummy: Others remind of financial obligations	0.56	0.50	0.62	0.37	0.31	0.43	0.00***
Dummy: Would immediately spend 100 rupees if found	0.27	0.21	0.32	0.27	0.22	0.33	0.44
Dummy: Follows a strict schedule on finances	0.77	0.72	0.82	0.49	0.43	0.55	0.00***
Dummy: Follows a tight routine	0.61	0.55	0.66	0.41	0.35	0.47	0.00***
Patience measure (higher is more patient)	4.76	4.49	5.04	6.63	6.38	6.87	0.00***
Patience measure in future frame	4.59	4.31	4.87	6.75	6.51	6.99	0.00***

This table provides a cluster analysis of all the baseline covariates used for the machine learning analysis for Phase 2. Specifically, we describe the characteristics of those respondents in the ‘most affected’ and ‘least affected’ groups, defined in terms of estimated probability of adopting. We provide average characteristics, confidence intervals and a p-value on a test of equality of means (‘DIFF. (p)’) using the methodology proposed by Chernozhukov et al. (2018).

Figure A6: **Group Average Treatment Effects (sorted by take-up propensity)**

TAKE-UP BY CONTRACTUAL TERMS (PAYMENT AND TIMING)



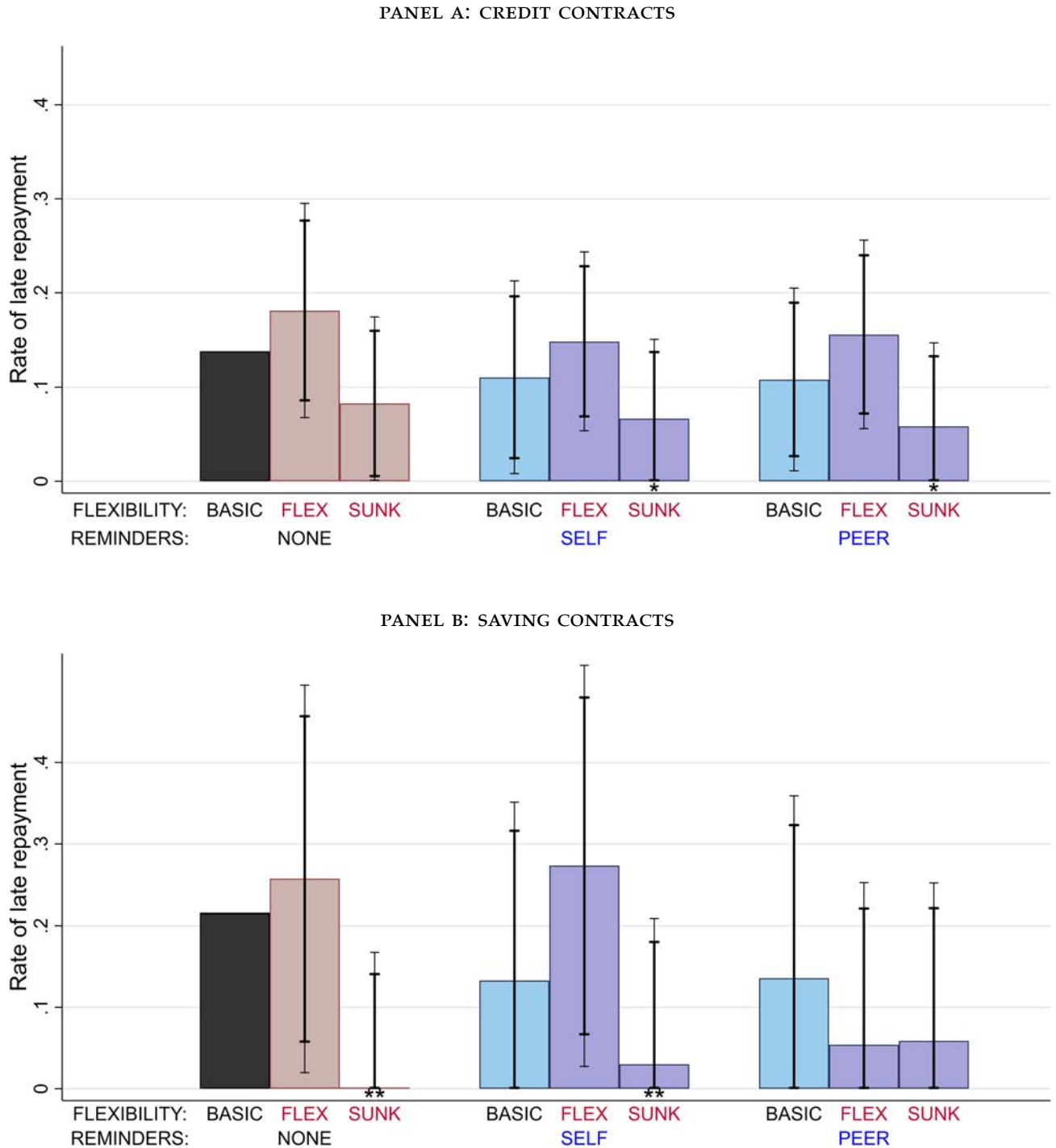
This figure shows the Group Average Treatment Effects, sorted by the take-up propensity estimated in the main text. In each figure, the leftmost (black) lines for each group show the average probability of take-up across all contract types; note that these leftmost lines are identical across figures (allowing for a different scaling of the vertical axis). The six lines in each group (in color) show the average take-up across the six different variations on contract payment and timing. For each category, the graphs show point estimates and 90% confidence intervals (both formed using the bootstrap methodology proposed by [Chernozhukov et al. \(2018\)](#)).

Table A8: Cluster Analysis: Description of extreme groups (all covariates): Phase 1

	20% LEAST LIKELY TO ADOPT			20% MOST LIKELY TO ADOPT			DIFF. (p)
	ESTIMATE	90% CONFIDENCE		ESTIMATE	90% CONFIDENCE		
Dummy: Age ≤ 28	0.29	0.17	0.40	0.17	0.07	0.26	0.16
Dummy: Age ≤ 34	0.47	0.35	0.60	0.46	0.34	0.59	0.52
Dummy: Age ≤ 40	0.62	0.50	0.74	0.63	0.50	0.75	0.31
Dummy: Age ≤ 48	0.82	0.72	0.92	0.82	0.72	0.92	0.58
Digit span test score	4.65	4.39	4.91	4.53	4.21	4.86	0.53
Dummy: Education to class 5	0.24	0.13	0.34	0.15	0.06	0.25	0.24
Dummy: Education to matric	0.13	0.05	0.21	0.14	0.06	0.23	0.33
Household size	5.18	4.65	5.69	6.31	5.76	6.84	0.01***
Dummy: Household head	0.19	0.10	0.30	0.08	0.01	0.15	0.10
Dummy: Literate	0.57	0.44	0.69	0.52	0.39	0.64	0.51
Dummy: Married	0.62	0.50	0.74	0.93	0.87	0.99	0.00***
Dummy: Correct on math question	0.49	0.36	0.62	0.48	0.35	0.61	0.44
Dummy: Self-employed	0.19	0.09	0.30	0.93	0.87	0.99	0.00***
Dummy: Spouse of household head	0.49	0.37	0.62	0.78	0.68	0.89	0.00***
Dummy: Currently in a savings committee	0.23	0.12	0.33	0.42	0.29	0.54	0.05**
Dummy: Has experience in a savings committee	0.34	0.22	0.46	0.70	0.59	0.82	0.00***
Monthly household consumption (z-score)	-0.25	-0.47	-0.03	0.13	-0.10	0.37	0.05*
Dummy: Has a bank account	0.13	0.04	0.21	0.19	0.09	0.29	0.35
Dummy: Currently owes family or friends	0.03	-0.00	0.08	0.48	0.36	0.61	0.00***
Dummy: Currently owes an MFI	0.03	-0.00	0.08	0.23	0.12	0.34	0.00***
Dummy: Currently owes NRSP	0.02	0.00	0.05	0.47	0.35	0.60	0.00***
Number of minutes to walk to NRSP (z-score)	-0.15	-0.37	0.06	0.14	-0.14	0.43	0.21
Dummy: Usually makes final decision on spending	0.86	0.78	0.95	0.79	0.69	0.89	0.30
Dummy: Finds it hard to save	0.75	0.63	0.85	0.53	0.40	0.65	0.02**
Dummy: Faces pressure to share	0.55	0.42	0.68	0.58	0.45	0.70	0.62
Risk aversion measure (higher is more risk-tolerant)	5.79	5.12	6.44	5.35	4.64	6.06	0.52
Patience measure (higher is more patient)	4.04	3.61	4.48	3.63	3.21	4.05	0.28
Patience measure in future frame	4.02	3.61	4.45	3.69	3.28	4.10	0.36

This table provides a cluster analysis of all the baseline covariates used for the machine learning analysis for Phase 1. Specifically, we describe the characteristics of those respondents in the ‘most affected’ and ‘least affected’ groups, defined in terms of estimated probability of adopting. We provide average characteristics, confidence intervals and a p-value on a test of equality of means (‘DIFF. (p)’) using the methodology proposed by [Chernozhukov et al. \(2018\)](#).

Figure A7: Late repayment by behavioral variations (including respondent using the 'flex' option)



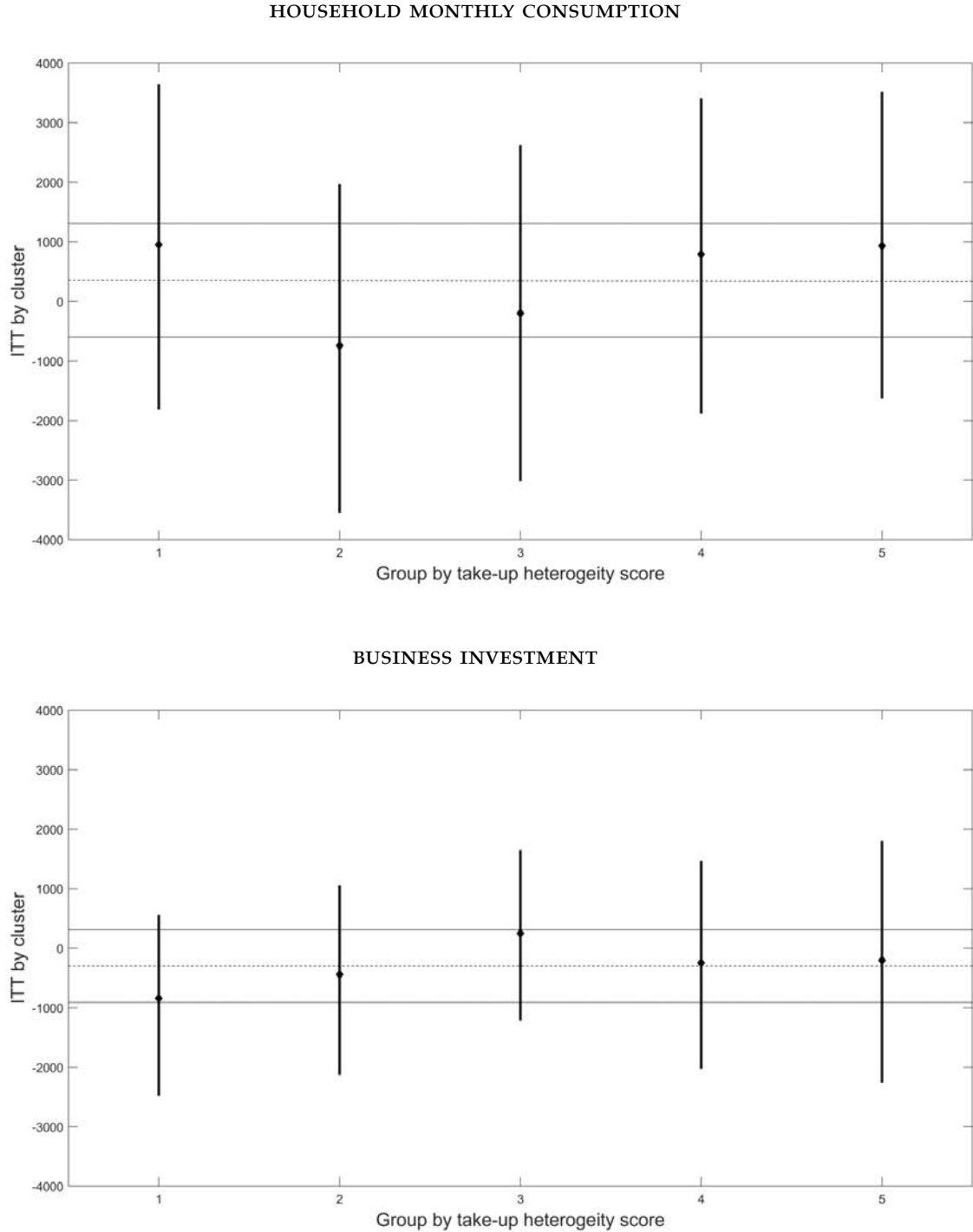
This figure shows the rate of late repayment for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'peer' variation), and for each of the eight possible variations. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic contract. Stars indicate a significant difference from the basic contract.

Table A9: Summary of ITT and LATE estimates of mobile phone data: Phase 2 experiment

	Control mean	ITT	LATE	Observations
<u>Business/employment outcomes:</u>				
Runs a business	0.118	-0.001 (0.005)	-0.006 (0.040)	9115
Number of businesses	0.144	0.002 (0.007)	-0.027 (0.056)	9115
Value of capital invested in business	26.379	5.141 (3.871)	-4.017 (26.322)	9115
Has a wage job	0.124	0.001 (0.005)	0.005 (0.042)	9115
<u>Household material outcomes:</u>				
Value of household assets	18633	-19.876 (791.485)	-246.266 (6460.983)	9115
Household monthly consumption	6737	166.658 (275.633)	-115.492 (1436.094)	9115
Total respondent debt	4147.403	-196.318 (253.633)	-559.797 (2441.105)	9115

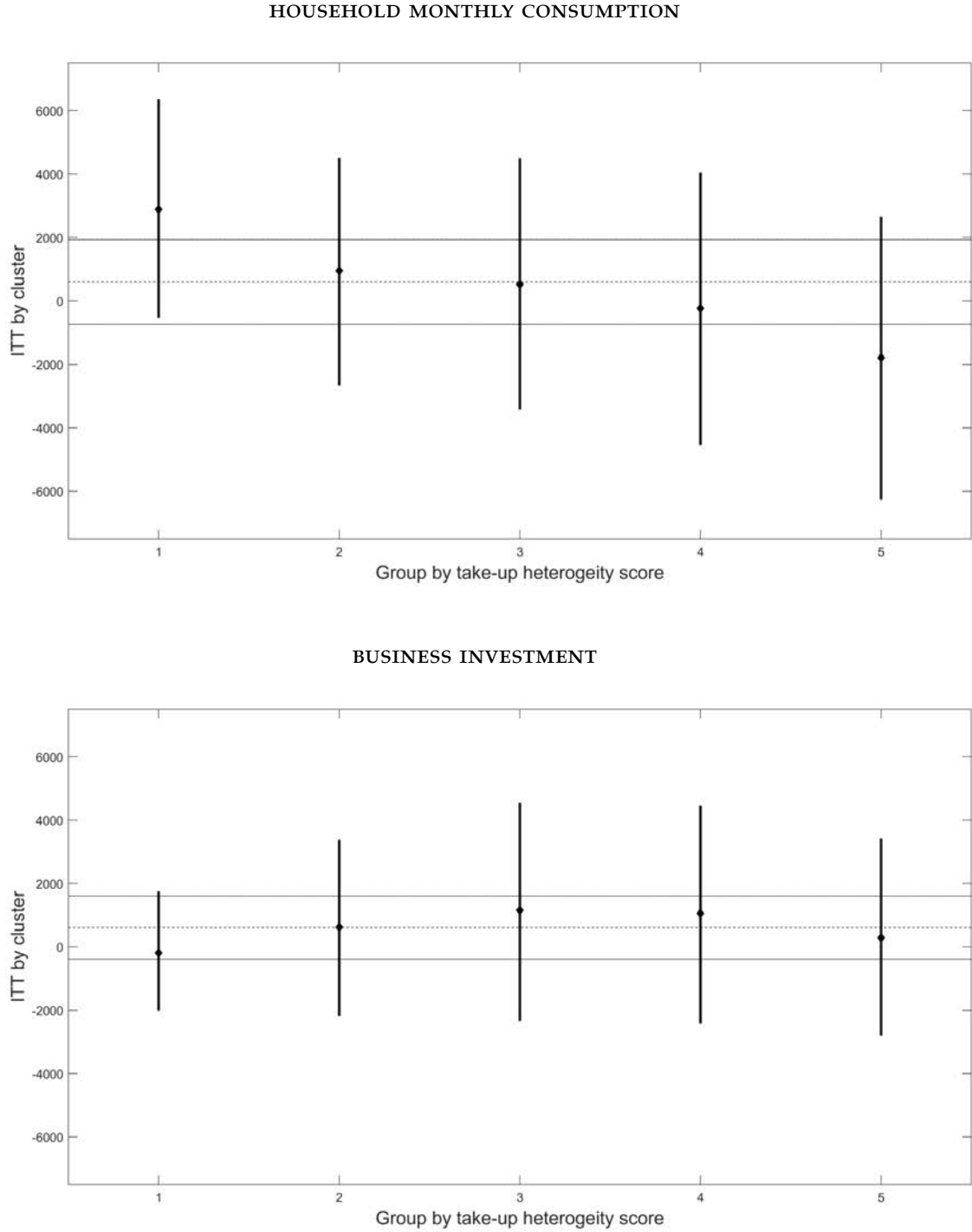
This table reports regression estimates of equation 2. We report standard errors under each coefficient in parentheses. All values are in Pakistani rupees. Confidence: * $\leftrightarrow p < 0.1$; ** $\leftrightarrow p < 0.05$; *** $\leftrightarrow p < 0.01$.

Figure A8: **Group Average Treatment Effects (sorted by take-up propensity): Consumption and investment (Phase 2)**



This figure shows the Group Average Treatment Effects, sorted by the take-up propensity estimated in the main text. For each of the five groups separately, we estimate equation 2; the graph shows the estimated ITT and 90% confidence intervals (both formed using the bootstrap methodology proposed by [Chernozhukov et al. \(2018\)](#)). The horizontal lines show the point estimates and 90% confidence intervals for the ITT across the sample (as earlier reported in Table 7 and Table 8).

Figure A9: **Group Average Treatment Effects (sorted by take-up propensity): Consumption and investment (Phase 1)**



This figure shows the Group Average Treatment Effects, sorted by the take-up propensity estimated in the main text. For each of the five groups separately, we estimate equation 2; the graph shows the estimated ITT and 90% confidence intervals (both formed using the bootstrap methodology proposed by Chernozhukov et al. (2018)). The horizontal lines show the point estimates and 90% confidence intervals for the ITT across the sample (as earlier reported in Table 7 and Table 8).