Implicit and Explicit Commitment in Credit and Saving Contracts: A Field Experiment*

Uzma Afzal
University of Nottingham†

Giovanna d'Adda
University of Milan‡

Marcel Fafchamps
Stanford University§

Simon Quinn
University of Oxford¶

Farah Said
Lahore School of Economics‖

September 2020

Abstract

We conduct a field experiment to test the demand for flexibility and for soft and hard commitment among microfinance clients. 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 instalment), and flexibility (an option to postpone an instalment). We find substantial demand for both credit and savings contracts but no demand for additional commitment features; this shows that demand for commitment depends on whether commitment features are implicit or explicit.

*We thank Josh Dean, Esther Duflo, 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 (for the first phase) and by the Economic and Social Research Council (UK) (for the second phase). We thank Mahreen Mahmud and Muhammed Meki, both for excellent field assistance and for valuable comments. Our pre-analysis plans are available at https://www.socialscienceregistry.org/trials/684 and https://www.socialscienceregistry.org/trials/1916. IRB approval was obtained from the University of Oxford (ECONCIA13-038, 9 September 2014) and from the Lahore School of Economics (RERC-042015-02, 7 May 2015).

†**School of Economics**: Uzma.Afzal@nottingham.ac.uk.

‡**Department of Economics, Management and Quantitative Methods**: giovanna.dadda@unimi.it.

§**Freeman Spogli Institute for International Studies**: fafchamp@stanford.edu.

¶**Department of Economics and Centre for the Study of African Economies**: simon.quinn@economics.ox.ac.uk.

‖**Department of Economics and Centre for Research in Economics and Business**: farahs@lahoreschool.edu.pk.

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 as impediments to the repayment of loans (Casaburi and Macchiavello, 2019; Dupas and Robinson, 2013b; Duflo et al., 2011). So too are problems of inattention, which have been identified as an alternative psychological factor behind saving behavior (Karlan et al., 2016). Indeed, several key 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 and default options – relying either on psychological or economic motivations– have proven effective in encouraging savings and reducing loan defaults (Ashraf et al., 2006; Dupas and Robinson, 2013b; Stango and Zinman, 2014; Karlan et al., 2016; Brune et al., 2016; Somville and Vandewalle, 2018).

The evidence on commitment and inattention 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 from 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 term these variations – that is, random variations in the timing and the size of the lumpsum payment – as '*contractual variations*'.

We then augment this standard product with a set of '*contractual add-ons*', designed to isolate each of the main barriers to saving and timely debt repayment identified in the literature: self-commitment issues, inattention, and intra-household dynamics. We address each barrier through the major tools found to be effective by studies on saving and borrowing. We vary the level of commitment built into the contract in two ways. On one hand, we increase commitment through the introduction of a penalty for missing an instalment (e.g., John (2020)); on the other hand, we decrease commitment by allowing higher flexibility in the repayment schedule through the possibility of deferring one instalment (e.g., Field et al. (2013)). We vary the salience of the repayment schedule by sending reminders to the clients. Finally, we vary the extent to which household members can influence clients' repayment efforts by targeting reminders to family members.

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, on whether it is increased by explicit contractual add-ons – through variation in commitment and flexibility, through assistance in keeping track, and through exogenous variation in exposure to intrahousehold pressure. Crucially, we can test whether demand for add-ons, added by previous studies either to saving or to credit contracts (Bryan et al., 2010), varies when the same feature is added to a credit or 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 our understanding of the behavioral foundations of microfinance (Bauer et al., 2012), and help to guide future development of next-generation microfinance products informed by behavioral insights.

Our results show substantial demand for credit contracts and, to a lower extent, saving contracts. Take-up rates are on average about 30% for our credit contracts, and about 8% for our saving contracts. This is in line with demand for similar commitment saving products (Cole et al., 2011), although somewhat lower than the average take-up rate of financial products with commimtent features found in the literature (Karlan et al., 2014).

When offered both a credit and a saving contract, 46 percent of the participants who take up at least one offer 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. To shed light on these apparently contradictory behaviors, we develop a behavioral model of lumpsum accumulation from which we to derive testable hypotheses tailored to our experimental design. The version of the model that best accounts for the combined evidence is one in which subjects cannot hold onto cash. This creates a demand for commitment contracts, either in the form of credit or savings, when a lumpsum expenditure is needed. When the need for a lumpsum is occasional and sometimes unforseen, demand for commitment does not manifest itself all the time. This seems to explain the observed variation in take-up across cycles by the same subject. Furthermore, when the financial need is unanticipated and urgent, it favors takeup of the credit contract. Only when it is is not immediate does it trigger also a demand for savings commitment contracts. A decion-maker who can hold onto cash would deal with this situation by accumulating precautionary saving – and would not take up some of our less appealing products. But this is not true for our subjects, many of whom seem unable to accumulate on their own and thus are willing to take unappealing contracts when the need arises.

Given this, it is perhaps unsurprising that we find no demand for contractual add-ons – such as flexibility and reminders. This is particularly in evidence for credit contracts, where take-up falls when explicit commitment features are imposed. In the saving domain, take-up does not increase with contract add-ons, but we cannot reject that it is the same as without the added features. These findings complement the existing literature: even though explicit commitment features may be effective at increasing compliance with contract terms, our results indicate that they are not valued by clients. The commitment that is implicit in a credit or savings contract seems sufficient for the needs of our subject population.

To further our understanding of take-up motives, we characterize heterogeneity in the demand for our product. To this effect we adapt the machine learning method recently proposed by Chernozhukov et al. (2018). We find significant demand heterogeneity across respondents, and show that those most likely to adopt our products are, on a variety of measures, more likely to have reported difficulties in saving, facing pressure to share their wealth, and beliefs 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 several distinct contributions. 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' – or more precisely to accumulate a lumpsum – thereby demonstrating a demand for implicit commitment. In previous pilot work, we showed this pattern of behavior for small financial products with daily repayments (Afzal et al., 2018). This paper substantially extends that result in several key ways. First, in the earlier paper, we focussed on daily repayment contracts only – a contract quite unlike standard microfinance products offered on the market – and we were left to speculate as to the generalisability of our result. To the best of our knowledge, our current experiment is the first to randomly offer 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.[1] Second, the earlier experiment did not involve any cross-cutting contractual add-ons. The fact that individ-

[1] 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 literature already cited on borrowing to save among microfinance clients (Morduch, 2010; Collins et al., 2009; Armendáriz and Morduch, 2010). 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.

uals' take-up patterns appear to be driven by behavioral motivations raises two questions: what is the relative role of the main behavioral barriers identified by the literature on credit and saving? And how is demand for these products affected by features that are added to standard microfinance contract to explicitly address these barriers? Third, the earlier work said almost nothing about heterogeneity, whereas the current paper uses recent machine learning techniques to provide a rich characterisation of heterogeneity patterns. Finally, our earlier work did not involve any follow-up interviews – and was therefore unable to provide any estimates of the consequences of being offered this kind of rotating product. This shortcoming is corrected here.

More generally, this is the first time that demand for multiple types of commitment devices have been tested within the same microfinance field experiment. This is important, because it allows us to quantify the relative effect of different contractual features, in the context of a common underlying contract and a common sample.[2] Further, our cross-cut treatments allow us to test the relevance of these add-ons for both 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; Allcott et al., 2020) 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 add-ons), 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, 2020; Bai et al., 2017), which drives firms to make commitment features implicit.

[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 (see, for example, DellaVigna and Pope (2017), (Allcott and Greenstone, 2017) and (Bertrand et al., 2010)).

[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., 2019; 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 formal flexibility, added penalty for cancellation, reminders, or intra-household pressure – also complements a recent literature documenting the hidden welfare cost of nudges. Recent empirical research shows that, while nudges can encourage intended behavior, they also increase avoidance behavior (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.

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 balance is positive, always refuse a savings contract if the net cash balance 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

[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).

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.

This result then begs other questions that our experiment addresses. First, *if microfinance demand is driven by behavioral foundations, what is the relative role of the main drivers of financial decisions identified by the literature on microfinance and saving by the poor?* Different streams of literature show the importance of commitment problems, inattention and intra-household pressures in determining both take-up of financial products and their impact on saving and income. Typically, this is done by testing the impact of a small set of contractual add-ons: hard and soft commitment features imposing or reducing the cost of missing instalments to test for self-control issues; individual reminders to test for inattention and lack of salience of the cost and benefits of saving; and information disclosure of financial obligations or decisions to family members to test for the enabling or constraining role of intra-household pressures. We combine for the first time, to the best of our knowledge, examples of all these add-ons within the same experimental design, to evaluate the relative importance of these factors.

Second, how does their influence differ between the credit and saving domains? We address this question by embedding these contract add-ons in our basic product. Before add-ons are disclosed to participants before take-up, we are able to test specifically how demand varies with explicit commitment devices, individual reminders, and peer pressure from other members of the household that are added to the implicit commitment built within basic credit or saving contracts. The fact that we find positive demand for the basic product does not guarantee that it sets the right level of commitment for the customer, nor that explicit contractual features addressing self-control or other barriers to saving may be valued when added to implicit commitment. Our design allows us to answer these questions.

[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.

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; we randomly draw different contractual variations in each cycle. 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 week (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), Karlan et al. (2016) 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 $M = PKR$ 1000 and we let the lumpsum payment take three possible values: PKR 5000, PKR 4500 or PKR 5500. Since participants pay $N - 1 = 5$ instalments of PKR 1000 each, a lumpsum of PKR 5000 simply returns the five instalments to the subject. A lumpsum of 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 a net balance of -10%.

Since there are three possible lumpsum values and two possible disbursement weeks, there are six possible contractual variations. 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 provide a positive net balance: credit is subsidized. Similarly, some saving contracts yield a negative net balance: 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 Contractual add-ons: Flexibility and reminders

Beside traditional explanations relying on credit constraints, transaction costs and lack of financial literacy, the recent behavioral literature on saving and borrowing by the poor has focused on three main explanations for undersaving and take-up of microfinance products: commitment issues, inattention and intra-household dynamics. Existing evidence on these barriers typically tests one, or at most two of these factors within the same design, or relies on the analysis of heterogeneous treatment effects to assess their empirical relevance (Dupas and Robinson, 2013a; Karlan et al., 2016). Moreover, the existing evidence focuses either on saving or borrowing behavior.

We contribute to this literature in two key ways: first, we augment our basic product with a set of contractual add-ons, specifically aimed at isolating each of the three barriers; second, we apply these treatments both to the credit and to the saving domain within the same experiment. Our design allows us to understand the relative role of commitment problems, inattention and intra-household pressures on take-up of financial products; it then allows us to test how their effects change between the credit and saving domains. Our design also allows us to make other contributions to the existing literature. By increasing or reducing the level of commitment, relative to that implicit in the basic product – through additional penalties for defaulting or flexibility in the repayment schedule, respectively – we can identify the level of commitment preferred by clients in our sample. Futher, the combination of the basic contract with individual reminders allows us to examine the interaction between self-commitment and inattention (Karlan et al., 2016). Similarly, we can study the interaction between self-commitment and intra-household pressures through the combination of the basic product and reminders to family members. Finally, the fact that our design not only includes treatments adding explicit commitment or reminders, but also a novel form of contractual flexibility, means that we combine for the first time the broad literature on commitment features with the new emerging literature on flexibility in microfinance contracts.

2.2.1 Reminders

Reminders are a common tool studied in the behavioral literature on savings. Their purpose is to increase the salience of saving goals or payment obligations and of the benefits from meeting them, and through this to help participants follow a 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 'family 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

[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 family 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 were instructed that, for reminder calls to respondents in the 'family reminder' treatment, it was not permissible to leave reminder messages with any other person who might answer the call.

investigate whether subjects are more willing to accept a financial contract that includes reminders.

Respondent reminders and family reminders are aimed at isolating the impact of two different sources of saving or repayment issues: inattention and intra-household pressures. Inability to meet financial obligations may derive from their lack of salience: in such instances, personal reminders increase commitment attainment (Karlan et al., 2016). The influence of household members on financial discipline can be positive or negative: peer pressure and demands from household members to share available resources may limit individual ability to save or repay a loan (Ashraf, 2009; Jakiela and Ozier, 2015), but household members may also act as 'saving monitors', to help the respondent to stick to the savings commitment (Breza and Chandrasekhar, 2019). By comparing the impact of reminders on take-up and repayment when sent to the respondent or to a household member, we are able to assess the direction and 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 allowing for 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 (2020)). 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.

[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.'

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, 2020) 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., 2019).[8] We combine these features in an unique experimental design and test their impact on take-up and default both in the saving and credit domains.

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.

Participants are drawn from clients of microfinance products offered by NRSP. 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.

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, family reminders and no reminders. As showed in that table, Phase 2 respondent faced a single form of treatment throughout the experiment (either the basic contract or one of the contractual add-ons), and this was known in advance of the wave 1 take-up decision. Respondents then faced random wave-to-wave variation in contractual terms (that is, the timing of the lumpsum and the contractual balance); this was known in advance of the take-up decision for each wave. We explain the specific details of randomization in Online Appendix A.

Tables A14 and A15 describe main characteristics of the sample in the two phases. Monthly household consumption averages PKR 25,000 (at the time, equivalent to about \$250) in Phase 1 and PKR 20,000 (approximately \$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 A14 and A15 also report p-values for randomization

[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.

balance across treatments. This is done by regressing each variable on the assigned treatment status in a saturated specification. We also test for randomization 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

3.1 Conceptual framework

Our empirical analysis – both of product demand and consequences of adopting – follows two Pre-Analysis Plans.[9] To frame that empirical analysis, we begin by outlining three stylised theoretical benchmark scenarios: first, when subjects can hold onto cash; second, when they cannot hold onto cash but have no need for bunching their expenditures into a lumpsum; and third, when they cannot hold onto cash and have a demand for lumpsum accumulation.

3.1.1 Scenario 1: Subjects can hold onto cash

We first describe what happens when subjects are able to save in a liquid asset. Given the context and time horizon of the experiment, this asset can be understood as money. When people can hold money on their own, a simple arbitrage argument implies that taking up a saving contract with a negative return can never be optimal. This is because the decision-maker can exactly replicate the cash flow pattern of this contract by saving all the installments and spending the lumpsum at the end – and be left with a positive net balance. For instance, instead of taking a savings contract with a payout of PKR 4500 in exchange for five installments of PKR 1000, the subject could simply set aside the installments each week and end up with PKR 5000 instead: a strategy that dominates the contract. Thus for such a saving contract to be taken up, the subject must not be capable of holding onto cash – for instance, because of self-commitment issues. The same reasoning applies to savings contracts with a zero return: subjects who able to save on their own can mimic the contract without incurring the time cost of visiting the MFI to pay each installment. It follows that take-up of saving contracts with a zero or negative return represent a lower bound on the demand for commitment: they are the least appealing commitment contracts.

[9] Our pre-analysis plan for Phase 1 (filed on 10 May 2015) is available at https://www.socialscienceregistry.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.socialscienceregistry.org/trials/1916, with extensive implementation available at http://www.simonrquinn.com/research/Microfinance_PreAnalysis_Phase2.pdf.

By a similar arbitrage argument, credit contracts in which the lumpsum exceeds the value of the installments should always be accepted by subjects who can hold onto cash. This is because a subject who can hold onto cash can use the loan to pay the installments and keep the difference. For instance, the subject could take an upfront payment of *PKR* 5500, repay the five installments of *PKR* 1000, and be left with *PKR* 500. Hence take-up is the optimal decision, provided the cost of visiting the MFI to pay the installments is small enough relative to the surplus. This kind of loan may nonetheless be rejected by subjects who have difficulties holding cash – i.e., subjects who are sophisticated about their self-commitment problems but for whom the credit contract is not a sufficiently strong external commitment. Take-up of these loans therefore represent an upper bound on the demand for commitment: they are the most attractive commitment contract. In particular, such loans will not be taken by subjects who cannot hold onto cash but have no regular income to service the debt or no need for lumpsum accumulation. We also note that subjects who can hold onto cash but refuse subsidized credit only because of transaction costs should refuse all other contracts as well since, by construction, these contracts are less advantageous – they either do not include the subsidy, or pay the lumpsum later, or both.

3.1.2 Scenario 2: Subjects cannot hold onto cash

We now examine predicted take-up for subjects who *cannot* hold to cash (and for whom the contracts we offer are the *only* available way of moving funds across periods). To do this, we use a standard framework of expected utility with exponential discounting (with weekly discount parameter β). As before, N is the duration of the contract, L is the lumpsum, and M is the installment. We denote weekly income as y, and assume that y is drawn from a stationary distribution. This framework implies the following utility-maximising behaviour:

(i) Take a borrowing contract if and only if:

$$\sum_{t=1}^{N} \beta^t \cdot \mathbb{E}\left[U(y)\right] \leq \beta \cdot \mathbb{E}\left[U(y+L)\right] + \sum_{t=2}^{N} \beta^t \cdot \mathbb{E}\left[U(y-M)\right]$$

$$\Leftrightarrow \frac{\mathbb{E}\left[U(y+L)\right] - \mathbb{E}\left[U(y)\right]}{(N-1)(\mathbb{E}\left[U(y)\right] - \mathbb{E}\left[U(y-M)\right])} \geq \frac{\sum_{t=2}^{N} \beta^t}{(N-1)\beta} = \frac{\beta - \beta^N}{(N-1)(1-\beta)}. \tag{1}$$

(ii) Take a savings contract if and only if:

$$\sum_{t=1}^{N} \beta^t \cdot \mathbb{E}\left[U(y)\right] \leq \sum_{t=1}^{N-1} \beta^t \cdot \mathbb{E}\left[U(y-M)\right] + \beta^N \cdot \mathbb{E}\left[U(y+L)\right]$$

$$\Leftrightarrow \frac{\mathbb{E}\left[U(y+L)\right] - \mathbb{E}\left[U(y)\right]}{(N-1)(\mathbb{E}\left[U(y)\right] - \mathbb{E}\left[U(y-M)\right])} \geq \frac{\sum_{t=1}^{N-1} \beta^t}{(N-1)\beta^N} = \frac{\beta - \beta^N}{(N-1)\beta^N \cdot (1-\beta)}. \tag{2}$$

To simplify, we now impose linear utility. From this, we can therefore rewrite the decision rules in equations 1 and 2 as follows (where, for convenience, we express in terms of the ratio of payments received by the client over payments made by the client):

(i) Take a borrowing contract if and only if:

$$\frac{L}{(N-1)\cdot M} \geq \frac{\beta - \beta^N}{(N-1)\cdot(1-\beta)}. \tag{3}$$

(ii) Take a saving contract if and only if:

$$\frac{L}{(N-1)\cdot M} \geq \frac{\beta - \beta^N}{(N-1)\cdot \beta^N \cdot (1-\beta)}. \tag{4}$$

Figure 1 shows the implications. In Panel A and Panel B of that figure, we graph the indifference curves implied by equations 3 and 4, for Phase 1 and Phase 2 respectively. In each case, the horizontal axis shows variation in β, and the vertical axis shows the payout ratio of L over $(N-1)\cdot M$, that is, what the client receives divided by what she paid in total. We use a log-log space for clarity. We show on each Figure the three $L/[(N-1)\cdot M]$ values used in the experiment: 1.1, 1, and 0.9 for Phase 1; and 1.086, 1 and 0.914 for Phase 2. The downward-sloping line in the upper section of each figure shows the indiffence curve for saving; points above the line imply take-up of a saving contract with payout ratio $L/[(N-1)\cdot M]$. The upward-sloping line in the lower section of each figure is the indifference curve for borrowing: points above it imply take-up of a loan with $L/[(N-1)\cdot M]$. The graph shows that, for all $\beta < 1$, there exists values of $L/[(N-1)\cdot M]$ at which a client accepts a loan but not a saving contract. We see that in Phase 1, clients with $\beta > 0.9695$ take a saving contract with a 1.1 payout ratio (the 1.1 line is above their indifference curve) as well as loans with $L/[(N-1)\cdot M] = 1$ or higher. But they do not take loans with $L/[(N-1)\cdot M] = 0.9$ (the 0.9 line is below their indifference curve). Clients with $\beta < 0.965$ take all loans, but they do not take up any of the savings contracts. Clients with β in between only take the loans with $L/[(N-1)\cdot M]$ equal to 1 or 1.1. No subject with $\beta < 1$ takes savings contracts with payout ratios of 1 or below. All subjects take borrowing contracts with a payout ratio of 1 but none of the saving contracts with the same payout ratio.

There are two further points worth noting from the two top panels of Figure 1. First, in each case, the cutoff value of β at which a respondent is indifferent as to taking a savings contract with a 1.1 payout ratio is very close to the value of β at which a respondent is indifferent as to taking a borrowing contract with a 0.9 payout ratio.[10] This means that equations 3 and 4 predict that the proportion of subjects who *take* a loan with a *low* L is approximately the same as the proportion of subjects who *reject* a savings contract with a *high* L. Second, if subjects have stable time preferences – and thus a time-invariant β – Figure 1 generate testable consistency conditions on choices made across rounds. For instance, someone who takes a loan with payout ratio 0.9 in one rounds (and thus has a $\beta < 0.965$) should reject a savings contract with a 1.1 payout in another round.

3.1.3 Scenario 3: Subjects cannot hold onto cash but have a preference for a lumpsum

Predictions are different if subjects have a specific desire to accumulate a lumpsum. Put differently, subjects may value receiving a lumpsum L over and above its monetary value because a lumpsum allows them to purchase a durable good producing a flow of services with future discounted value larger than L, as in, for example, Besley et al. (1993).[11] Alternatively, someone may wish to spend L on a ceremonial expenditure that also generates memories and social capital, thereby raising its value above L. The effect on take-up of a preference for a lumpsum can thus be examined by multiplying the payout ratio $L/[(N-1)\cdot M]$ in equations (1) and (2) by a parameter $\theta \geq 1$.

Such a change is illustrated in Panels C and D of Figure 1 for $\theta = 1.05$. We see that setting $\theta > 1$ shifts both indifference curves down, expanding the range of values of β at which taking up some of the contracts is optimal. The shift is different for loan and savings contracts, however. The range of β's for which it is optimal to *take* a loan with a *low* L is substantially larger than the range of β's that *reject* a savings contract with a *high* L. Put differently, the proportion of subjects who take a credit contract with $L/[(N-1)\cdot M] < 1$ is substantially larger than the proportion of subjects who reject a savings contract with a positive return. Further, for a contract with a payout ratio of 1, there is now a range of β values for which individuals take both loan and savings contracts. This is empirically testable: conditional on having stable preferences, these subjects would take up both savings and credit contracts across experiment waves.

Alternatively, some subjects may have a $\theta < 1$ instead, for example because they have no particular use of a lumpsum and wish to avoid the cost of making the installments or, alternatively, because they wish to smooth their consumption.[12] In those cases, the indifference curves shift upwards and this logic reverses: only very impatient subjects take a costly loan (i.e., for which

[10] This is not a coincidence: in logs, the right-hand side of equation 3 has a slope of approximately $0.5N$ while the right-hand side of equation 4 has a slope of approximately $-0.5N$.

[11] A relevant example in the context of our experiment is bulk purchases that reduce the unit cost of, say, flour, oil, or kerosene, relative to small daily purchases.

[12] To see the latter, focus on the case where $L = (N-1)M$ and consider the take-up inequalities (1) and (2). The concavity of $U(.)$ implies that, unlike in the linear case, the utility gain from receiving a large transfer L (the numerator) is less than $N-1$ times the utility loss of making installment M (the denominator) – hence their ratio is less than 1. The effect of this on take-up can be mimicked by multiplying the lefthand side of equations 3 and 4 by $\theta < 1$.

$L/[(N-1) \cdot M] < 1$) and only very patient subjects take a savings contract with a positive return. This is illustrated in Panels E and F. If β and θ are time-invariant, these predictions can be tested by comparing subjects' take-up behavior across experiment waves.

If θ varies across rounds – for example, because of fluctuations in the utility of lumpsum accumulation or in the anticipated utility cost of installments – it might be possible to observe subjects borrowing in some rounds and saving in others. It has long been noted that liquidity constraints can distort measurement of discount factors – for example, by affecting experimental measurements of time preference (Cassidy, 2019) or by causing respondents to turn down profitable savings opportunities or to take expensive credit (Noor, 2009; Gerber and Rohde, 2015; Epper, 2017; Dean and Sautmann, 2020). In this framework, where β denotes an individual-specific and time-invariant parameter, any immediate demand for funds due to unforeseen circumstances (Frederick et al., 2002) manifests itself as a sudden and temporary increase in the demand for a lumpsum and thus in θ. We revisit this point in the empirical section when we discuss changes in take-up behavior across rounds and the motives behind the demand for lumpsums.

3.1.4 Testing strategy on take-up

Our analysis of the take-up data is organized around the above ideas. We first check whether observed choices contradict the pure arbitrage argument of scenario 1 – i.e., subjects take savings contracts with $L < (N-1) \cdot M$ or refuse credit contracts with $L \geq (N-1) \cdot M$. Such evidence indicates that some subjects find it difficult to hold onto cash, thereby justifying a closer examination of scenarios 2 and 3.

We then use the cutoff values of β's reported for each of the models in 1 to draw inference about the range of β's consistent with observed choices. Since different values of θ generate very constrasted patterns of cutoff β's, we do this separately for different θ's. We then look for the θ that generates an implied distribution of β's that makes the most sense. The maintained assumption throughout this analysis is that the distribution of time preference parameters β across subjects remains the same across the three contract cycles.

Next we extend this approach to the behavior of individual subjects across cycles. Here the maintained assumption is that each subject has a stable time preference parameter. This assumption is common to most of the literature on credit and savings. We then check whether the range of β's implied by a subject's contract choices in one cycle overlaps with the range of β's implied by the choices they make in subsequent cycles. We do this separately for different values of θ. If β's overlap over the three cycles for a particular value of θ, we conclude that the observed behavior is consistent with a data generating process with that constant θ. If, however, choices made cannot be reconciled with a single value of θ, we conclude that observed behavior implies that θ changes over time –

e.g., because the expected benefit from accumulating a lumpsum L has increased ($\mathbb{E}[U(y+L)] - \mathbb{E}[U(y)]$ is higher), or because the ability to pay the installments M varies with income and other shocks (i.e., $(N-1) \cdot (\mathbb{E}[U(y)] - \mathbb{E}[U(y-M)])$ is larger).

3.2 Average take-up

We start by documenting average take-up frequencies for the six combinations of lumpsum amount and lumpsum week offered in the two phases of experimental sessions. We omit respondents in the control groups since they were not offered the contracts. Take-up frequencies are obtained by estimating a linear probability model of the form:

$$a_{it} = \sum_{w=1}^{2} \sum_{r=1}^{3} \beta_{wr} \cdot T_{it}^{w} \cdot T_{it}^{r} + \varepsilon_{it}, \tag{5}$$

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 net balance.[13]

Table 4 shows results for all subjects, from both experimental phases.[14] We first note that take-up is positive for all six contracts in both experimental phases. This is consistent with earlier results obtained by Afzal et al. (2018) using a similar contract design but a much shorter contract duration. Comparable qualitative results are obtained when we only include subjects who were offered the basic contract. As predicted by theory, take-up responds to contract terms: demand for the product is higher when payment is in week 1 instead of week N; and demand is higher when the lumpsum is larger. These differences are all statistically significant at the 1% level.[15] We find a larger sensitivity of take-up to the size of the lumpsum for credit than for savings contracts, in line with existing

[13] In both phases, some subjects said they were not interested in any contract and, consequently, staff members did not insist that they draw out 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. We examine the characteristics of automatic refusers in both phased in Tables A1 and A2 and find that automatic refusal is less likely among subjects who are currently participating in a committee, those who have higher debt and, in the case of phase 1, those who are currently running a business.

[14] In the online appendix (Table A3) we show the same analysis excluding 'automatic refusers' – that is, respondents who refused the contract before learning the contractual terms. The proportion of automatic refusers in each cycle of each phase is reported in Table A16. Across all three cycles, automatic refusers account for 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 subjecs are never automatic refusers in phase 1 and 2, respectively. The proportion of subjects who automatically refuse in all experimental cycles is 25% and 20% in phase 1 and 2, respectively.

[15] For phase 2 we repeated the analysis for subjects who were offered the basic contract without reminders, Sunk or Flex treatments. Although take-up is on average slightly higher in the basic treatment, all other patterns are qualitatively similar to those described here, with and without refusers.

evidence (Karlan et al., 2010).

3.2.1 Arbitrage and ability to hold onto cash

Next we examine the observed take-up behaviour in light of the three stylised scenarios described earlier. We begin by considering arbitrage and the ability to hold cash across periods. First, we note that the take-up of saving commitment contracts is positive even when they have a negative return. This is true in the first phase – we observe a 2.7% take-up for contracts offering a payment of *PKR* 4500 in week 6 after five payments of *PKR* 1000 – as well as in the second phase – 4.1% of participants accept a contract that pays *PKR* 3200 in week 8 after 7 installments of *PKR* 500. In both cases, participants could have accumulated the installments themselves and end up with a surplus. We also find that 4.3 to 8.9% of subjects take a zero-balance savings contract, which they could reproduce on their own without incurring the transactions cost of a contract.[16] By the arbitrage argument presented earlier, take-up of these contracts requires that the subjects be unable to hold onto cash.

Second, as in Afzal et al. (2018), we also find that a large proportion of subjects refrain from taking a subsidized credit contract: in phase 1, more than half of the subjects (53%) refuse to receive *PKR* 5500 in week 1 in exchange for *PKR* 5000 in 5 installments of *PKR* 1000; in phase 2, 62.8% of subjects refuse *PKR* 3800 in week 1 in exchange for *PKR* 3500 in seven installments. This constitutes prima facie evidence that a large proportion of subjects either cannot hold onto cash or consider that the time cost of visiting the MFI to pay the installments exceeds the value of the subsidy. This rejects the arbitrage model (scenario 1) we discussed in the conceptual section and motivates us to examine scenarios 2 and 3 in which subjects find it difficult to hold onto cash.

3.2.2 Implied discount factors for subjects who cannot hold onto cash

We now turn to the second model scenario – namely, a situation in which subjects cannot save on their own but react to contract offers in a way consistent with a standard model of saving and borrowing with stable time preferences. We have seen in Table 4 that take-up responds to contract term, as predicted in Figure 1. However, contrary to predictions of the second scenario, take-up of the positive-balance saving contract and take-up of the negative-balance loan contract do not sum to anything close to 100 percentage points. For instance, take-up of the low-balance loan is 8.2% in phase 1 while take-up of the positive-balance savings contract is 11% – i.e., a sum of 19.2%, very far from 100%. In phase 2, the figures are 11.0% and 11.3% – a sum of just 22.3%. (Similarly, for the

[16] These proportions are not massively different from the 11% of subjects who take a high return savings contract, making it unlikely that they are driven purely by decision mistakes.

zero-balance loan contract, the sums are 30.2% + 4.3% = 34.5% in phase 1 and 26.0% + 8.9% = 34.9% in phase 2.)

To ascertain which model scenario best accounts for the data, we use the pattern of take-up to infer the distribution of the time preference parameter β under different scenarios. We do so separately for credit and savings contracts. Since contract offers are randomized across subjects within each cycle, the populations of subjects offered credit contracts and savings contracts are comparable and thus they should have the same distribution of time preference parameters β. Hence for a model scenario to fit the data, the two distributions should overlap. To implement this idea, we compute, for each of the six graphs in Figure 1, the cut-off values of β for each of the six possible contracts. We then use actual take-up to infer the proportion of subjects that are *below* these cut-off values, separately for the credit and savings contracts. Hence the implied cumulative distribution of β's should line up in a monotonic and smooth fashion. If it does not when we use the β cutoff values corresponding to a particular scenario, this means that this model scenario is rejected by the data.

The results are presented in Figure 2, combining phase 1 and phase 2 subjects. Four sets of markers are shown, corresponding to four values of θ, i.e.: $\theta = 1$ (linear model); $\theta = 1.1$ (high demand for lumpsum accumulation); and $\theta = 0.9$ and $\theta = 0.8$ (low lumpsum demand + desire for consumption smoothing). Each set of markers represents the estimated cumulative distribution of β's in the study population under a particular scenario. We immediately see that markers do not line up for $\theta = 1$ or $\theta = 1.1$ (they overlap instead) in both phases of Table 4. In contrast, the markers for $\theta = 0.9$ and 0.8 line up in a monotonic fashion. We further note that the implied distribution of β's is smoother when assuming $\theta = 0.8$. These results imply that if we impose a constant θ across cycles, this θ is inconsistent with a demand for lumpsum accumulation.

3.2.3 Consistency of behavior across rounds

To summarize our results so far, we have found evidence that a large fraction of subjects cannot hold onto cash on their own (otherwise they would arbitrage), and that, on average across cycles, demand for our contracts is dominated by the time and utility cost of making installments, i.e., by $\theta < 1$. At first glance, this seems to imply that take-up is dominated by consumption smoothing considerations and concerns about the transaction costs of meeting installment obligations. But one feature of Figure 2 casts serious doubt on this interpretation: in order the reconcile the data with a model imposing a constant θ over cycles, the distribution of the intertemporal discount factor β is forced to include a non-negligible range of values above 1, suggesting that some subjects have a strong preference for delaying consumption. This seems a priori implausible for our study population. A more likely possibility is that subjects only have an *occasional* demand for lumpsum accumulation. This would tally with the large proportion of subjects who refuse the contact before even seeing the contract terms, but do so only in some cycles, not all.

To test this idea formally, we investigate whether subjects vary their take-up decisions across contract cycles in ways that agree with having a single θ. To this effect, we can use a key model prediction illustrated in Figure 1: if demand for lumpsum accumulation is high ($\theta > 1$), subjects can demand both credit and savings contracts; on the other hand, if $\theta < 1$, they demand either one or the other, never both. Hence if subjects' θ is constant over time and less than 1, they should not switch between credit and saving contracts across cycles: either they only take loans, or only saving contracts, or neither.

To investigate this idea, we report in Table 5 the proportion of time that the same subject takes both a saving and a credit contract across cycles. In the first panel of the graph, we focus on individuals who take a loan with a negative net balance – implying impatience – as well as a savings contract with a zero or negative return – implying either a strong desire to postpone consumption or a desire to accumulate for a lumpsum. After dropping automatic refusers (for whom contract terms were not drawn), we have 107 individuals in phase 1 and 350 in phase 2. In both cases, the majority of subjects rejected both contracts. Of those who take at least one loan contract with zero or positive interest (i.e., $L \leq (N-1) \cdot M$), 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 kind of behavior cannot be accommodated by a model with a constant $\theta \leq 1$: as shown in Figure 1, when $\theta \leq 1$ there is no overlap in the ranges of β's that allow taking up both borrowing and savings contract. It follows that these subjects must have, in at least one of the cycles, a desire for lumpsum accumulation, that is, $\theta > 1$.

In the second panel of Table 5, we pursue this idea by focusing instead on individuals who *accept* either of the two low payout contracts discussed in panel 1 (i.e., with $L \leq (N-1) \cdot M$) while at the same time *refusing* a subsidized credit contract with $L > (N-1) \cdot M$. There are 101 individuals in phase 1 and 399 individuals in phase 2 who are offered both types of contracts. The majority of those who are offered low payout 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 high payout credit contract. In spite of the different sample sizes, the proportions are almost identical in the two subject populations.

As we have demonstrated in Figure 1, such behavior is impossible to account for with a constant θ.[17] The only way to account for it is to assume that, when they refuse the high payout credit

[17] Note that this behavior cannot be explained by transactions costs: individuals who face transactions costs large enough to deter them from taking a high payout credit contract should *a fortiori* refuse any other contract, especially contracts with a low payout.

contract, subjects are primarily concerned with smoothing consumption ($\theta < 1$) while when they take a low payout credit or savings contact, they are more concerned with lumpsum accumulation ($\theta > 1$). In other words, their θ changes over time.

3.2.4 Demand for lumpsum accumulation

To summarize the evidence so far, the model of behavior that accounts for the evidence in the simplest way is one in which subjects cannot hold onto cash, which means they cannot accumulate a lumpsum by themselves. This creates a possible demand for a commitment contract of the type we offer in order to accumulate a lumpsum. This demand can be met either through a saving commitment contract or through a credit contract. The fact that take-up is higher for credit contracts with a low or negative interest suggests that subjects prefer to receive the lumpsum early. But many are also willing to take less advantageous contracts, while others are unwilling to accept even the most generous contract at least part of the time. As shown by Afzal et al. (2018), this behavior is consistent with a model in which individuals sometimes have a demand 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. This is because they have no interest in the lumpsum and would find it hard to hold onto cash at home in order to repay a subsidized loan. The fact that take-up is much higher when the lumpsum is paid early could also indicate a desire to respond to a known and immediate financial need (as discussed, for instance, by Frederick et al. (2002)) or to take advantage of a purchase opportunity, rather than a desire to accumulate a lumpsum for an expenditure that could easily be delayed.

To investigate the motives behind the demand for lumpsum accumulation, we examine respondents' description of how they used the lumpsum payments in Phase 2. The top reported categories are presented in Figure A7 (appendix). Of the top eight categories (which together cover about 80% of respondents), seven unambiguously correspond to 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 lumpsum food purchases often attract quantity discounts (see, for example, Brune and Kerwin (2019) and Attanasio and Pastorino (2020)). Several of these expenditure categories may also be unforeseen and urgent (e.g., home repairs, medical expenses) or driven by anticipation of an income shortfall (e.g., food), as in Frederick et al. (2002).

Taken altogether, our findings suggest that the take-up of our contracts is best explained by an occasional – and sometimes unforseen – desire to accumulate a sum of money to cover a lumpy expenditure or a sudden cash need.

3.3 Demand for contractual add-ons

We now turn to the various add-ons. 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 implicit-commitment contracts, we are interested in finding out whether demand by these same participants varies when the level of commitment built into the contract is raised or decreased by explicit contract features; or when the contract is augmented by nudges designed to address other barriers to saving, such as individual and family reminders. The literature has showed that such features reduce the probability of default. What is less clear is whether subjects have an *ex ante* demand for these additional add-ons. 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 contract.

Figure 3 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' / 'family reminder' variation); it then shows take-up rates for each of the eight possible contractual add-ons. 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 net balance. 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 A17 and A18 in the appendix.[18] Panel A of Figure 3 (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

[18] 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 family members. 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 family members, the positive effect on take-up in the 'flex' treatment is smaller in magnitude and no longer significant.

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

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 contacts. 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 family’), 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 A17). In the saving domain, demand does not appear to decrease with the added contractual features – 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.[19]

These results are indicative of the relative importance of the various barriers to saving and financial discipline identified by the literature. The fact that both increasing and reducing commitment, through the sunk and flex treatments respectively, reduces take-up, suggests that the level of commitment built into the basic product is probably right for our sample of participants. Support for the fact that inattention issues are relevant in our setting is given by the fact that increased flexibility is associated with higher demand only when coupled with reminders. This result suggests that the rigid repayment structure of the basic product also serves the function of reducing the attention costs of meeting payment obligations. It is when these cognitive costs increase due to flexibility that reminders become valuable to participants.[20] Finally, the negative effect of family reminders on demand suggests that intra-household dynamics are perceived by individuals as having a negative rather than a positive influence of repayment: especially when the cost of missing an instalment is high, in the sunk treatment, knowing that your family members will be informed of your payment obligations significantly reduce take-up of the product.

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. On the other hand, these institutions generally don’t market these commitment features – i.e., the

[19] 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 A18. 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 $p = 0.206$.

[20] This result echoes findings from the behavioral literature on planning prompts, showing that planning is valued and effective when the cognitive cost of following through is higher (Rogers et al., 2015); see also Bonan et al. (2020).

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 particular contract being offered. In particular, both Barboni and Agarwal (2018) and Battaglia et al. (2019) 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).

3.4 Robustness

3.4.1 Respondent understanding

In this section, we test the robustness of these results. First, we check for respondent misunderstanding of the contract; is it possible, for example, that our earlier results are driven by respondents having been confused about the contracts being offered?

In our view, there are several reasons to have a strong prior against this interpretation. In particular, we were well aware of this issue from the outset, and took several steps to ensure that the products were well understood. First, we designed the new products around a contract that is familiar to almost all of our subjects (namely, the ROSCA). In particular, variation in the timing of lump-sum payout occurs naturally in ROSCAs; variation in interest rates similarly occurs naturally.[21] Second, we conducted the experiment in close collaboration with an established microfinance organisation, already known and trusted by our subjects – who were, at the time of the experiment, all current or recent clients of that organisation. Third, the general forms of behavioral variation that we introduced have been tested in other related field contexts, without generating evidence of substantial subject confusion; this is true in respect of reminders (see, for example, Karlan

[21] In many ROSCAs, this variation tends to occur at the time of the payout, which introduces uncertainty; in this respect, at least, our contract is actually simpler than many ROSCAs.

et al. (2016)), of repayment flexibility (see, for example, Field et al. (2013), Czura (2015), Barboni and Agarwal (2018), Battaglia et al. (2019) and Castellanos et al. (2019)) and of the sunk repayment feature (for example, see John (2020) – and, by analogy to life insurance contracts, Anagol et al. (2017)).

Empirical evidence supports this prior. First, when asked at baseline, our respondents overwhelming agreed that they were familiar with the concept of a savings committee (96% agreed in Phase 1; 92% in Phase 2). Indeed, a substantial share had participated in a committee themselves: 51% in Phase 1, and 27% in Phase 2. Second, when we asked respondents directly for reasons that they refused the product, the overwhelming majority (about 85% in Phase 1 and 75% in Phase 2) attributed this to a lack of funds on hand to pay; almost nobody blamed a lack of understanding of the product.[22] Third, we conducted an explicit 'right/wrong' test for a hypothetical contract in Phase 1; despite asking this question approximately six months after the product was initially explained (and at least six weeks after the final take-up decision had been elicited), we found that about 85% of respondents answered correctly. Similarly, at the same time, we asked respondents whether they agreed with the statement that 'I understand how the new contracts work'; about 60% agreed (or strongly agreed), while only about 20% disagreed (or strongly disagreed). We provide further details on these figures in Online Appendix B.

3.4.2 Dynamic effects

In the same appendix, we provide further analysis on the dynamics of respondent behavior. First, we disaggregate our take-up patterns by experimental wave – and show that the general take-up patterns observed in the aggregate are also observed for each experimental wave separately. Second, we test for persistence effects across waves; to do this, we regress take-up in a given wave with take-up in the previous wave (instrumenting this lagged take-up by the contract terms randomly offered in the previous wave). We find a significant causal effect of lagged take-up; respondents who take up in a given wave are about 50 percentage points more likely to take up in the following period as a result. We interpret this as a strong familiarity effect (Mehrotra et al., 2016). This is an interesting separate finding in its own right – but has no implications for our earlier estimates on sensitivity to offered contractual terms. Because we randomized the contractual offers in each wave, the offer terms are uncorrelated to lagged take-up – and, therefore, the inclusion or omission of lagged take-up does not change our regression results. We show this empirically in the same appendix. Finally, we show the effect of dropping individuals who ever defaulted (to check that our conclusions are not driven by defaulters having been progressively excluded from the experiment). We show that

[22] Of those giving reasons, about 2% provided this reason in Phase 1. In Phase 2, respondents had the option to report this in the 'other' category, but not a single respondent did so.

overall take-up patterns are unaffected by this.

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 subgroups, 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 parameters from this model to predict take-up in the main sample, for post-processing.[23] 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 4 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 4, are a direct analog to Figure 4 in Chernozhukov et al.

[23] 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).

(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%.[24]

We augment this analysis in two ways. First, in the top panel of Figure 4, we add take-up rates and confidence intervals for (i) products offering the 'flex' variation, (ii) products offering the 'sunk' variation, (iii) products offering the 'respondent reminder' variation and (iv) products offering the 'family reminders' variation.[25] 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 contractual add-ons. Second, in the bottom panel of Figure 4, 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 A20 (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 shows large and highly significant differences in respondent characteristics for almost 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

[24] 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).

[25] 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.

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 A20 appears in the appendix as Table A21. 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 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 4 in the appendix, as Figure A9. 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 add-ons and payment of installments

How do contract features affect the payment of installments? This question is important for shedding light on our earlier take-up analysis. If, for example, late payment problems are widespread, this would have implications for the practical viability of the products studied; similarly, if late payment 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, 2020).

In Figure 5, we show the rate of late payment by contractual add-ons. The structure of the figure mirrors that of the earlier figure showing take-up rates (Figure 3): we show the rate of late payment both for the basic contract and for the contractual add-ons, and we divide the analysis between credit contracts (Panel A, top) and saving contracts (Panel B, below). The figure shows late payment 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 A10 (appendix) repeats the analysis, but recording those women as making a late payment.

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 contractual add-ons – and is significantly lower in several of those cases. If we compare different basic treatments with or without reminders, we see that, in both the credit and the saving domains, late payment falls with reminders. Although not statistically significant, this result is in line with the literature. We do, however, find that when reminders are combined with the sunk or flex treatment, they statistically reduce late payment – often by quite a large margin relative to the basic contract. What is important to recall here is that, in our experiment, we see the effect of these contractual add-ons for those subjects who do accept them, thereby allowing us to see the combined effect on both take-up and repayment performance. This is important for policy makers. We also note that subjects do enjoy quite a bit of *de facto* flexibility in meeting their installment obligations – much more flexibility than what our formal and rigid flex treatment provides. Given this, it should not be a surprise that the flex treatment shows little appeal. In fact, the apparent fall in late payment associated with the flex and self reminder treatments is a mirage: when we include late payments allowed by the flex contract (see Figure A10 in appendix), we find that, if anything the flex treatment leads to *more* late payments – although the difference remains small and is never statistically significant. Given that our flex treatment does not improve take-up – showing clients don't care for it – and that it possibly increases late payments – which raises collection costs for the lender – it is quite clear that it is not a desirable feature to add to a setting where de facto flexibility is already prevalent.

Third, we note that – with the important exception of the 'sunk' contracts – the rate of late payment 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 reneging 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

Finally, we estimate the impact of treatment on business and household outcomes. We do this by exploiting the random assignment to the control group: we compare outcomes for control participants (who were not invited to take up any of our commitment contracts) with treated participants (who were). Given the relatively small size of our lumpsum – and given previous experimental results in the literature on microfinance – it would be surprising if this product were to have large effects on business or household outcomes. However, were we to find that product has large effects, this would shed a different light on our earlier explanations for product demand; for this reason, it is important to estimate these impacts.

We provide a detailed analysis in Online Appendix C. In short, we find no robust effects on business or 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 (2018a) and Meager (2018b)) and is consistent with our preferred interpretation of the demand for lumpsum. Further, we check for heterogeneity in these effects, by the quintiles of take-up rates estimated earlier. Specifically, we estimate treatment effects separately for each of those quintiles, using 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.

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 annoyed 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 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. First, we test directly whether the rotating structure of a ROSCA 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 higher for credit contracts than for commitment savings contract, a finding we attribute to the unforseen and urgent demand for lumpsum accumulation. 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 use additional 'behavioral' add-on features in the form of reminders (both for respondents and for respondents' family members), formal flexibility in installments, and a cancellation fee. Our design allows to compare how demand for these features varies between the saving and credit domain. Our findings show that all these contract add-ons 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 payment schedule built into a standard microcredit or into a savings commitment contract of the type studied here.

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*.

Allcott, H., J. Kim, D. Taubinsky, and J. Zinman (2020). Payday Lending, Self-Control, and Consumer Protection. *Working paper*.

Anagol, S., S. Cole, and S. Sarkar (2017). Understanding the Advice of Commissions-motivated Agents: Evidence from the Indian Life Insurance Market. *Review of Economics and Statistics 99*(1), 1–15.

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 (2020). Nonlinear pricing in village economies. *Econometrica 88*(1), 207–263.

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 (2019). Repayment Flexibility and Risk Taking: Experimental Evidence from Credit Contracts. *Working paper*.

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.

Bonan, J., G. d'Adda, M. Mahmud, and F. Said (2020). The Role of Flexibility and Planning in Repayment Discipline: Evidence from a Field Experiment on Pay-as-You-Go Off-Grid Electricity. *Working paper*.

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 Revue of Economics 2*(1), 671–698.

Casaburi, L. and R. Macchiavello (2019). Demand and supply of infrequent payments as a commitment device: Evidence from kenya. *American Economic Review 109*(2), 523–55.

Cassidy, R. (2019). Are the Poor So Present-Biased? *Institute for Fiscal Studies Working Paper W18/24*.

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*.

Cole, S., T. Sampson, and B. Zia (2011). Prices or Knowledge? What Drives Demand for Financial Services in Emerging Markets? *The Journal of Finance 66*(6), 1933–1967.

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.

Dean, M. and A. Sautmann (2020). Credit Constraints and the Measurement of Time Preferences. *Working paper*.

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). 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.

Epper, T. (2017). Income Expectations, Limited Liquidity, and Anomalies in Intertemporal Choice. *Working paper*.

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.

Frederick, S., G. Loewenstein, and T. O'donoghue (2002). Time Discounting and Time Preference: A Critical Review. *Journal of economic literature 40*(2), 351–401.

Gerber, A. and K. I. Rohde (2015). Eliciting Discount Functions When Baseline Consumption Changes Over Time. *Journal of Economic Behavior & Organization 116*, 56–64.

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.

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. (2020). When Commitment Fails: Evidence from a Field Experiment. *Management Science 66*(2), 503–529.

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., A. L. Ratan, and J. Zinman (2014). Savings By and For the Poor: A Research Review and Agenda. *Review of Income and Wealth 60*(1), 36–78.

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. *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). Aggreating 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*.

Mehrotra, R., V. Somville, and L. Vandewalle (2016). Increasing Trust in the Bank to Enhance Savings: Experimental Evidence from India. *Working paper*.

Morduch, J. (2010). Borrowing to Save. *Journal of Globalization and Development 1*(2).

Noor, J. (2009). Hyperbolic Discounting and the Standard Model: Eliciting Discount Functions. *Journal of Economic Theory 144*(5), 2077–2083.

Rogers, T., K. L. Milkman, L. K. John, and M. I. Norton (2015). Beyond good intentions: Prompting people to make plans improves follow-through on important tasks. *Behavioral Science & Policy 1*(2), 33–41.

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*.

Somville, V. and L. Vandewalle (2018). Saving by Default: Evidence from a Field Experiment in Rural India. *American Economic Journal: Applied Economics 10*(3), 39–66.

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	*take up*		1000	1000	1000	1000	1000
Bank pays	*decision*	4500					

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

Table 2: **Sample structure across phases and locations**

	District	Offices	Respondents
Phase 1	Bhakkar	3	418
	Chakwal	5	372
Total		8	790
Phase 2	Khushab	5	725
	Mandi Bahauddin	4	674
	Jhelum	6	296
	Rawalpindi	2	721
Total		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**

Phase 1

Basic treatment (1/2) n = 394
Control group (1/2) n = 396

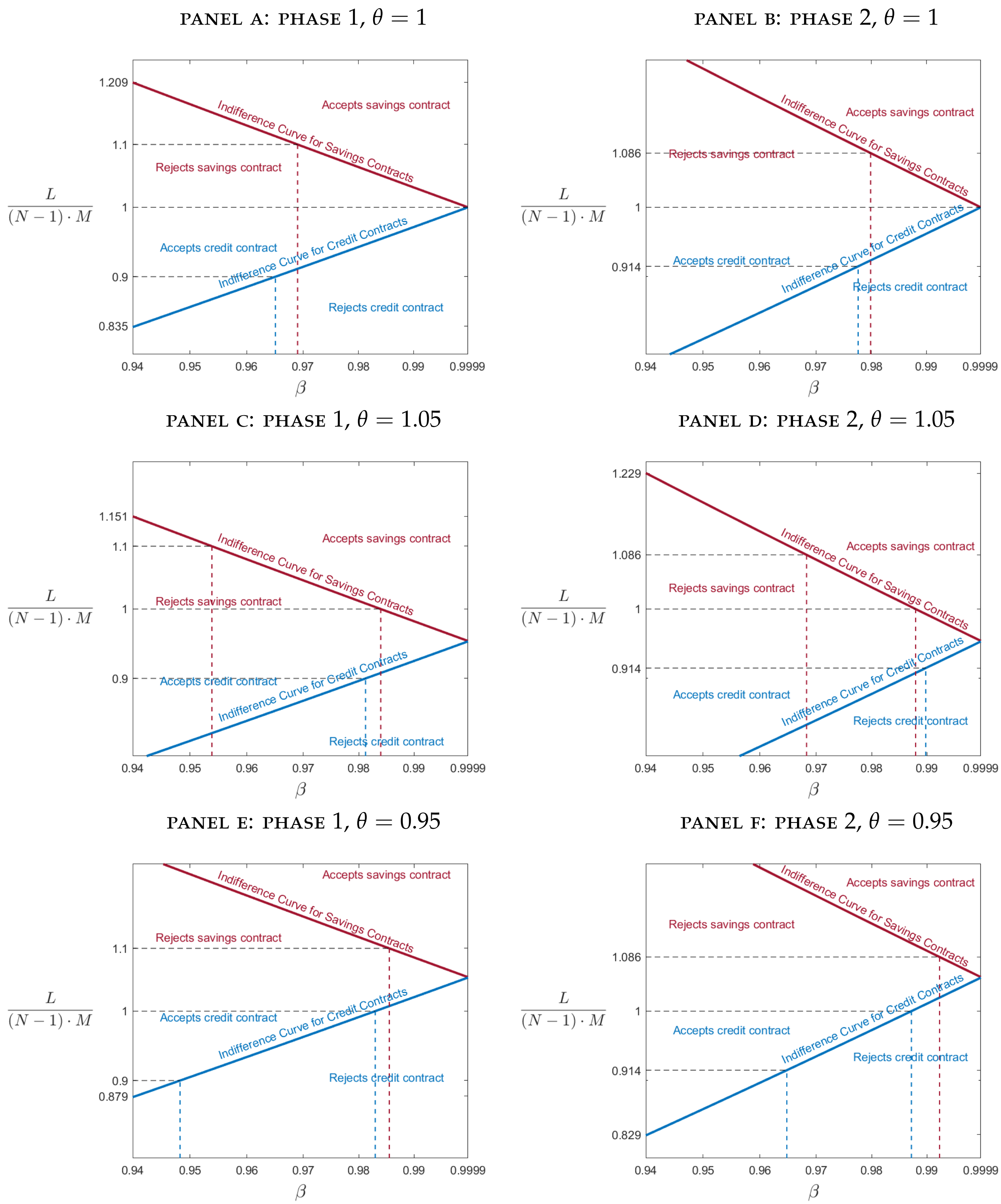
Phase 2

Basic treatment with **no reminders** (1/12) (n = 197)	**Basic treatment** with **respondent reminders** (1/12) (n = 204)	**Basic treatment** with **family 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 **family 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 **family 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.*

Figure 1: **Take-up predictions: Indifference curves**



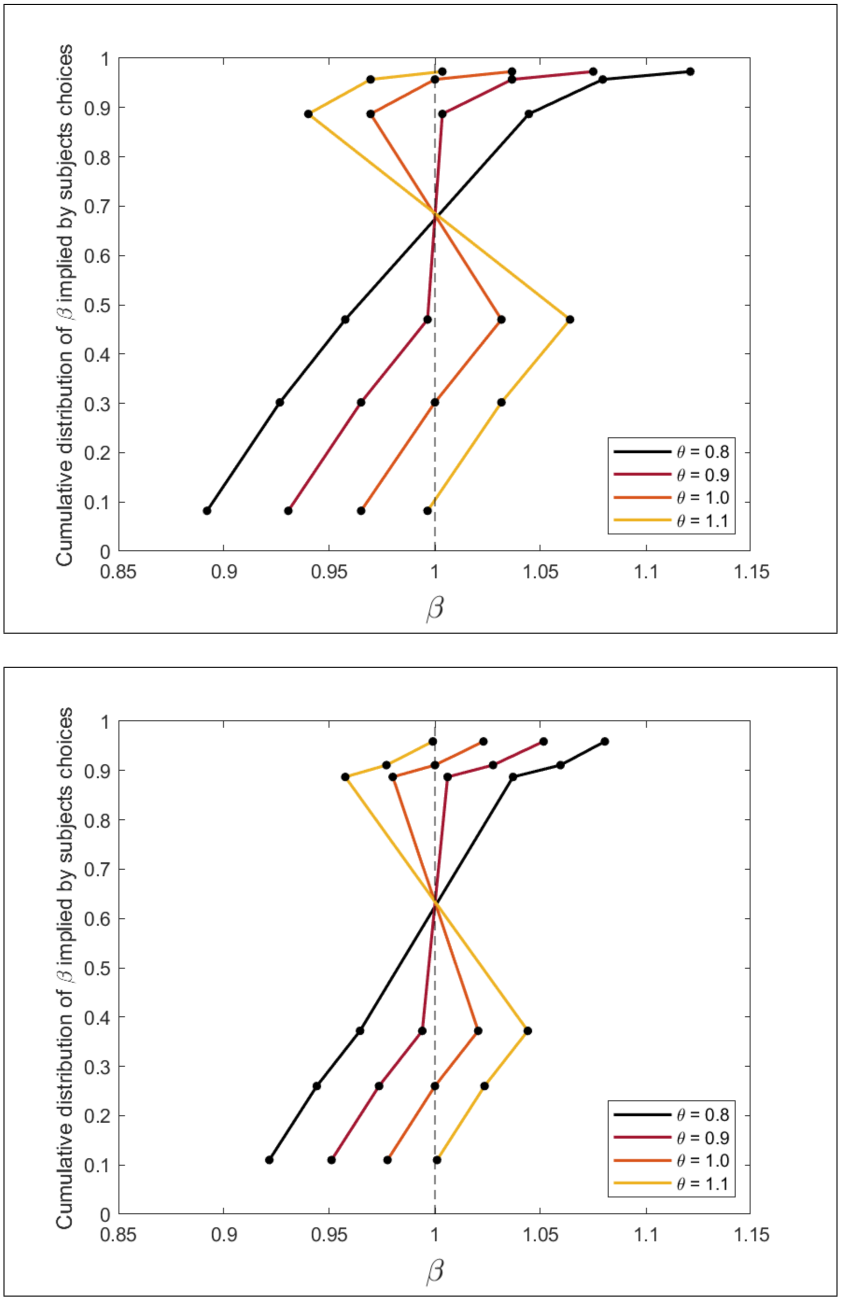
Note: Each graph shows the two indifference curves from equations 3 and 4. Graphs on the left related to Phase 1 of the experiment; graphs on the right related to Phase 2. In each case, the horizontal axis shows variation in β, and the vertical axis shows the payout ratio $L/[(N-1) \cdot M]$. We use a log transformation for β because, with that transformation, the indifference curves are approximately linear. Each graph shows the three values of $L/(N-1)M$ used in the experiment: 1.1, 1, and 0.9 for Phase 1; and 1.086, 1 and 0.914 for Phase 2. The downward-sloping line in the upper section of each graph shows the indiffence curve for saving; points above the line imply take-up of a saving contract with payout ratio $L/(N-1)M$. The upward-sloping line in the lower section of each graph is the indifference curve for borrowing: points above it imply take-up of a loan with $L/(N-1)M$.

Table 4: **Average take-up by contract terms**

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

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.

Figure 2: **Cumulative Distribution of β for different posited values of θ**



These Figures show the cumulative distribution of discount factor β that is implied by take-up choices made by subjects across all contract cycles. Four distributions are constructed based on four different assumptions regarding parameter θ, i.e.: $\theta = 1$ (linear model); $\theta = 1.1$ (desire for lumpsum accumulation); and $\theta = 0.9$ and 0.8 (concerns for consumption smoothing). The top Figure reports the findings form the Phase 1 experiment; the bottom Figure reports the findings from the Phase 2 experiment. Each point on the x-axis represents the maximum value of β for which a choice would be made (either take-up or rejection), conditional on a value of θ. All cut-off values of β are obtained as the value that equalizes the two sides of either equations 3 and 4, with the left-hand side multiplied by the relevant θ. Some implied values of β exceed 1, implying that only individuals who wish to postpone consumption would make the relevant choice, conditional on θ and the other assumptions of the model. The y-axis shows the proportion of choices that imply at most a particular β cutoff. Each line therefore represents the implied cumulative distribution of β across subjects. Normally, that implied cumulative distribution should be monotonic. Hence failure of monotonicity (and smoothness) implies rejection of a particular value of θ.

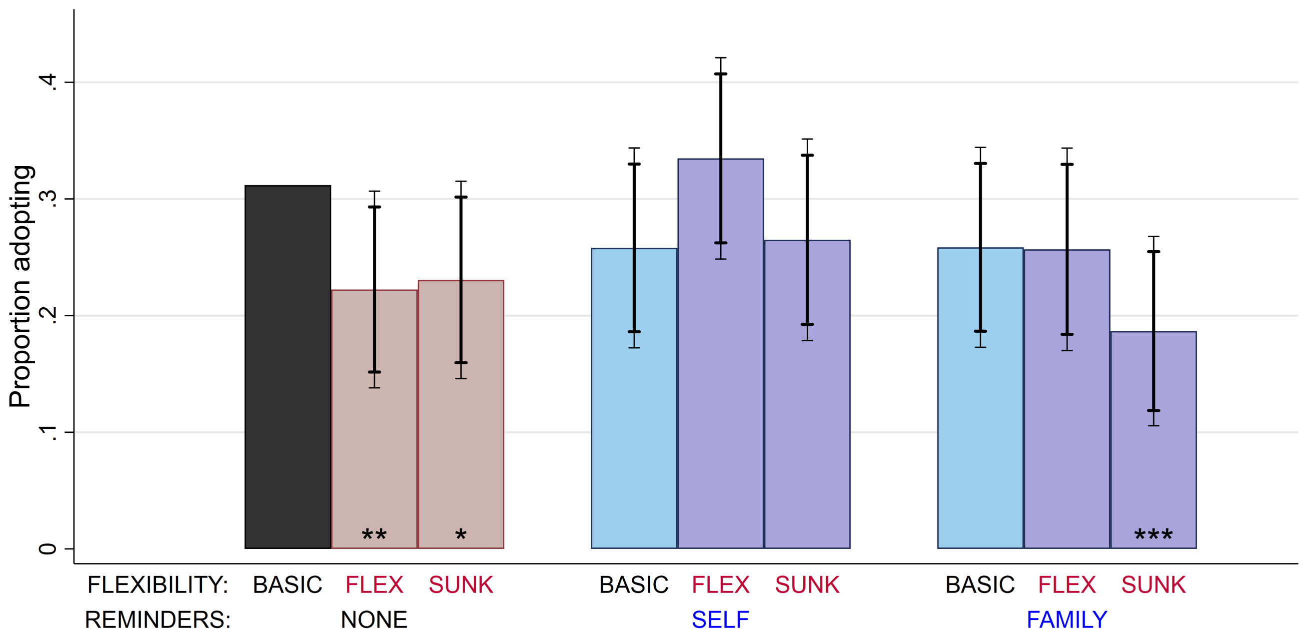
Table 5: **Consistency of Behavior across Contract Cycles**

	Phase 1	Phase 2
Subjects who were offered at least one high payout loan (i.e., with $L \geq (N-1)M$) and a low payout savings contract (i.e., with $L \leq (N-1)M$):		
took neither	76	225
took the high payout loan	30	102
took the low payout savings contract	7	47
took both	6	24
Total:	107	350
conditional on taking the loan once, the subject takes the savings contract	20%	24%
conditional on taking the savings contract once, the subject takes the loan	86%	51%
Subjects who were offered at least one high payout loan ($L > (N-1)M$) and at least one low payout contract ($L \leq (N-1)M$) – either a loan or a savings contract:		
took both	6	39
took the high payout loan	89	315
took the low payout loan or saving contract	18	123
took neither	0	0
Total:	101	399
conditional on taking a low payout contract once, refused the high payout loan	67%	68%
conditional on refusing a high payout loan once, took a low payout contract	100%	100%

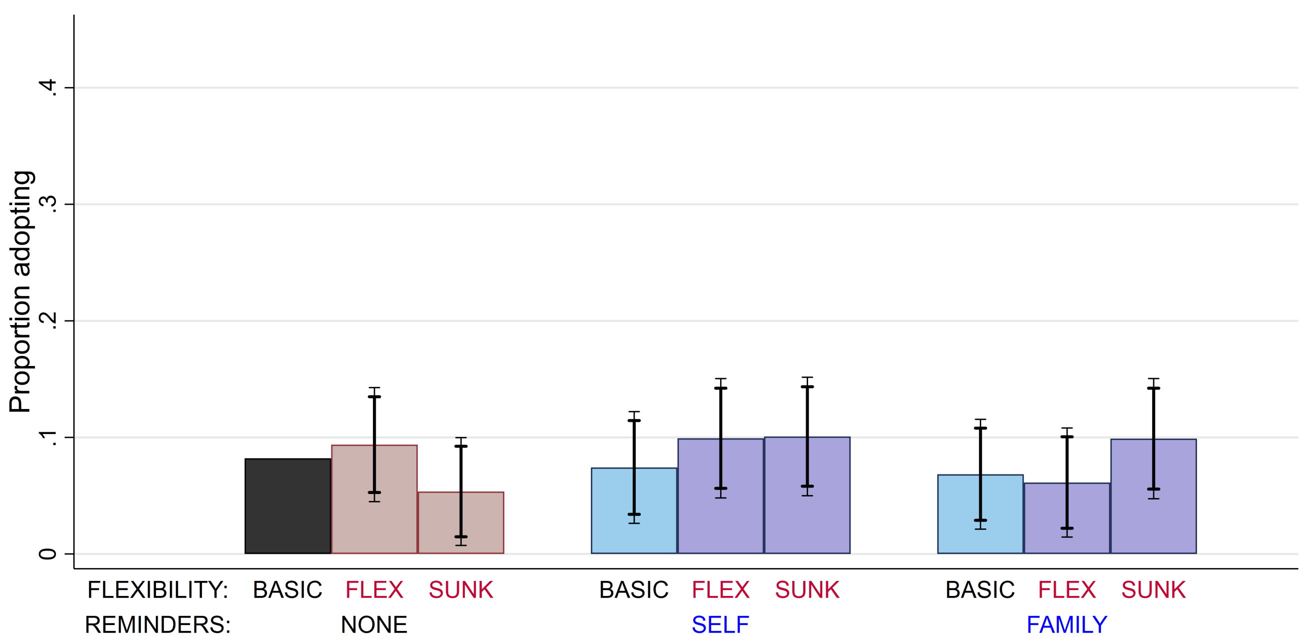
This table reports the take-up decisions made by subjects across contract cycles. Automatic refusers are omitted from these calculations.

Figure 3: **Average take-up by contractual add-ons**

PANEL A: CREDIT CONTRACTS



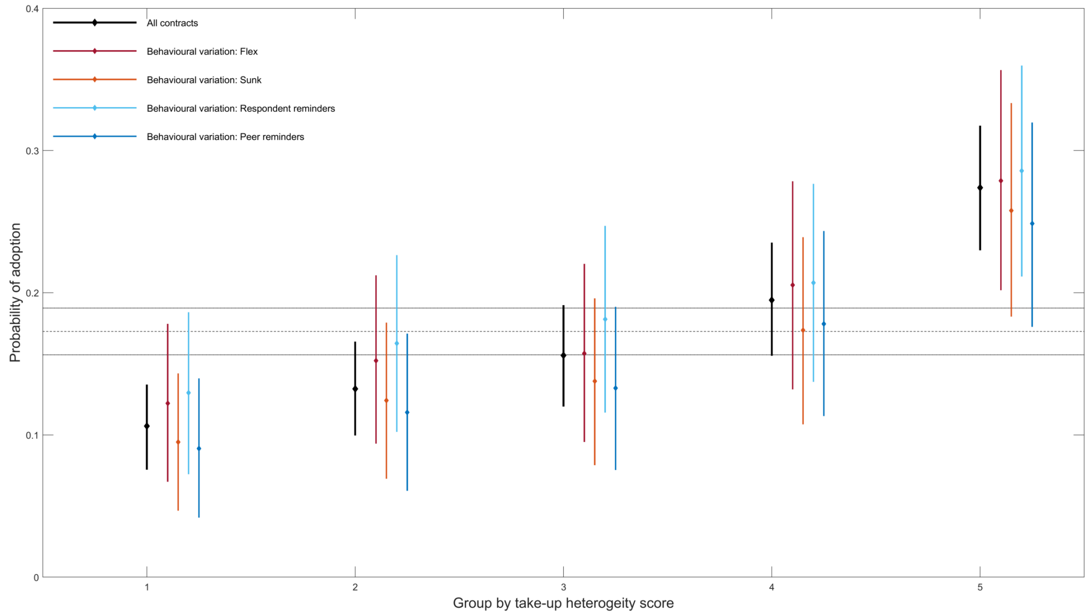
PANEL B: SAVING CONTRACTS



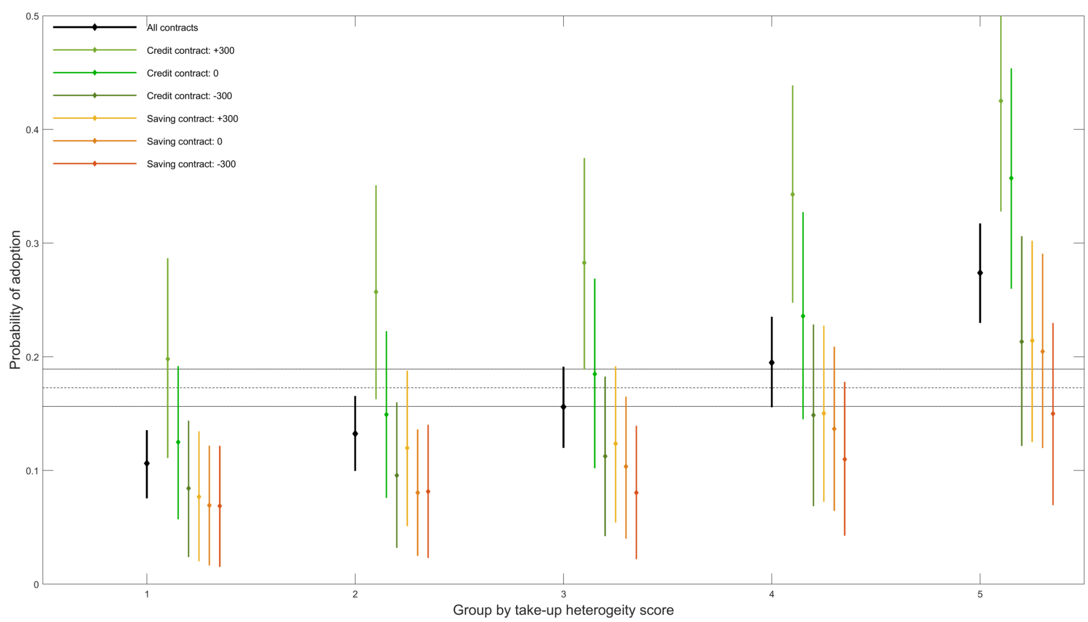
This figure shows the average take-up for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'family' variation), and take-up for each of the eight possible add-ons. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic 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 family' variation, each at the 5% significance level.

Figure 4: **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)



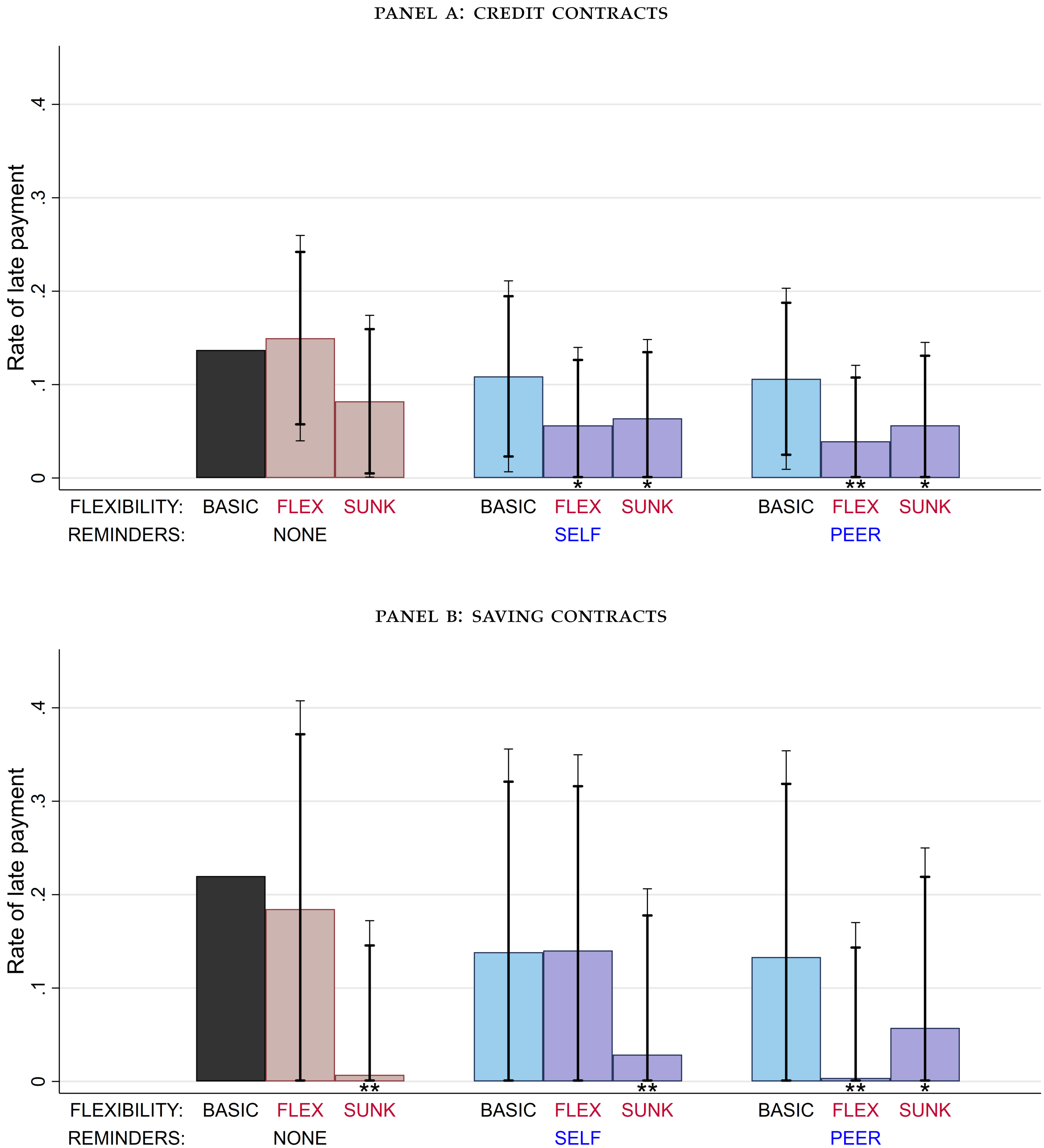
This figure shows the Group Average Treatment Effects, sorted by the take-up propensity estimated in the main text. In each figure, the leftmost (black) lines for each group show the average probability of take-up across all contract types; note that these leftmost lines are identical across figures (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 'family 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			
	ESTIMATE	90% CONFIDENCE		ESTIMATE	90% CONFIDENCE		DIFF. (p)
Saving challenges:							
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***
Keeping track:							
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***
Women empowerment:							
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 5: **Rate of late payment by contractual add-ons**



This figure shows the rate of late payment for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'family' variation), and for each of the eight possible variations. Note that we are here studying the rate of late payment; 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.