Demand for Commitment in Credit and Saving Contracts: A Field Experiment

Appendix for Online Publication

January 2023

Contents

List of Figures

List of Tables

A Theoretical framework

A.1 Demand for credit and saving contracts

To provide a theoretical frame to the empirical analysis of take-up of credit and saving products, we present three stylised benchmark scenarios: (1) when subjects can hold on to cash; (2) when they cannot hold on to cash but have no need for bunching their expenditures into a lumpsum; and (3) when they cannot hold on to cash and have a demand for lumpsum accumulation. We then draw on the insights gained from this framework to guide our empirical analysis – both our initial analysis of take-up patterns and our subsequent analysis of demand for behavioral features.

Scenario 1: Subjects can hold on to cash: We first discuss the case where subjects are able to save in a liquid asset (*i.e.* cash). When people can hold cash on their own, a simple arbitrage argument implies that taking up a saving contract with a negative return can never be optimal.[1] For take-up of a saving contract to be rational in this case, the subject must face meaningful economic costs of holding onto cash – for instance, because of self-commitment issues. The same reasoning applies to savings contracts with a zero return: subjects able to save on their own can mimic the contract without incurring the time cost of visiting the MFI to pay each instalment. It follows that take-up of saving contracts with a zero or negative return represents a lower bound on the demand for commitment: they are the least appealing commitment contracts.

By a similar arbitrage argument, credit contracts in which the lumpsum exceeds the value of the instalments should always be accepted by subjects who can hold on to cash.[2] This kind of loan may nonetheless be rejected by subjects who have difficulties holding cash – for example, 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 on to cash but have no regular income to service the debt or no need for lumpsum accumulation. We also note that subjects who can hold on to cash but refuse subsidized credit only because of transaction costs should refuse all other contracts as well since, by construction, these contracts are less

[1] This is because the decision-maker can exactly replicate the cash flow pattern of this contract by saving all the instalments 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 instalments of *PKR* 1000, the subject could simply set aside the instalments each week and end up with *PKR* 5000 – a strategy that dominates the contract.

[2] That is, a subject who can hold on to cash can use the loan to pay the instalments and keep the difference. For instance, the subject could take an upfront payment of *PKR* 5500, repay the five instalments 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 instalments is small enough relative to the surplus.

advantageous – they either do not include the subsidy, pay the lumpsum later, or both.

Scenario 2: Subjects cannot hold on to cash: We now examine predicted take-up in the stylised case of a subject who *cannot* hold cash and for whom the contracts we offer are the *only* available way of moving funds across periods. To do this, we use the standard framework of expected utility with exponential discounting and weekly discount parameter β.[3] We denote weekly income as y, and assume that y is drawn from a stationary distribution. This framework implies the following utility-maximising behaviour, comparing the net present value of the two alternative utility streams:

1. Take a credit contract if and only if the:

$$\sum_{t=1}^{N} \beta^t \cdot \mathbb{E}\left[U(y)\right] \leq \beta \cdot \mathbb{E}\left[U(y+L)\right] + \sum_{t=2}^{N} \beta^t \cdot \mathbb{E}\left[U(y-M)\right]. \tag{A1}$$

2. Take a savings contract if and only if:

$$\sum_{t=1}^{N} \beta^t \cdot \mathbb{E}\left[U(y)\right] \leq \sum_{t=1}^{N-1} \beta^t \cdot \mathbb{E}\left[U(y-M)\right] + \beta^N \cdot \mathbb{E}\left[U(y+L)\right]. \tag{A2}$$

Figure A10 shows the predictions of the model for linear utility.[4] In Panel A and Panel B of that figure we graph the indifference curves implied by equations A12 and A13 for Phase 1 and Phase 2, respectively (the remaining panels introduce variation in a parameter θ, which we introduce shortly). In each case, the horizontal axis shows the 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.[5] 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 graph shows that, for all $\beta < 1$, there exists values of $L/[(N-1) \cdot M]$ at which a client accepts a

[3] N is the duration of the contract, L is the lumpsum, and M is the instalment.

[4] Specifically, the condition in equation A12 implies:

$$\frac{\mathbb{E}\left[U(y+L)\right] - \mathbb{E}\left[U(y)\right]}{(N-1)(\mathbb{E}\left[U(y)\right] - \mathbb{E}\left[U(y-M)\right])} \geq \frac{\beta - \beta^N}{(N-1)(1-\beta)},$$

where, under linear utility, the lefthand side of this expression simplifies to $L/[(N-1) \cdot M]$. Note that the condition in equation A13 implies:

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

[5] The downward-sloping line in the upper section of each figure shows the indifference 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 $L/[(N-1) \cdot M]$.

loan contract but not a saving contract.[6] All subjects take credit 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 A10. 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 credit contract with a 0.9 payout ratio.[7] The model thus predicts 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. Put differently, with no particular demand for a lumpsum, the model predicts that the take-up rate of the positive-balance saving contract and the take-up rate of the negative-balance loan contract should sum to approximately 100 percentage points. Second, if subjects have stable time preferences – and thus a time-invariant β – Figure 1 generates testable consistency conditions on choices made across waves. For instance, someone who takes a loan with payout ratio 0.9 in one wave (and thus has a $\beta < 0.965$) should reject a savings contract with a 1.1 payout in another wave.

Scenario 3: Subjects cannot hold on to cash but have a preference for a lumpsum: Predictions are different if subjects have a specific desire to accumulate a lumpsum. Subjects may value receiving a lumpsum L above its monetary value because it allows them to purchase a durable good producing a flow of services with future discounted value larger than L – as, for example, in Brune and Kerwin (2019), Attanasio and Pastorino (2020) and Besley et al. (1993).[8] Alternatively, someone may wish to spend L on a ceremonial expenditure that also generates memories and social capital of value greater than L.

The effect on take-up of a preference for a lumpsum can be examined by multiplying the payout ratio $L/[(N-1)\cdot M]$ in equations (A12) and (A13) by a parameter $\theta \geq 1$ and redrawing Figure A10. Such a change is illustrated in Panels C and D. We see that setting $\theta > 1$ shifts both indifference curves lower. This makes intuitive sense: multiplying by $\theta \geq 1$ makes the product more desirable, and thus expands the range of values of β at which taking up some of the contracts is optimal. For θ clearly greater than 1 (for example, $\theta = 1.05$, as in Panels C and D), the range of

[6] For example, the model predicts 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 $L/[(N-1)\cdot M] = 1.1$. 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.

[7] This is not a coincidence: in logs, the expression $\frac{\beta-\beta^N}{(N-1)(1-\beta)}$ has a slope of approximately $0.5N$ while the expression $\frac{\beta-\beta^N}{(N-1)\beta^N\cdot(1-\beta)}$ has a slope of approximately $-0.5N$.

[8] 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.

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. This makes intuitive sense: multiplying by $\theta > 1$ increases the value of a lumpsum, so increases overall demand in both the credit and the savings domain. 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 instalments or, alternatively, because they wish to smooth their consumption.[9] In those cases, the indifference curves shift upwards and this logic reverses: only very impatient subjects take a costly loan (*i.e.*, for which $L/[(N-1)\cdot M] < 1$) 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 waves – for example, because of fluctuations in the utility of lumpsum accumulation or in the anticipated utility cost of instalments – it might be possible to observe subjects borrowing in some waves 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, 2021). In our 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 waves and the motives behind the demand for lumpsums.

A.2 Demand for behavioral add-ons

In this appendix section we formalize the intuition provided in Section 4 and provide a formal derivation of the take-up predictions presented in Table A1 under the reminder treatments and

[9] To see the latter, focus on the case where $L = (N-1)M$ and consider the take-up inequalities (A12) and (A13). 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 instalment M (the denominator) – hence their ratio is less than 1. The effect of this on take-up can be mimicked by multiplying the left-hand side of equations A12 and A13 by $\theta < 1$.

the sunk and flex treatments. To do this, we need to amend the model to allow for contractual breach in response to shocks. This is needed to clarify the conditions under which the flex and sunk treatments represent an improvement relative to the flexibility already present in the standard contract.

A.2.1 Increasing or reducing flexibility

Credit contracts: For credit contracts, as discussed in section 2.1, default is *de facto* not allowed by our partner MFI. Default on credit contracts is, as a result, not observed in our data. It immediately follows that we should observe no demand for credit contracts with sunk instalments: should a subject fail to spontaneously pay one of the instalments on time, the MFI would insist that the instalment be paid immediately to avoid the entire debt becoming due.[10] This means that a borrower would derive no immediate benefit from an outright default – and, in the sunk treatment, she would have to pay a separate penalty (*i.e.*, pay the first instalment a second time) for failing to spontaneously pay an instalment.

Because the penalty incentivizes spontaneous repayment, less debt collection effort is required of MFI staff. This reduces the lender's cost but it does not benefit the borrower. As a result, theory predicts that the credit contract with a sunk instalment is a weakly dominated contract for the borrower: only those who expect to *never* default would be indifferent between a standard credit contract and one with a default penalty, such as a sunk instalment. The above reasoning does not stop lenders from including default penalties into credit contracts in order to reduce debt collection costs. It is therefore important to verify empirically if, indeed, borrowers have little demand for these additional penalties.

In contrast, a flexible credit contract may be beneficial if the borrower benefits from delaying an instalment by a week, *i.e.*, if, for some week t:[11]

$$U_t(y_t) + \beta \cdot \mathbb{E}_t\left[U_{t+1}(y_{t+1} - 2M)\right] > U_t(y_t - M) + \beta \cdot \mathbb{E}_t\left[U_{t+1}(y_{t+1} - M)\right]. \tag{A3}$$

Rearranging, we see that a borrower benefits from flexibility in week t if and only if $U_t(y_t) - U_t(y_t - M) > \beta \cdot \mathbb{E}_t\left[U_{t+1}(y_{t+1} - M) - U_{t+1}(y_{t+1} - 2M)\right]$. This arises when the marginal utility of M is much larger in week t than it is expected to be in week $t+1$ – for example, because of a large negative income shock or because of an emergency that requires an urgent outlay M. Borrowers who anticipate such occurrences would express a higher demand for a flexible credit contract than for a standard one. In practice, as in standard microfinance contracts, clients are sometimes offered some *de facto* flexibility with respect to the exact date of repayment. Based

[10] This is because any debt in arrears is immediately callable – meaning that all instalments are immediately due.
[11] The notation here is the same as before, except that we have added time subscripts to income and utility.

on these observations, we expect, other things being equal, the take-up of credit contracts to (weakly) increase in the flexible treatment and to fall or, at best, remain unchanged in the sunk treatment.

Saving contracts: Things are different for savings contracts. This is because, if a subject fails to pay one instalment on time, the MFI regards the savings contract as breached and stops collecting the remaining instalments. If this occurs, in week N the MFI simply returns to the subject the sum of the instalments already collected. The behavior of the MFI therefore means that, unlike for credit contracts, default is possible. In this context, subjects who have a desire to accumulate the lumpsum L may have a demand for contractual features that reduce the likelihood of commitment failure. Whether this is the case or not depends critically on whether they see future default as desirable or not.

To capture this idea in a simple way, imagine that subjects face two types of shocks: those that they *ex ante* consider as being for a justifiable purpose (for example, taking a child to the hospital); and those they consider *ex ante* as unjustified or 'sinful' (for example, an impulse purchase). Intuitively, if the probability of a desirable breach is high and the probability of *ex ante* undesirable breach is low, at take-up the subject values the flexibility offered by the possibility of breaking the contract. This is because, in this case, the savings contract is most likely to be breached in situations that are regarded as optimal from an *ex ante* point of view – for example, because of a negative income shock or an unanticipated but welfare-enhancing expenditure. In contrast, if the probability of an undesirable breach is high relative to that of a desirable breach, a sophisticated subject will reject a flexible contract and welcome a contract that reduces the probability of breach. Based on this reasoning, the take-up of flexible savings contracts should increase if subjects believe a justifiable breach is more likely than an undesirable one – and fall if the reverse is true and they are sophisticated about it. Similarly, we expect more take-up of savings contract in the sunk treatment if subjects believe an undesirable breach is more likely than a desirable one, *and if* the sunk treatment reduces the probability of undesirable breach. We investigate this formally in Appendix A.

A.2.2 Reminders

Reminders can be thought of as a way of avoiding mistakes due to forgetfulness. For reminders to be useful, two conditions need to be satisfied: (1) limited attention: people forget things but remember when reminded; and (2) mistakes are costly – the larger the cost of missing a deadline, the more useful a reminder is. Reminders may be particularly helpful in the presence of biased neglect: people may remember pleasant information (for example, to celebrate a birthday) more easily than unpleasant information (for example, to pay an instalment). When these conditions are satisfied, a demand for reminders naturally arise in borrowing or savings contracts, irrespective of the reason for taking them up in the first place.

Reminders may be particularly helpful for certain individuals – for example, because they play a chastising role for those tempted to spend the instalment on an unjustifiable purchase. People who anticipate being tempted should be more willing to take up commitment contracts that include reminders. Since our commitment contracts are designed to appeal to sophisticate agents, reminders may increase take-up among our target population. They are not without cost, however – for example, because they are perceived as unnecessary or insulting; or because the implied threat is emotionally draining. Negative effects are more likely to dominate for people who have more financial self-discipline. It follows that, on balance, reminders should be most valued by subjects with a self-commitment problems.

Credit contracts: Applying these principles to our context, it is immediately clear that reminders serve little purpose in our standard *credit* contract since the MFI vigorously pursues instalments in arrears. As a result, reminders do not add anything of value for subjects with a self-commitment problem since the MFI shows up anyway to remind them. Subjects without self-commitment difficulties may however perceive reminders as annoying, intrusive, or intimidating because they unnecessarily foreshadow the arrival of the MFI debt collector. Following the same reasoning, we expect no additional take-up of standard credit contracts with reminders by subjects lacking self-discipline, and a possible fall in take-up by self-disciplined subjects.

Saving contracts: Reminders are more useful in the savings contract because the MFI regard being in arrears as a breach of contract. Based on this, we expect reminders to increase the take-up of *savings* contracts among those unsure they will remember to make the instalments on their own and those worried about indulging in an undesirable expenditure instead of meeting an instalment.

A.2.3 Combining reminders with variations in flexibility

Should we expect the demand for reminders – and thus for contracts with reminders – to vary across sunk and flex contracts? For credit contracts in the sunk treatment, missing an instalment may theoretically result in a penalty (the loss of the first instalment). But the diligent debt recovery behavior of the MFI *de facto* rules it out. For this reason, we do not expect reminders to increase demand for the credit contract in the sunk treatment relative to the base treatment.[12] For savings contracts, combining the sunk treatment with reminders should increase take-up –

[12] This prediction is specific to settings such as ours in which the lender is active and diligent in collecting instalments. If the lender relies instead on penalties for arrears to incentivize timely repayment, we expect a demand for reminders.

relative to sunk alone – among people who wish to avoid an undesirable breach of contract. For subjects without such difficulties, reminders may just be an annoyance that subjects prefer to avoid by not taking up the contract.

What about combining the flex treatment with reminders? We have argued that only subjects unconcerned about undesirable breach may have a demand for the flex treatment. Reminders may be seen as beneficial in this case because the flex treatment creates more risk when used. It is therefore possible that take-up of the flex treatment is higher with than without reminders among subjects with little or no repayment discipline problems. While subjects who lack self-discipline are less likely to be attracted to the flex treatment, they may nonetheless prefer it accompanied by reminders. This is particularly true for savings contracts.

Regarding reminders to family members, peer pressure adds a further encouragement to follow through with contractual obligations. For this reason, subjects with low financial self-discipline may prefer them to reminders to self. The others may instead resent reminders being sent to others and prefer to conduct their affairs in private; they may also perceive family reminders to be harmful to repayment discipline if family members make demands on clients' money. It is therefore possible that take-up with reminders may be higher among less disciplined respondents when sent to family members. Otherwise we expect reminders to family members to lower take-up relative to reminders to self.

A.3 High- and low-discipline subjects

We now offer a formal treatment of the heterogeneity of the treatment depending on the high or low level of financial discipline of the respondents. For the purpose of this exercise, we define a low-discipline subject as someone who has a non-zero probability of breaching a contract with regular instalments for a reason that is, ex ante, sub-optimal – e.g., to indulge in a frivolous or unnecessary expenditure. We call these 'unjustifiable breaches'. A high-discipline subject is someone who would only breach a contract if doing so is optimal, i.e., justifiable – e.g., to deal with a bona fide emergency. We only consider sophisticated subjects here, since naive subjects would not anticipate deviating from the contract and thus have no demand for add-ons.

We start by considering the case of (ex ante) justifiable breaches. We focus on savings contracts since, as pointed out earlier, the strict debt recovery behavior of the MFI de facto precludes default in credit contracts.

Default in the standard savings contract Let vector $\{p_t\} \equiv \{p_1, p_2, ..., p_{N-1}\}$ denote the subject's beliefs at take-up regarding the probability that they will default from the contract in each of the $N-1$ weeks in which they have to pay instalment M. Since they can only end the contract once, these probabilities are mutually exclusive – e.g., they can end the contract either in period 1 or 2, not in both. Hence they expect to fulfill the contract with probability $1-\sum_{i=1}^{N-1} p_i$. The expected value of fulfilling the standard contract can be written as in equation (A13):

$$\sum_{t=1}^{N-1} \beta^t E\left[U_t(y_t - M)|P_t = 0\right] + \beta^N E\left[U_N(y_N + L)\right], \tag{A4}$$

where the notation is as before, $P_t = \{0,1\}$ is the realization of the probability p_t, and the expectation $\mathbb{E}\left[U_t(y_t - M)|p_t = 0\right]$ is taken over values of y_t or U_t for which the subject decides not to breach the contract. By construction, the subject breaches the contract with probability p_t in week $t < N$. The expected value of taking up the contract but breaching it in week t is:

$$\sum_{s=1}^{t-1} \beta^s E\left[U_s(y_s - M)|P_s = 0\right] + \beta^t E\left[U_t(y_t)|P_t = 1\right] + \beta^N E\left[U_N(y_N + (t-1)M)\right]. \tag{A5}$$

Putting the two together yields the value of the standard savings contract when assuming only justifiable defaults:

$$\begin{aligned} B_J = {} & (1 - \sum_{i=1}^{N-1} p_i) \left(\sum_{t=1}^{N-1} \beta^t E\left[U_t(y_t - M)|P_t = 0\right] + \beta^N E\left[U_N(y_N + L)\right] \right) \\ & + \sum_{i=1}^{N-1} p_i \left(\sum_{s=1}^{t-1} \beta^s E\left[U_s(y_s - M)|P_s = 0\right] + \beta^t E\left[U_t(y_t)|P_t = 1\right] + \beta^N E\left[U_N(y_N + (t-1)M)\right] \right). \end{aligned} \tag{A6}$$

The values of each default probability is determined by backward induction. In week $N-1$, the subject defaults iff:

$$U_{N-1}(y_{N-1} - M) + \beta E[U_N(y_N + L)] < U_{N-1}(y_{N-1}) + \beta E[U_N(y_N + (N-2)M)] \tag{A7}$$

where, since subjects cannot save on their own, the unpaid instalment is consumed during the default week. The probability of default p_{N-1} is the share of all possible joint realizations of y_{N-1} and U_{N-1} for which inequality (A7) is satisfied. The probability of defaulting in week $N-2$ can be determined in a similar fashion, conditional on not having defaulted yet but taking p_{N-1} into account, etc.

Now we consider the polar opposite case with only (ex ante) unjustifiable breaches. Default probabilities are derived in the same manner,[13] but the *ex ante* value of the contract is different.

[13] I.e., we only consider sophisticated subjects here. Naive subjects do not anticipate defaulting and thus have no demand for commitment.

Since the undesirable consumption is not valued *ex ante*, the expected utility when defaulting is $E\left[U_t(y_t - M)|P_t = 1\right]$, i.e., M is not included in the utility function. It follows that, from an ex ante point of view, strict inequality (A7) is never satisfied as long as $E[U_N(y_N + L)] > E[U_N(y_N + (N-2)M)]$ and hence default is never *ex ante* optimal in week $N-1$ – and, by backward induction, in any week. The value of the contract when all defaults are undesirable can thus be written:

$$B_U = (1 - \sum_{i=1}^{N-1} p_i)\left(\sum_{t=1}^{N-1} \beta^t E\left[U_t(y_t - M)|P_t = 0\right] + \beta^N E\left[U_N(y_N + L)\right]\right)$$
$$+ \sum_{i=1}^{N-1} p_i \left(\sum_{s=1}^{t-1} \beta^s E\left[U_s(y_s - M)|P_s = 0\right] + \beta^t E\left[U_t(y_t - M)|P_t = 1\right] + E\left[U_N(y_N + (t-1)M)\right]\right). \tag{A8}$$

It immediately follows that, as intuition would suggest, $B_J > B_U$: take-up is higher among individuals who only expect to default from the savings contract for a reason that is regarded as justifiable *ex ante*.

Default in the sunk treatment In the sunk treatment, default is penalized by the loss of the first instalment. The value of this saving contract under desirable default becomes:

$$S_J = (1 - \sum_{i=1}^{N-1} p_i)\left(\sum_{t=1}^{N-1} \beta^t E\left[U_t(y_t - M)|P_t = 0\right] + \beta^N E\left[U_N(y_N + L)\right]\right)$$
$$+ \sum_{i=1}^{N-1} p_i \left(\sum_{s=1}^{t-1} \beta^s E\left[U_s(y_s - M)|P_s = 0\right] + \beta^t E\left[U_t(y_t)|P_t = 1\right] + E\left[U_N(y_N + max\{0, (t-2)\}M)\right]\right), \tag{A9}$$

where S_J denotes the value of the sunk contract to a subject who only faces justifiable detault. Since $(t-1)M > \max\{0, (t-2)\}M$ for any $t > 1$, the value of default is clearly lower, and this reduces the probability of default. But since desirable defaults are *ex ante* optimal for the decision maker, the sunk treatment implies a deviation from optimality and thus a reduction in the value of the contract. If follows that $S_J < B_J$: there will be be less take-up by subjects who anticipate defaulting on the saving contract for valid reasons. This was to be expected because they have no demand for the sunk treatment since they do not face a self-commitment problem.

The situation is different for sophisticated subjects who anticipate an undesirable default. Their ex ante value of the saving contract is:

$$S_U = (1 - \sum_{i=1}^{N-1} p_i) \left(\sum_{t=1}^{N-1} \beta^t E\left[U_t(y_t - M)|P_t = 0\right] + \beta^N E\left[U_N(y_N + L)\right] \right)$$
$$+ \sum_{i=1}^{N-1} p_i \left(\sum_{s=1}^{t-1} \beta^s E\left[U_s(y_s - M)|P_s = 0\right] + \beta^t E\left[U_t(y_t - M)|P_t = 1\right] + E\left[U_N(y_N + max\{0,(t-2)\}M)\right] \right), \tag{A10}$$

which shows a similar loss of consumption in case of default. Can this loss be compensated by a reduction in undesired defaults? The default condition for week $N-1$ is now:

$$U_{N-1}(y_{N-1} - M) + \beta E[U_N(y_N + L)] < U_{N-1}(y_{N-1}) + \beta E[U_N(y_N + (N-3)M)], \tag{A11}$$

where the decision maker anticipates losing an extra M of consumption a week later, relative to the standard contract. This reduces p_{N-1}, that is, the share of values of y_{N-1} and U_{N-1} for which inequality (A11) is satisfied. By extension, the sunk treatment should also reduce default in previous weeks – although, other things being equal, the induced reduction in default probability falls as t gets closer to the start of the contract. This is because the loss of M in week N happens further and further into the future. It follows that a sophisticated agent can use the sunk treatment as an imperfect commitment device to reduce default – but she must trade off this gain against the loss of M when default occurs. If follows that S_U can be larger or smaller than B_U depending on parameters only known to the subject – i.e., $S_U >$ or $< B_U$.

Default in the flex treatment We have shown that, for credit contracts, repayment flexibility is welcomed by borrowers who ex ante anticipate to benefit from delaying an instalment by a week. Other dislike flexibility because it makes temptation to indulge in an undesirable breach more likely. The same reasoning applies to savings contracts for subjects who only face valid shocks. Without the flexibility to delay, the subject may choose to default when a shock occurs. With the flex treatment, the subject can now 'buy back' into the contract in the following week, if conditions allow. The expected utility gain is, however, likely to be small given that the delayed instalment is due in full in the following week, increasing the utility cost in that week. We therefore expect that, for individuals who only expect to experience justifiable expenditure shocks, the flex contract is at least as good as a standard savings contract – and perhaps slightly more desirable: $F_J \geq B_J$.

The situation is different for sophisticated subjects who worry about undesirable default. For these subjects, there is, by assumption, no breach of contract that is *ex ante* justifiable. The introduction of the flex treatment therefore serves no useful purpose. Instead, it creates increased temptation for subjects to succumb to an impulse purchase in one week by convincing themselves that they will pay it back in the next. Consequently, for these subjects, the expected discounted value of a savings contract is lower with the flex option and they should have a lower demand for savings contracts in the flex treatment: $F_U < B_U$.

Summary of results All the predictions regarding high and low discipline subjects are summarized in Table A1. The different treatments are denoted by letters: B stands for the basic contract; S stands for the sunk treatment; F for the flex treatment; and R^s for reminders to self. Subscripts J and U denote subjects with high and low financial discipline, respectively.[14]

Table A1: **Model predictions for subjects without and with financial self-discipline issues**

High financial discipline (justifiable breach only)		Low financial discipline (undesirable breach only)	
Credit contract	Savings contract	Credit contract	Savings contract
$F_J \geq B_J \simeq S_J$	$F_J \geq B_J > S_J$	$F_U \leq B_U \simeq S_U$	$F_U < B_U < S_U$ or $> S_U$
$B_J > BR^s_J$	$B_J > BR^s_J$	$B_U \simeq BR^s_U$	$B_U < BR^s_U$
$S_J \geq SR^s_J$	$S_J > SR^s_J$	$S_U < SR^s_U$	$S_U < SR^s_U$
$F_J \leq FR^s_J$	$F_J \leq FR^s_J$	$F_U < FR^s_U$	$F_U < FR^s_U$
$BR^s_J \simeq SR^s_J$		$BR^s_U \simeq SR^s_U$	

B Implementation details

B.1 Randomization

The sample in both phases consists of past and current NRSP borrowers. There were no further credit-worthiness checks to determine eligibility for participating in the experiment. In both phases, we used two separate mechanisms to assign eligible respondents to treatment. First, we assigned each respondent to either the treatment or control group. Second, we randomly assigned those in the treated to group receive the lump-sum payment either in week 1 or week N; and randomly varied the interest payment (*i.e.* zero, negative or positive).

In Phase 1, we first stratified using a method of the kind described in Bruhn and McKenzie (2009). We first formed four blocks based on baseline variables measuring 'loan status', i.e, whether the respondent has a currently outstanding loan or if the loan had closed in the past 12 months, and 'whether the loan will be used for investment in the business'. We then sorted by business profit within each block and formed strata of four respondents within each block – i.e. the four respondents with the highest baseline business profits were assigned to one stratum, the four respondents with the next highest baseline profits were assigned to the next stratum, and so on.

[14] J stands for justifiable breach only and U for undesirable breach only – see Appendix A.

Within each stratum, we then randomly assigned two respondents to the treated group and two respondents to the control group, as described in Table 2. The results of this randomization was fixed over time – a respondent assigned to the treated group remained in this group throughout the duration of the experiment.

We use a similar stratification method in the first step of the randomization in Phase 2. First we assigned every respondent either to the control group, to the 'basic contract' group, or to one of eight separate contractual add-ons, as illustrated in Table 2. Specifically, we formed eight blocks based on the answers to the binary baseline variables 'running a business', 'whether the respondent makes the final decision on spending', and 'whether the respondent would use a loan for investment'. We then sorted by household income within each block, and formed strata of 12 respondents within each block – so, for example, the 12 respondents having the highest baseline household income were assigned to one stratum, the 12 respondents having the next highest baseline income were assigned to the next stratum, and so on. Within each stratum, we then randomly assigned three respondents to the control group, one respondent to the 'basic contract' group, and one respondent to each of the eight contractual add-ons described in Table 2. The results of this randomization were fixed over time – a respondent who was placed into the 'sunk treatment with respondent reminders' was informed of this fact before her wave 1 take-up decision, and remained in this variation throughout the experiment.

In both phases, NRSP field officers would visit eligible respondents at their homes or place of work and offer them the product assigned to them. In the second step of randomization in Phase 1 and 2, every respondent faced random variation in both the interest charge (*i.e.*, zero, negative, or positive) and the week of the lumpsum payment (Week 1 or Week N) at this offer. This assignment was implemented by inviting participants to draw a card at random at the beginning of each wave. The card was drawn in Week 0, at which time subjects were asked whether they take the contract or not. If they agreed to take the contract, NRSP field officers returned a week later to start the contractual implementation. The result of this randomization was not fixed over time – the card was drawn out in Week 0 at each wave. For example, a treated respondent had to draw out a card assigning week of lumpsum payment and the interest charge in week 0 of the first wave, then week 0 of the second wave and, finally, week 0 of the third and final wave in each Phase of the experiment.

B.2 Automatic refusers

In both phases of the experiment, some subjects said that they were not interested in the product. Consequently, staff members offering the contract did not ask for them to draw out any card to determine the net balance or timing of the lumpsum payment. In the analysis, we consider these subjects to have refused all six contractual terms, each of which would have been offered with 1/6 probability. The proportion of respondents who automatically refuse in each wave of phase 1 and 2 are given in Table A21.

Much of the variation in automatic refusals can be explained by staff member differences in how strictly they followed the product offer protocol, i.e. draw out cards to complete the product offer even if the subject says they are not interested in the product. We examine the variation in automatic refusal by subject characteristics for both phases of the experiment, controlling for staff member specific effects. In Phase 1, we do this by controlling for neighbourhood (or 'mohallah') effects - which is a close approximation of the staff member assigned to the area in which the subject resides; in Phase 2 we collect information on the staff member responsible for offering to each subject and control for staff member effects directly. Results are provided in Table A2 for phase 1 and Table A3 for phase 2. We find that after controlling for neighbourhood dummies in Phase 1, where we offer the product to a sample of microenterprise loan borrowers, the likelihood of automatic refusal is lower among subjects who currently own a business and among literate subjects. Subjects in both phases are less likely to automatically refuse if they are currently participating in a committee and if they have young children. We also find weak evidence to suggest that subjects are less likely to refuse if they have more debt, though this effect is economically and statistically small.

Table A2: **Describing the characteristics of automatic refusers in Phase 1**

	(1) OLS	(2) FE Logit	(3) RE Logit
Dummy: participates in a committee	-0.064 (0.033)**	-0.069 (0.042)*	-0.033 (0.031)
Total amount owed by individual (000's PKR)	-0.002 (0.001)**	-0.002 (0.001)	-0.001 (0.001)*
Total household consumption last month (000's PKR)	0.003 (0.002)*	0.002 (0.002)	0.004 (0.002)**
Total value of assets owned by household (000's PKR)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Dummy: runs a business	-0.179 (0.082)**	-0.155 (0.093)*	-0.233 (0.078)***
Total number of businesses owned by respondent or household	0.045 (0.042)	0.050 (0.045)	0.060 (0.044)
Total capital invested in respondent or household business(es) (000's PKR)	0.001 (0.001)	0.001 (0.003)	0.000 (0.001)
Total monthly sales of the business (000's PKR)	-0.004 (0.004)	-0.008 (0.006)	-0.007 (0.005)
Total monthly expense of the business (000's PKR)	0.002 (0.005)	0.003 (0.005)	0.003 (0.005)
Total monthly profit(2) of the business (000's PKR)	0.004 (0.006)	0.011 (0.010)	0.010 (0.008)
Dummy: finds it hard to save	0.014 (0.051)	0.004 (0.051)	0.025 (0.049)
Index: respondent opinions taken into account in household decisions	0.009 (0.036)	-0.000 (0.039)	0.038 (0.034)

Index: respondent needs to ask permission for making decisions	0.083	0.087	0.074
	(0.050)*	(0.064)	(0.048)
Age (years)	-0.001	-0.001	0.000
	(0.003)	(0.003)	(0.003)
Dummy: Respondent is currently married	-0.090	-0.038	-0.022
	(0.077)	(0.070)	(0.070)
Level of education	0.009	0.012	0.014
	(0.008)	(0.011)	(0.008)*
Dummy: Respondent can read and write	-0.161	-0.204	-0.163
	(0.076)**	(0.111)*	(0.078)**
Number of children	-0.024	-0.027	-0.039
	(0.014)*	(0.020)	(0.014)***
Head of the household	0.042	0.048	0.050
	(0.092)	(0.098)	(0.086)
Neighbourhood effects	yes	yes	yes
Observations	389	271	389

This table provides an analysis of automatic refusals in Phase II by subject characteristics, after controlling for neighborhood level effects. In each column, we show a regression of a subject automatically refusing the product in any wave, on individual characteristics. Specifically, we show results from an OLS regression with neighborhood dummies in column (1), a logit regression with neighbourhood fixed effects in column (2) and a logit regression with neighbourhood random effects in column (3). For the random-effect logit, we estimate $\rho = 0.20$*, with a 95% confidence interval of* $[0.08, 0.41]$*. Standard errors are clustered at the household level for the OLS regression. We use '*' to denote confidence at the 90% level.*

Table A3: **Describing the characteristics of automatic refusers in Phase 2**

	(1) OLS	(2) FE Logit	(3) RE Logit
Dummy: participates in a committee	-0.044	-0.079	-0.046
	(0.022)**	(0.040)**	(0.023)**
Total amount owed by individual (000's PKR)	-0.001	-0.001	-0.001
	(0.000)***	(0.000)***	(0.000)***
Total household consumption last month (000's PKR)	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)
Total monthly income (000's PKR)	-0.001	-0.002	-0.001
	(0.001)	(0.001)*	(0.001)*
Total value of assets owned by household (000's PKR)	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Dummy: runs a business	-0.080	-0.156	-0.093
	(0.061)	(0.103)	(0.061)
Total number of businesses owned by respondent or household	0.033	0.067	0.040
	(0.030)	(0.051)	(0.030)
Total capital invested in respondent or household business(es) (000's PKR)	0.000	0.000	0.000
	(0.000)	(0.001)	(0.000)
Total monthly sales of the business (000's PKR)	-0.003	-0.003	-0.002
	(0.005)	(0.009)	(0.005)
Total monthly expense of the business (000's PKR)	-0.001	-0.002	-0.001
	(0.006)	(0.010)	(0.006)
Total monthly profit of the business (000's PKR)	0.006	0.009	0.005
	(0.008)	(0.013)	(0.008)
Dummy: finds it hard to save	-0.014	-0.021	-0.014
	(0.018)	(0.032)	(0.020)
Index: respondent opinions taken into account in household decisions	0.025	0.035	0.022
	(0.019)	(0.036)	(0.021)
Dummy: faces pressure to share cash on hand	-0.005	-0.008	-0.006
	(0.020)	(0.034)	(0.021)

Index: respondent needs to ask permission for making decisions	-0.025	-0.037	-0.023
	(0.017)	(0.031)	(0.018)
Age (years)	-0.000	-0.001	-0.000
	(0.001)	(0.001)	(0.001)
Dummy: Respondent is currently married	0.001	0.004	0.003
	(0.011)	(0.018)	(0.011)
Level of education	-0.000	-0.001	-0.000
	(0.003)	(0.006)	(0.004)
Dummy: Respondent can read and write	-0.022	-0.045	-0.025
	(0.030)	(0.053)	(0.032)
Number of children	-0.008	-0.014	-0.008
	(0.004)*	(0.007)*	(0.004)*
Head of the household	0.016	0.021	0.012
	(0.027)	(0.041)	(0.025)
Neighbourhood effects	yes	yes	yes
Observations	1801	1603	1801

This table provides an analysis of automatic refusals in Phase II by subject characteristics, after controlling for neighborhood level effects. In each column, we show a regression of a subject automatically refusing the product in any wave, on individual characteristics. Specifically, we show results from an OLS regression with neighborhood dummies in column (1), a logit regression with neighbourhood fixed effects in column (2) and a logit regression with neighbourhood random effects in column (3). For the random-effect logit, we estimate $\rho = 0.57$*, with a 95% confidence interval of* $[0.39, 0.73]$*. Standard errors are clustered at the household level for the OLS regression. We use '*' to denote confidence at the 90% level.*

In Table 3, we show average take-up frequencies for all six combinations of lumpsum payment timing and net balance, for both phases. In Table A4, we repeat the analysis but exclude 'automatic refusers', and find that take-up patterns do not change in an economically meaningful way. Take-up is positive for all six contracts and responds to contractual terms: take-up is lower when lumpsum is paid out later and higher when lumpsum amount is higher.

Table A4: **Average take-up by contract terms, excluding automatic refusers**

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

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

C Robustness checks

C.1 Understanding of the product

In both Phase 1 and Phase 2 of our field experiments, we collected extensive data on respondents' understanding of the basic concept and structure of our microfinance product.

C.1.1 Familiarity with savings committees (*i.e.* ROSCAs)

We begin by summarising respondents' familiarity with the concept of a savings committee. As the following two tables show, familiarity with the concept of a savings committee was extremely high, in both Phase 1 and Phase 2 (*i.e.* above 90% in both phases).

Table A5: **Phase 1: Are you familiar with the concept of a savings committee?**

	Number	Percentage
Yes:	760	96.3%
No:	29	3.7%

Table A6: **Phase 2: Are you familiar with the concept of a savings committee?**

	Number	Percentage
Yes:	2210	91.5%
No:	206	8.5%

In both phases, we asked for experience of direct participation in a committee. The following two tables show that about half of our Phase 1 sample had participated, and about a quarter of our Phase 2 sample. (This difference, of course, is consistent with the different sampling schemes used; as discussed in the paper, the Phase 1 sample focussed on microenterprise owners.)

Table A7: **Phase 1: Have you ever participated in any committee?**

	Number	Percentage
Yes:	404	51.2%
No:	385	48.8%

Table A8: **Phase 2: Have you ever participated in any committee?**

	Number	Percentage
Yes:	640	26.5%
No:	1776	73.5%

From these four tables, we conclude that (i) the vast majority of the respondents were familiar with the concept of a savings committee, and (ii) a substantial share had participated in one. These results provide initial reassurance that respondents understood our product – given that it was closely based on the structure of a savings committee (and, indeed, explained to respondents by drawing a direct analogy to the committee structure).

C.1.2 Reasons for refusal

Further support for this conclusion comes from direct questions about the reasons for product refusal. The following three tables show the reasons given, in each of the three rounds of Phase 1, for refusing the offered product.

Table A9: **Phase 1: Reasons given for refusing (Round 1: 314 rejected; 192 provided reasons)**

	Number	Percentage
'I do not understand how the product works':	6	3.1%
'I cannot obtain the money each week to pay':	157	81.8%
Other:	29	15.1%

Table A10: **Phase 1: Reasons given for refusing (Round 2: 332 rejected; 211 provided reasons)**

	Number	Percentage
'I do not understand how the product works':	4	1.9%
'I cannot obtain the money each week to pay':	185	87.7%
Other:	22	10.4%

Table A11: **Phase 1: Reasons given for refusing (Round 3: 308 rejected; 165 provided reasons)**

	Number	Percentage
'I do not understand how the product works':	2	1.2%
'I cannot obtain the money each week to pay':	149	90.3%
Other:	14	8.5%

These tables show that, in each round, a negligible proportion of respondents complained that they did not understand how the product works. (This question was optional – and, of course, many respondents declined to provide a reason. However, if misunderstanding truly was an important reason, we would expect many more respondents to have complained of it.)

In Phase 2, we observed a similar pattern: that is, the vast majority of respondents who provided reasons for refusal indicated that this was due to not having regular access to money (about 75% in each round). In Phase 2, product misunderstanding was not offered as an explicit option (given that it was reported so rarely in Phase 1); respondents had the option to report this in the 'other' category, but not a single respondent did so.

C.1.3 Tests of understanding

In Phase 1, we also asked two questions explicitly on product understanding, at the endline survey. It is worth noting that our conclusions here are likely to be biased *against* respondents understanding the product, because these questions were asked approximately six months after the product was initially explained (and at least six weeks after the final take-up decision had been elicited). Nonetheless, we find strong evidence of understanding here, too.

First, we described a hypothetical contract, and asked respondents a 'right/wrong' question on when they would be paid. We found that:

- 315 of 363 (that is, 86.8%) answered *correctly*;
- 32 of 363 (that is, 8.8%) answered *incorrectly*;
- 16 of 363 (that is, 4.4%) *refused to answer*.

Second, we asked people whether they agreed with the simple statement: 'I understand how the new contracts work'. We found that:

- 68 of 363 (that is, 18.7%) *disagreed* (or strongly disagreed);
- 69 of 363 (that is, 19.0%) were *neutral*;
- 226 of 363 (that is, 62.2%) *agreed* (or strongly agreed).

We interpret both of these results as showing strong evidence that most respondents understood the product at the time of their decisions. As further support for this conclusion, we then regress each of these outcomes (that is, a dummy for answering the quiz correctly and a dummy for agreeing with the statement of understanding) on three measures of mental acuity: a dummy for the respondent being literate, a dummy for getting a numeracy question right, and a digitspan score. In each case, we find that the coefficients are small – and in none of the cases is the correlation significant.

C.2 Planned and actual spending

Finally, we can compare the way that the respondents reported at baseline that they would spend a hypothetical lump sum, and compare this to same pattern for respondents' reports of actual spending of the lump sums received through the product (at endline). If it were the case that respondents adopted our product without understanding its basic operation, one might expect these two patterns to differ substantially; instead, they are remarkably similar.

Figure A1: **Phase 2: Planned spending of the lump sum**

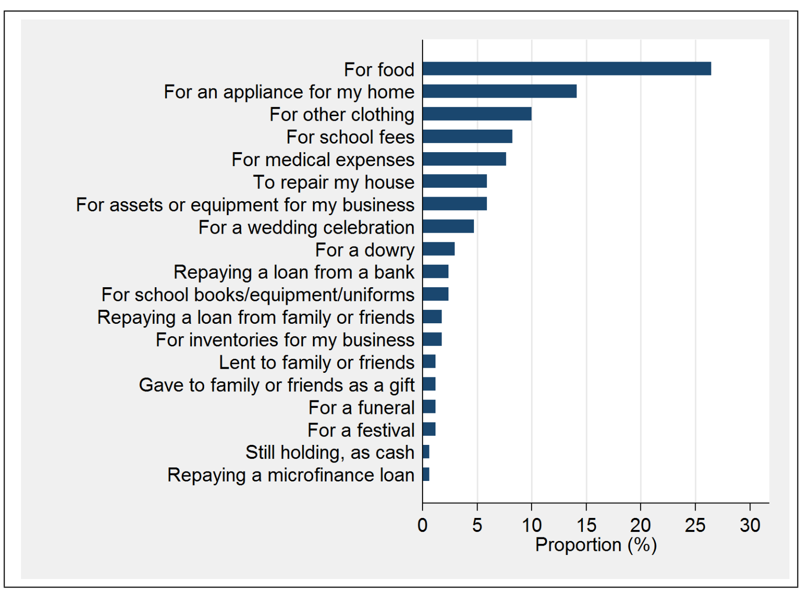
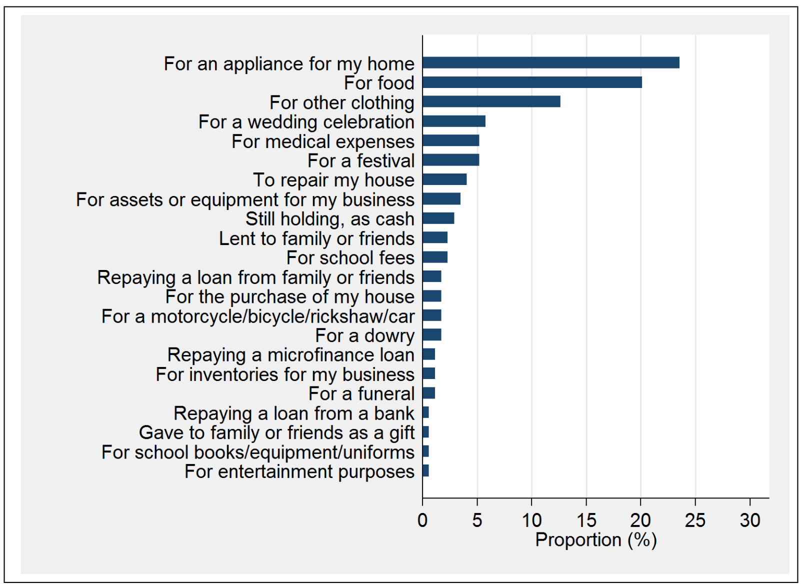


Figure A2: **Phase 2: Actual spending of the lump sum**



C.3 Additional analysis on dynamics

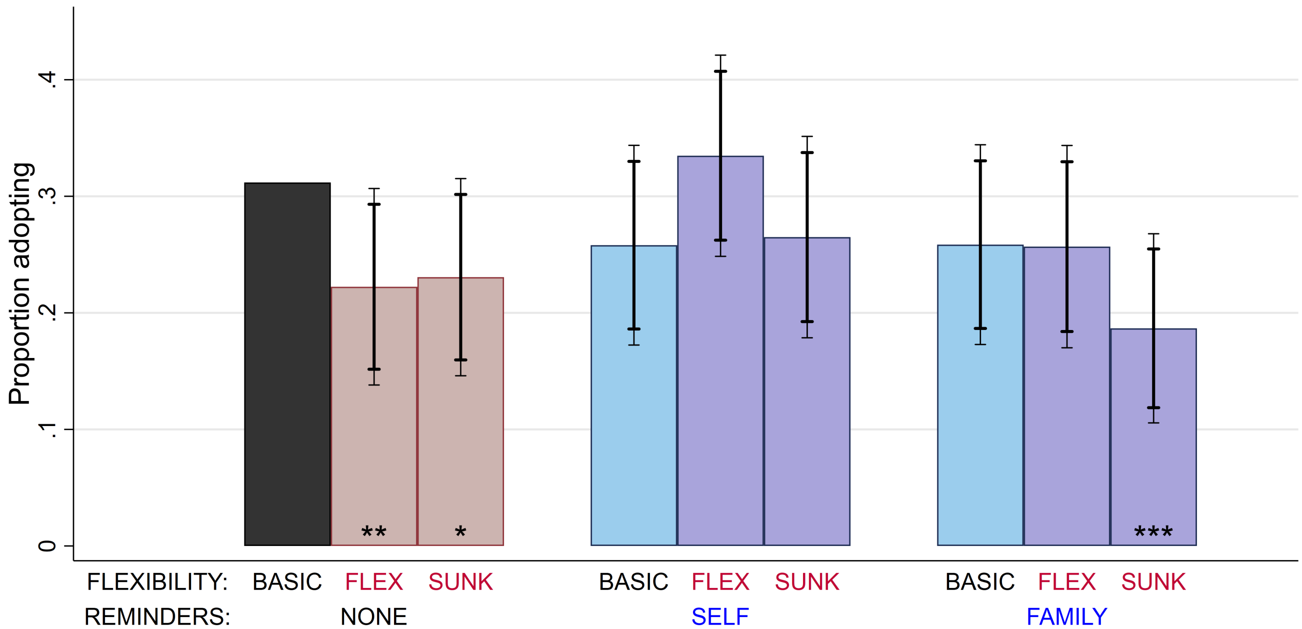
C.3.1 Disaggregating by wave

In this section, we provide additional analysis on take-up dynamics. In the original Figure A3, we show (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); the figure 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. Panel A of Figure A3 (at the top) shows these results for credit contracts (that is, contracts where the lumpsum is offered to be paid in the first period); Panel B (at the bottom) shows the results for savings contracts (where the lumpsum is to be paid in the final period).

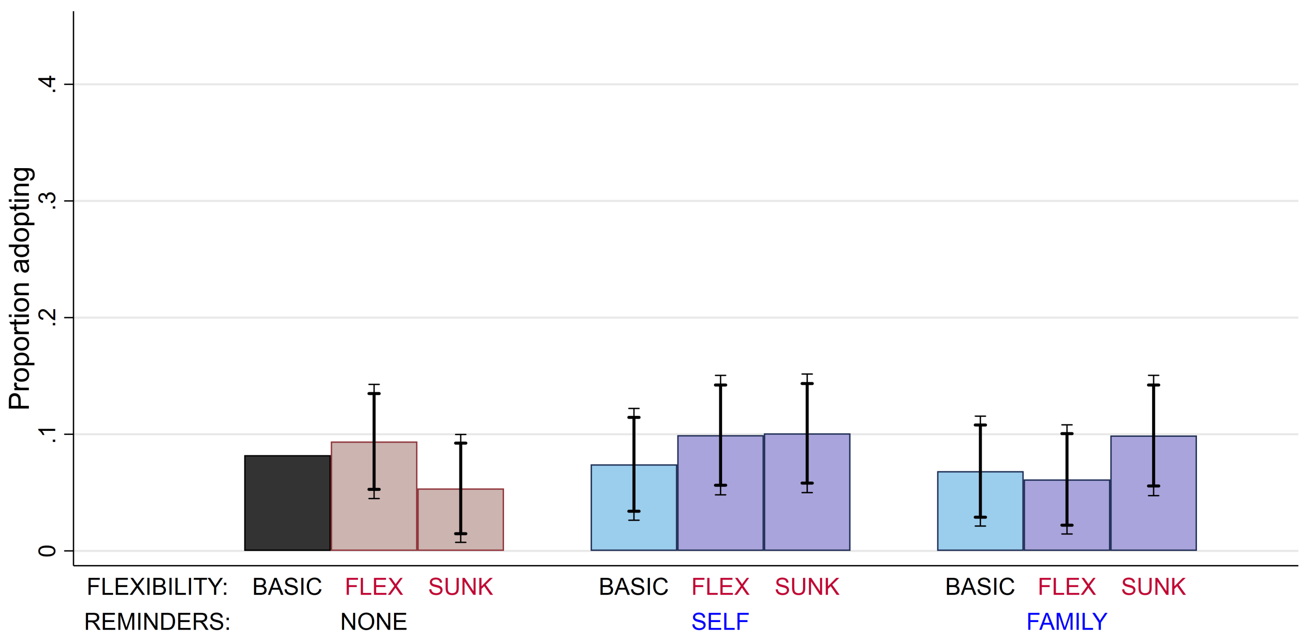
Figure A4 then disaggregates Panel A of Figure A3 by experiment wave; that is, it repeats the credit analysis, splitting the same into wave 1, wave 2 and wave 3. Figure A5 shows the same disaggregating for savings (that is, Panel B of Figure A3). While there is (inevitably) some variation in take-up patterns between waves – and, of course, a noticeable widening of the confidence bars due to the reduction in power – the graphs show that the general patterns observed in the pooled data are reflected in each of the three waves separately. As an additional test of relevance of wave effects, we re-estimate the main analysis with wave fixed effects. Coefficients across models with and without wave fixed effects are remarkably similar ($p - value$ of difference = 0.967).

Figure A3: **Average take-up by behavioral variations**

PANEL A: CREDIT CONTRACTS



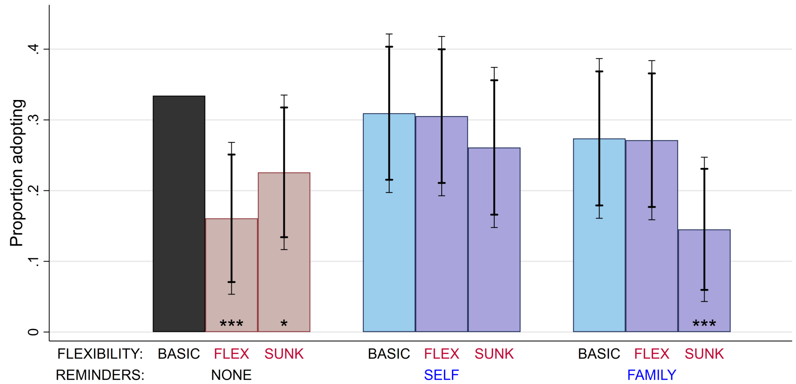
PANEL B: SAVING CONTRACTS



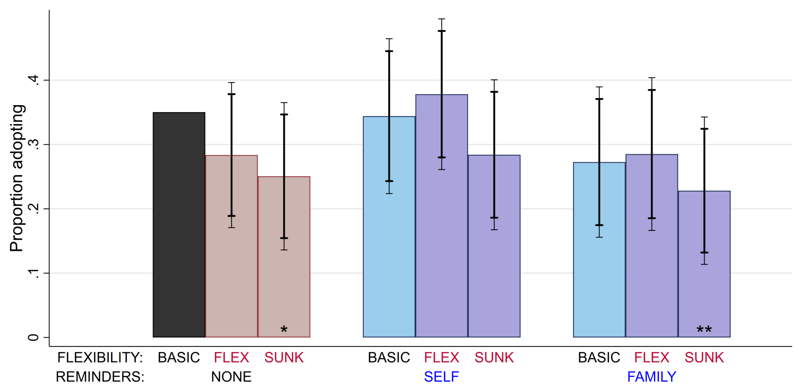
This figure shows the average take-up for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'family' variation), and take-up for each of the eight possible variations. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic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

Figure A4: **Average credit take-up by contractual add-ons: Disaggregating by wave**

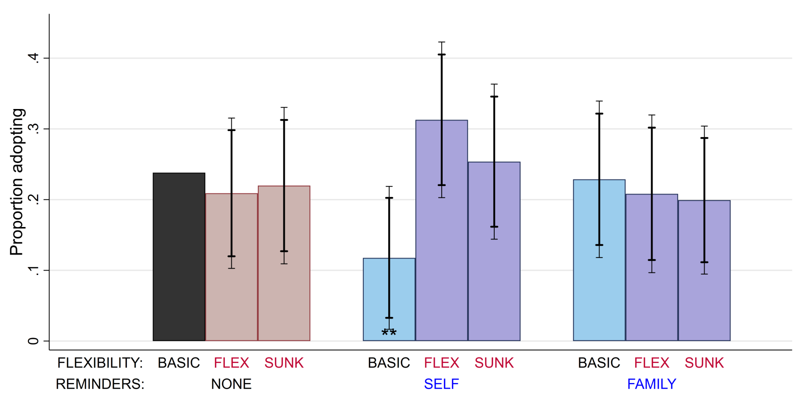
PANEL A: WAVE 1



PANEL B: WAVE 2



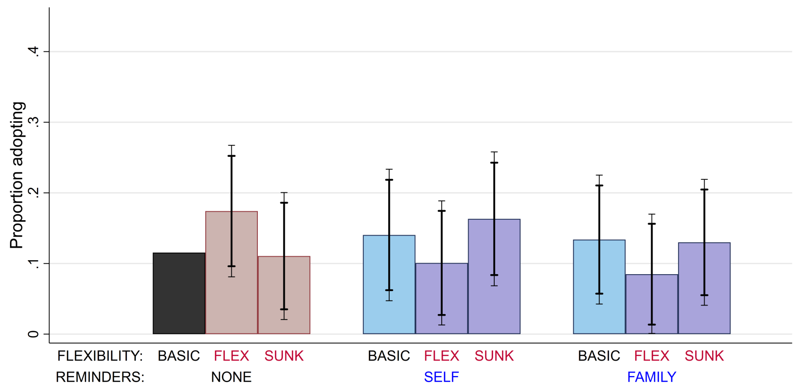
PANEL C: WAVE 3



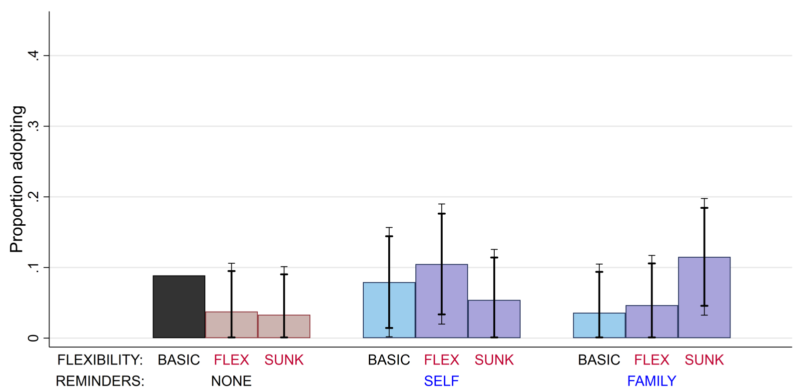
This figure shows the average take-up for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'family' variation), and take-up for each of the eight possible variations. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic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

Figure A5: **Average savings take-up by contractual add-ons: Disaggregating by wave**

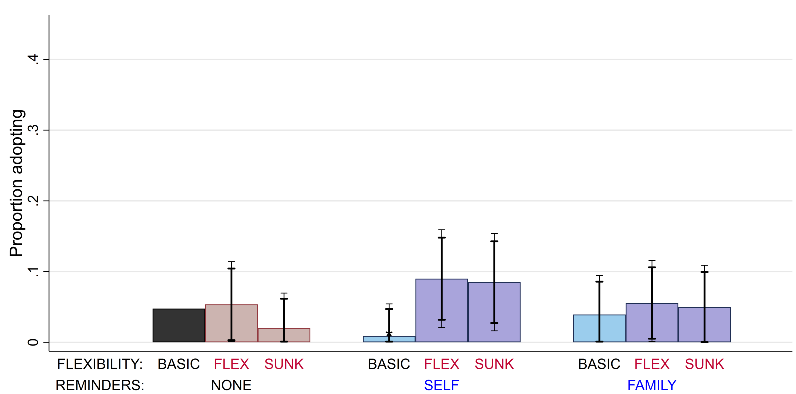
PANEL A: WAVE 1



PANEL B: WAVE 2



PANEL C: WAVE 3



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 variations. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic contract. Stars indicate a significant difference from take-up of the basic contract; that is, we reject a null hypothesis of equal take-up rates for the 'sunk' variation and for the 'sunk and family' variation, each at the 5% significance level.

C.3.2 Instrumenting for lagged take-up

Next, we explore the effect of take-up in a given wave on take-up in the following wave. To do this, we estimate a linear probability model, and we instrument lagged take-up with the lagged contract terms (that is, the interest rate and the time of payment). We do this for our Phase 2 data, and we show the results in Table A12. In column (1), we run this IV estimation for waves 2 and 3 (noting, of course, that we cannot include wave 1, because this wave is necessary to form the first lag). We estimate 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. (Note that, in total, about 12% of respondents are taking up in waves 1 and 2: the waves that then form the lags respectively for waves 2 and 3.) In columns (2) and (3), we show that this effect is remarkably stable when we disaggregate by wave (that is, when we estimate separately for wave 2 (with lagged take-up being wave 1) and for wave 3 (with lagged take-up being wave 2).

In columns (4), (5) and (6), we then explore the implications of this for our analysis of sensitivity of contract terms. In column (4), we show the percentage take-up in each of the six cells defined by interest rate and time of payment; that is, column (4) replicates exactly the Phase 2 figures in the top panel of Table 4 in the paper. In column (5), for completeness, we repeat this exercise just for take-up in waves 2 and 3; consistent with the graphs on the preceding pages, we find that the take-up rates are very stable over time. Finally, in column (6), we repeat the exercise while also including lagged take-up (instrumented, in the same way as in columns (1), (2) and (3)). We find that our estimates on the effect of contract terms are virtually unchanged. This makes strong intuitive sense: because we randomized the contractual offers, the offer terms are uncorrelated to lagged take-up – and, therefore, the inclusion or omission of lagged take-up does not change our conclusions. Nonetheless, it is useful to confirm this in our particular empirical context.

Table A12: **Phase 2: Effect of lagged take-up (instrumented with lagged contractual terms)**

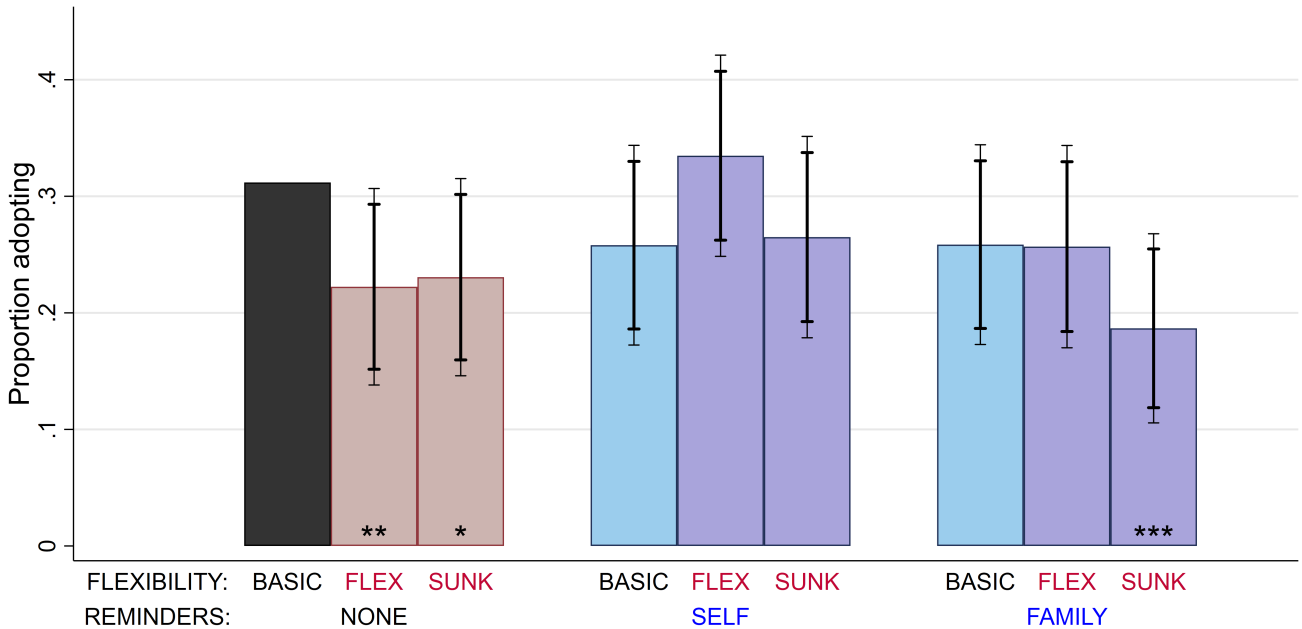
	(1) Waves 2 & 3	(2) Wave 2	(3) Wave 3	(4) Contract terms (all waves)	(5) Contract terms (waves 2 and 3)	(6) Contract terms adding lag
Lagged take-up	0.527	0.547	0.529			0.452
	(0.043)***	(0.074)***	(0.050)***			(0.041)***
Constant	0.061	0.079	0.039			
	(0.009)***	(0.016)***	(0.010)***			
Take-up by contract terms:						
$r < 0$, $p = 1$				11.0%	8.4%	11.1%
$r < 0$, $p = 8$				4.1%	2.3%	5.2%
$r = 0$, $p = 1$				26.0%	24.1%	22.6%
$r = 0$, $p = 8$				8.9%	6.5%	8.4%
$r > 0$, $p = 1$				37.2%	40.1%	34.5%
$r > 0$, $p = 8$				11.3%	8.4%	9.8%
Obs	3628	1814	1814	5442	3628	3628
K-P Wald test (F)	66.05	22.14	58.98			
Lagged take-up rate	12.1%	13.9%	10.3%			12.1%

C.4 Sensitivity to including defaulters

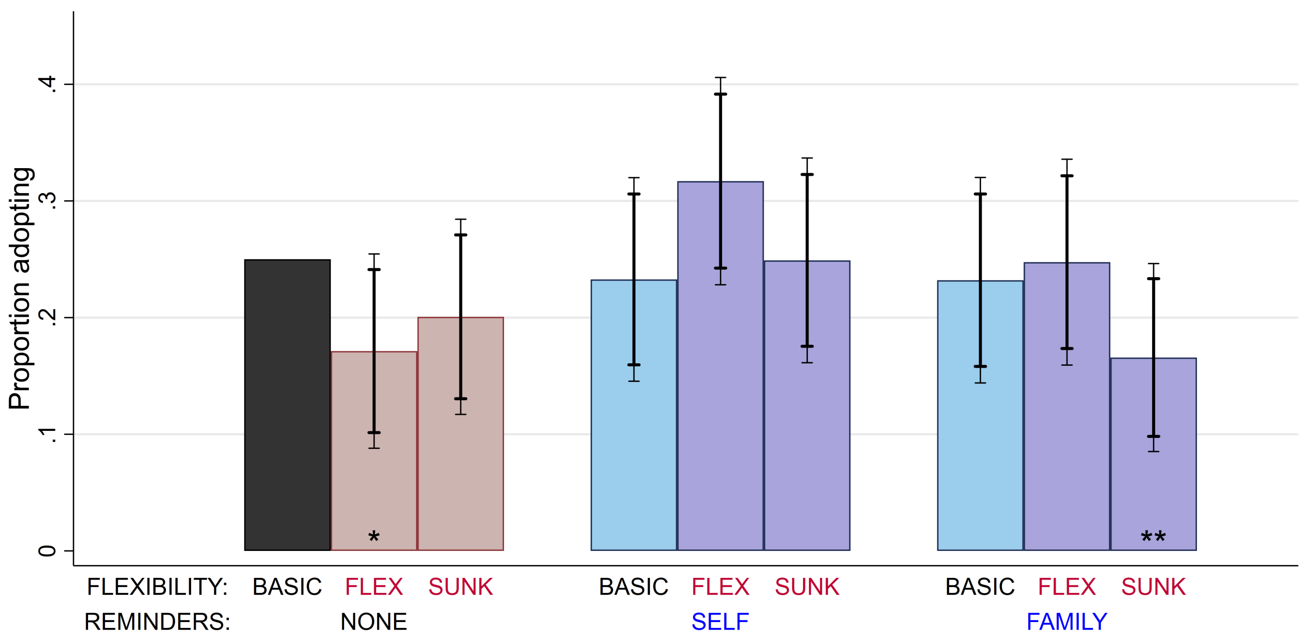
On the following two pages, we show the effect of dropping individuals who ever defaulted. (Thus, for example, if an individual defaulted in wave 3, we drop her observations in waves 1 and 2.) On each page, we show the original figure at the top (as reported in the paper), and the amended figure at the bottom (dropping 'ever-defaulters'). By construction, the take-up rates in the lower graph are reduced (because those who ever default are also more likely to have taken up), and the standard errors are slightly larger (because we are reducing the sample size). However, in both figures, we see that the overall take-up patterns and general conclusions are unaffected by this.

Figure A6: **Average take-up by contractual add-ons: Credit domain**

PANEL A: CREDIT CONTRACTS



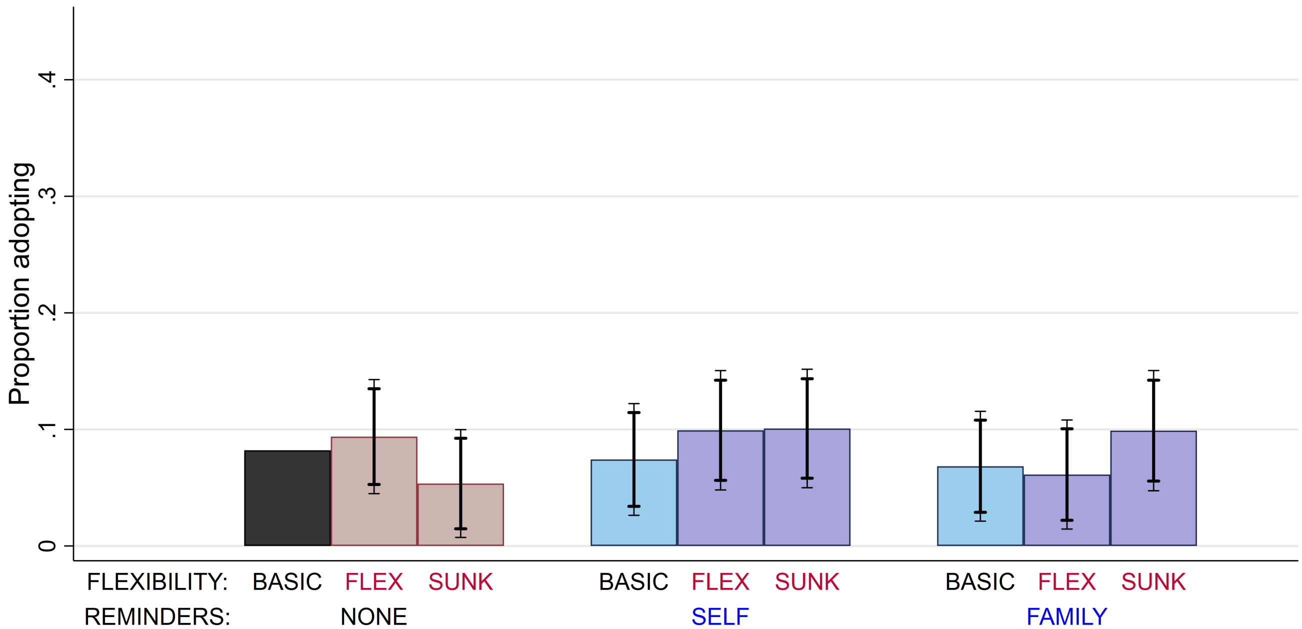
PANEL B: CREDIT CONTRACTS, DROPPING 'EVER DEFAULTERS'



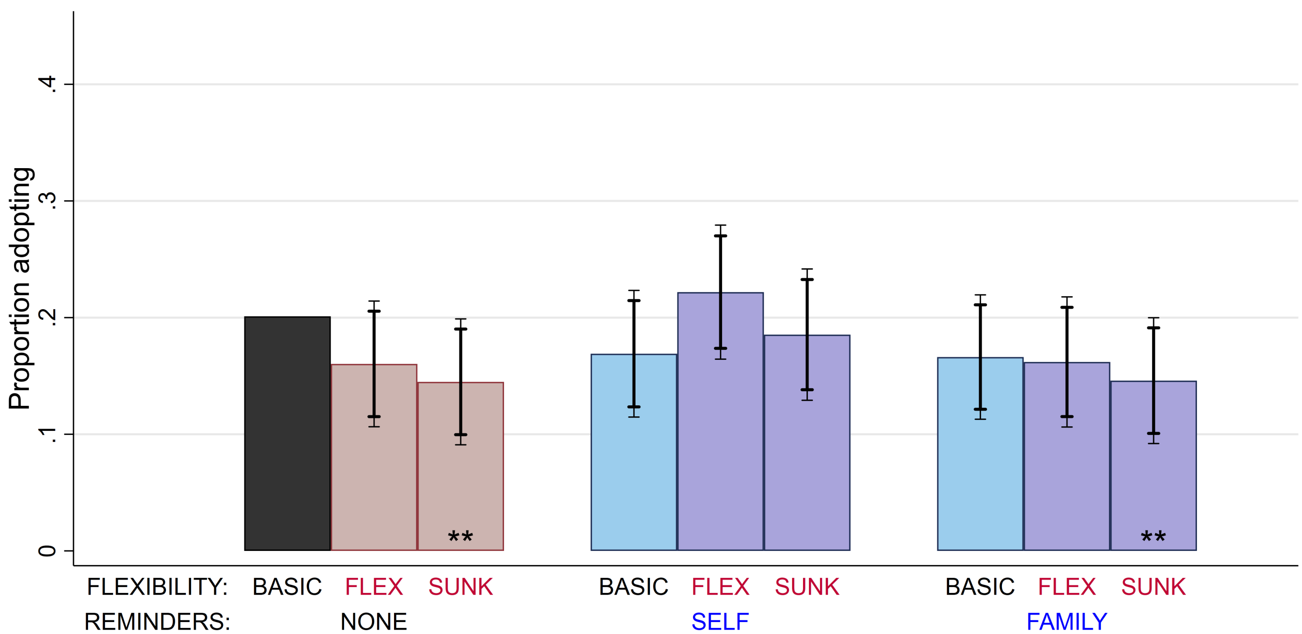
This figure shows the average take-up for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'family' variation), and take-up for each of the eight possible variations. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic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

Figure A7: **Average take-up by contractual add-ons: Saving domain**

PANEL A: SAVING CONTRACTS



PANEL B: SAVING CONTRACTS, DROPPING 'EVER DEFAULTERS'

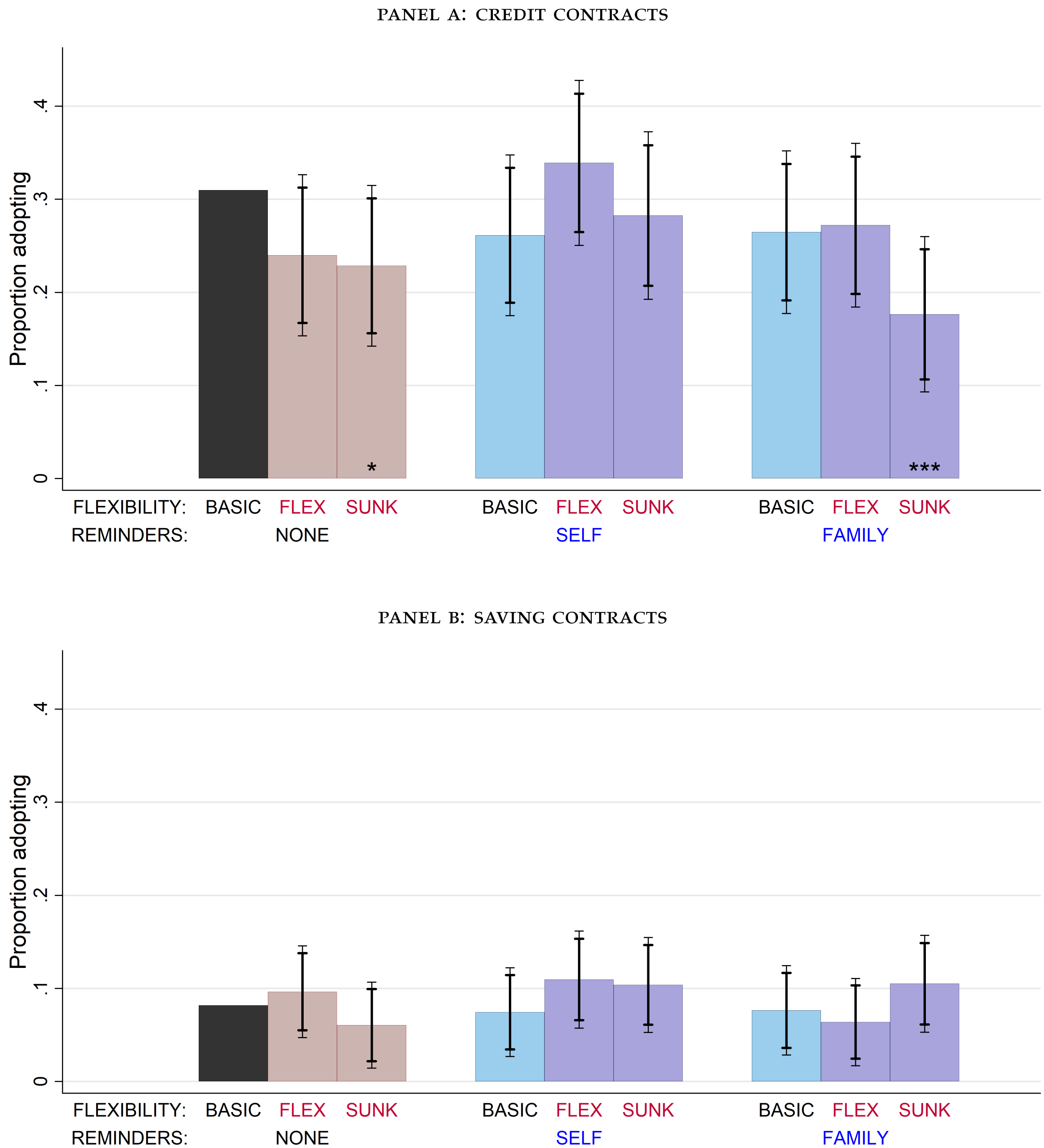


This figure shows the average take-up for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'family' variation), and take-up for each of the eight possible variations. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic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

C.5 Sensitivity to including baseline characteristics as controls

Next, we test the robustness of our estimation results to the inclusion of baseline controls and find that our main estimation results are unaffected. We show (on the far left) of Figure A8, take-up rates for the basic contract (that is, the product with neither the 'flex'/'sunk' variation nor the 'self reminder' / 'family reminder' variation). We then calculate point estimates of the difference in take-up rates for each of the eight possible contractual add-ons, while controlling for baseline characteristics using a post-double LASSO estimation. Error bars show 90% and 95% confidence intervals on the difference in take-up relative to the basic contract. In Figure A9, we similarly show the effects on rate of late payment.

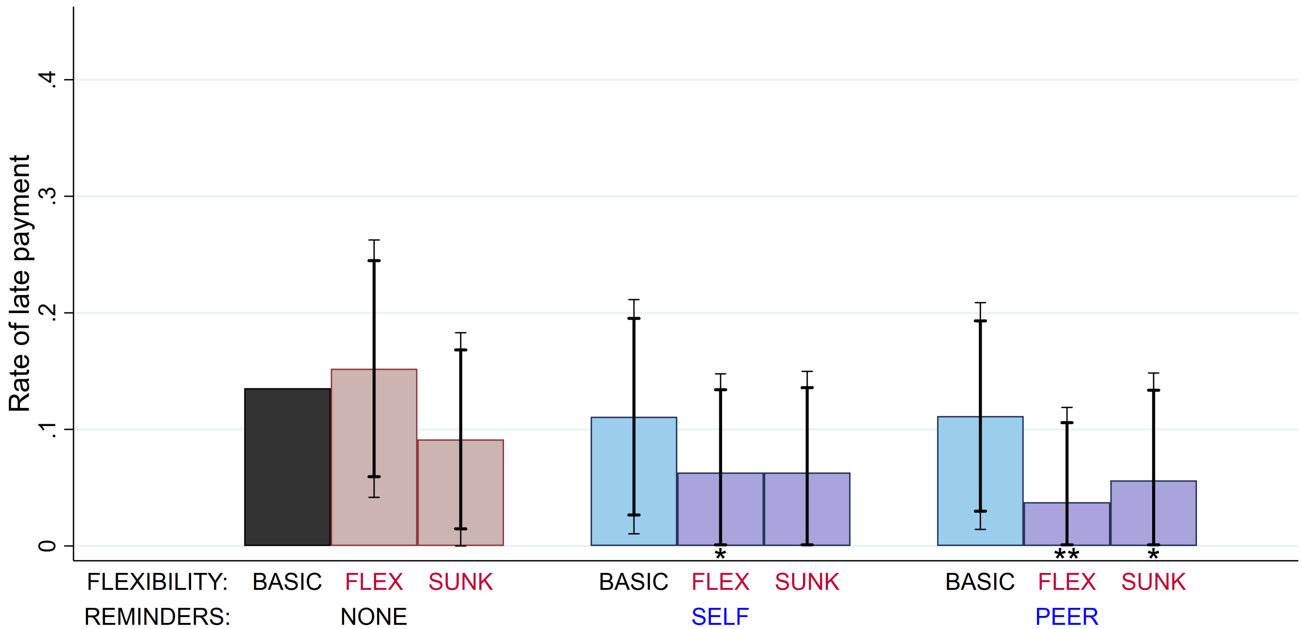
Figure A8: **Average take-up by contractual add-ons: including baseline controls**



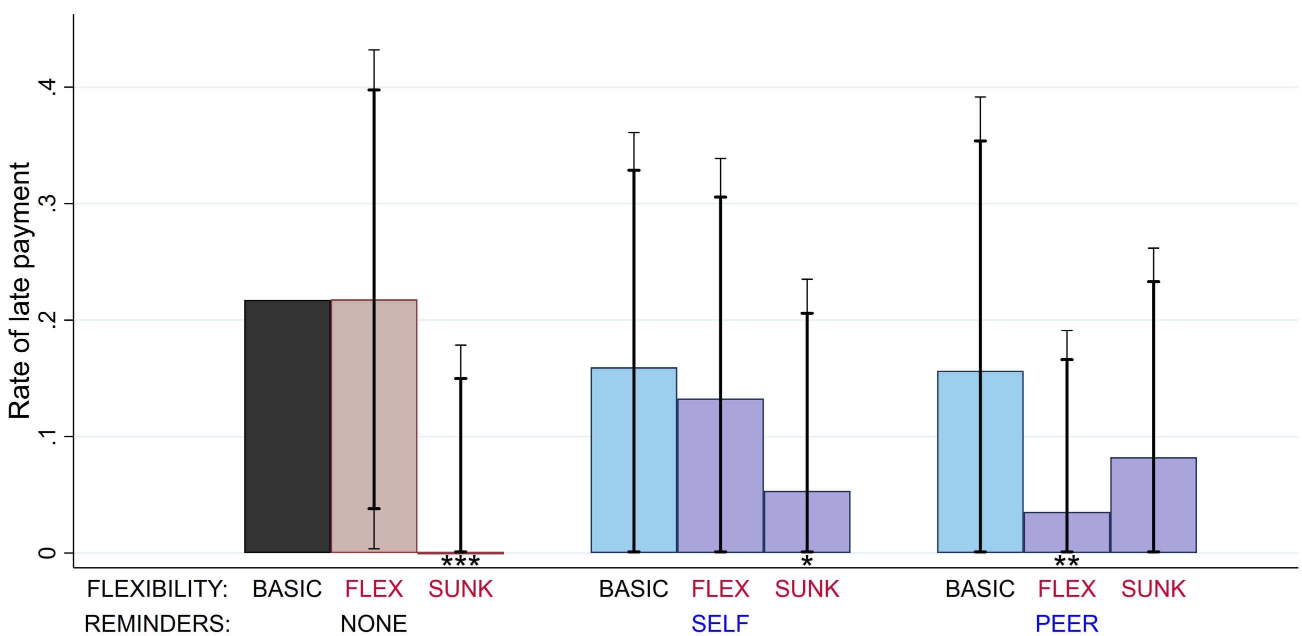
This figure shows the average take-up for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'family' variation), and take-up for each of the eight possible add-ons. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic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

Figure A9: **Rate of late payment by contractual add-ons: including baseline controls**

PANEL A: CREDIT CONTRACTS



PANEL B: SAVING CONTRACTS



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, where the minimum rate of late payment is set to 0. Error bars show 90% and 95% confidence intervals on the difference in rate of late payment to the basic contract. Stars indicate a significant difference from the basic contract.

D Conceptual Framework

To provide a theoretical frame to the empirical analysis, we present three stylised benchmark scenarios: (1) when subjects can hold on to cash; (2) when they cannot hold on to cash but have no need for bunching their expenditures into a lumpsum; and (3) when they cannot hold on to cash and have a demand for lumpsum accumulation. We then draw on the insights gained from this framework to guide our empirical analysis – both our initial analysis of take-up patterns and our subsequent analysis of demand for behavioral features.

Scenario 1: Subjects can hold on to cash: We first discuss the case where subjects are able to save in a liquid asset (*i.e.* cash). When people can hold cash on their own, a simple arbitrage argument implies that taking up a saving contract with a negative return can never be optimal.[15] For take-up of a saving contract to be rational in this case, the subject must face meaningful economic costs of holding on to cash – for instance, because of self-commitment issues. The same reasoning applies to savings contracts with a zero return: subjects able to save on their own can mimic the contract without incurring the time cost of visiting the MFI to pay each instalment. It follows that take-up of saving contracts with a zero or negative return represents a lower bound on the demand for commitment: they are the least appealing commitment contracts.
By a similar arbitrage argument, credit contracts in which the lumpsum exceeds the value of the instalments should always be accepted by subjects who can hold on to cash.[16] This kind of loan may nonetheless be rejected by subjects who have difficulties holding cash – for example, 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 on to cash but have no regular income to service the debt or no need for lumpsum accumulation. We also note that subjects who can hold on to 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, pay the lumpsum later, or both.

Scenario 2: Subjects cannot hold on to cash: We now examine predicted take-up in the stylised case of a subject who *cannot* hold cash and for whom the contracts we offer are the *only* available

[15] This is because the decision-maker can exactly replicate the cash flow pattern of this contract by saving all the instalments 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 instalments of *PKR* 1000, the subject could simply set aside the instalments each week and end up with *PKR* 5000 – a strategy that dominates the contract.

[16] That is, a subject who can hold on to cash can use the loan to pay the instalments and keep the difference. For instance, the subject could take an upfront payment of *PKR* 5500, repay the five instalments 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 instalments is small enough relative to the surplus.

way of moving funds across periods. To do this, we use the standard framework of expected utility with exponential discounting and weekly discount parameter β.[17] We denote weekly income as y, and assume that y is drawn from a stationary distribution. This framework implies the following utility-maximising behaviour, comparing the net present value of the two alternative utility streams:

1. Take a credit contract if and only if the:

$$\sum_{t=1}^{N} \beta^t \cdot \mathbb{E}\left[U(y)\right] \leq \beta \cdot \mathbb{E}\left[U(y+L)\right] + \sum_{t=2}^{N} \beta^t \cdot \mathbb{E}\left[U(y-M)\right]. \tag{A12}$$

2. Take a savings contract if and only if:

$$\sum_{t=1}^{N} \beta^t \cdot \mathbb{E}\left[U(y)\right] \leq \sum_{t=1}^{N-1} \beta^t \cdot \mathbb{E}\left[U(y-M)\right] + \beta^N \cdot \mathbb{E}\left[U(y+L)\right]. \tag{A13}$$

Figure A10 shows the predictions of the model for linear utility.[18] In Panel A and Panel B of that figure we graph the indifference curves implied by equations A12 and A13 for Phase 1 and Phase 2, respectively (the remaining panels introduce variation in a parameter θ, which we introduce shortly). In each case, the horizontal axis shows the 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.[19] 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 graph shows that, for all $\beta < 1$, there exists values of $L/[(N-1) \cdot M]$ at which a client accepts a

[17] N is the duration of the contract, L is the lumpsum, and M is the instalment.

[18] Specifically, the condition in equation A12 implies:

$$\frac{\mathbb{E}\left[U(y+L)\right] - \mathbb{E}\left[U(y)\right]}{(N-1)(\mathbb{E}\left[U(y)\right] - \mathbb{E}\left[U(y-M)\right])} \geq \frac{\beta - \beta^N}{(N-1)(1-\beta)},$$

where, under linear utility, the lefthand side of this expression simplifies to $L/[(N-1) \cdot M]$. Note that the condition in equation A13 implies:

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

[19] The downward-sloping line in the upper section of each figure shows the indifference 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 $L/[(N-1) \cdot M]$.

loan contract but not a saving contract.[20] All subjects take credit 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 A10. 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 credit contract with a 0.9 payout ratio.[21] The model thus predicts 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. Put differently, with no particular demand for a lumpsum, the model predicts that the take-up rate of the positive-balance saving contract and the take-up rate of the negative-balance loan contract should sum to approximately 100 percentage points. Second, if subjects have stable time preferences – and thus a time-invariant β – Figure 1 generates testable consistency conditions on choices made across waves. For instance, someone who takes a loan with payout ratio 0.9 in one wave (and thus has a $\beta < 0.965$) should reject a savings contract with a 1.1 payout in another wave.

Scenario 3: Subjects cannot hold on to cash but have a preference for a lumpsum: Predictions are different if subjects have a specific desire to accumulate a lumpsum. Subjects may value receiving a lumpsum L above its monetary value because it allows them to purchase a durable good producing a flow of services with future discounted value larger than L – as, for example, in Brune and Kerwin (2019), Attanasio and Pastorino (2020) and Besley et al. (1993).[22] Alternatively, someone may wish to spend L on a ceremonial expenditure that also generates memories and social capital of value greater than L.

The effect on take-up of a preference for a lumpsum can be examined by multiplying the payout ratio $L/[(N-1)\cdot M]$ in equations (A12) and (A13) by a parameter $\theta \geq 1$ and redrawing Figure A10. Such a change is illustrated in Panels C and D. We see that setting $\theta > 1$ shifts both indifference curves lower. This makes intuitive sense: multiplying by $\theta \geq 1$ makes the product more desirable, and thus expands the range of values of β at which taking up some of the contracts is optimal. For θ clearly greater than 1 (for example, $\theta = 1.05$, as in Panels C and D), the range of

[20] For example, the model predicts 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 $L/[(N-1)\cdot M] = 1.1$. 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.

[21] This is not a coincidence: in logs, the expression $\frac{\beta-\beta^N}{(N-1)(1-\beta)}$ has a slope of approximately $0.5N$ while the expression $\frac{\beta-\beta^N}{(N-1)\beta^N\cdot(1-\beta)}$ has a slope of approximately $-0.5N$.

[22] 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.

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. This makes intuitive sense: multiplying by $\theta > 1$ increases the value of a lumpsum, so increases overall demand in both the credit and the savings domain. 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 instalments or, alternatively, because they wish to smooth their consumption.[23] In those cases, the indifference curves shift upwards and this logic reverses: only very impatient subjects take a costly loan (*i.e.*, for which $L/[(N-1)\cdot M] < 1$) 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 waves – for example, because of fluctuations in the utility of lumpsum accumulation or in the anticipated utility cost of instalments – it might be possible to observe subjects borrowing in some waves 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, 2021). In our 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 waves and the motives behind the demand for lumpsums.

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, 2021). In our framework,

[23] To see the latter, focus on the case where $L = (N-1)M$ and consider the take-up inequalities (A12) and (A13). 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 instalment M (the denominator) – hence their ratio is less than 1. The effect of this on take-up can be mimicked by multiplying the left-hand side of equations A12 and A13 by $\theta < 1$.

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 waves and the motives behind the demand for lumpsums.

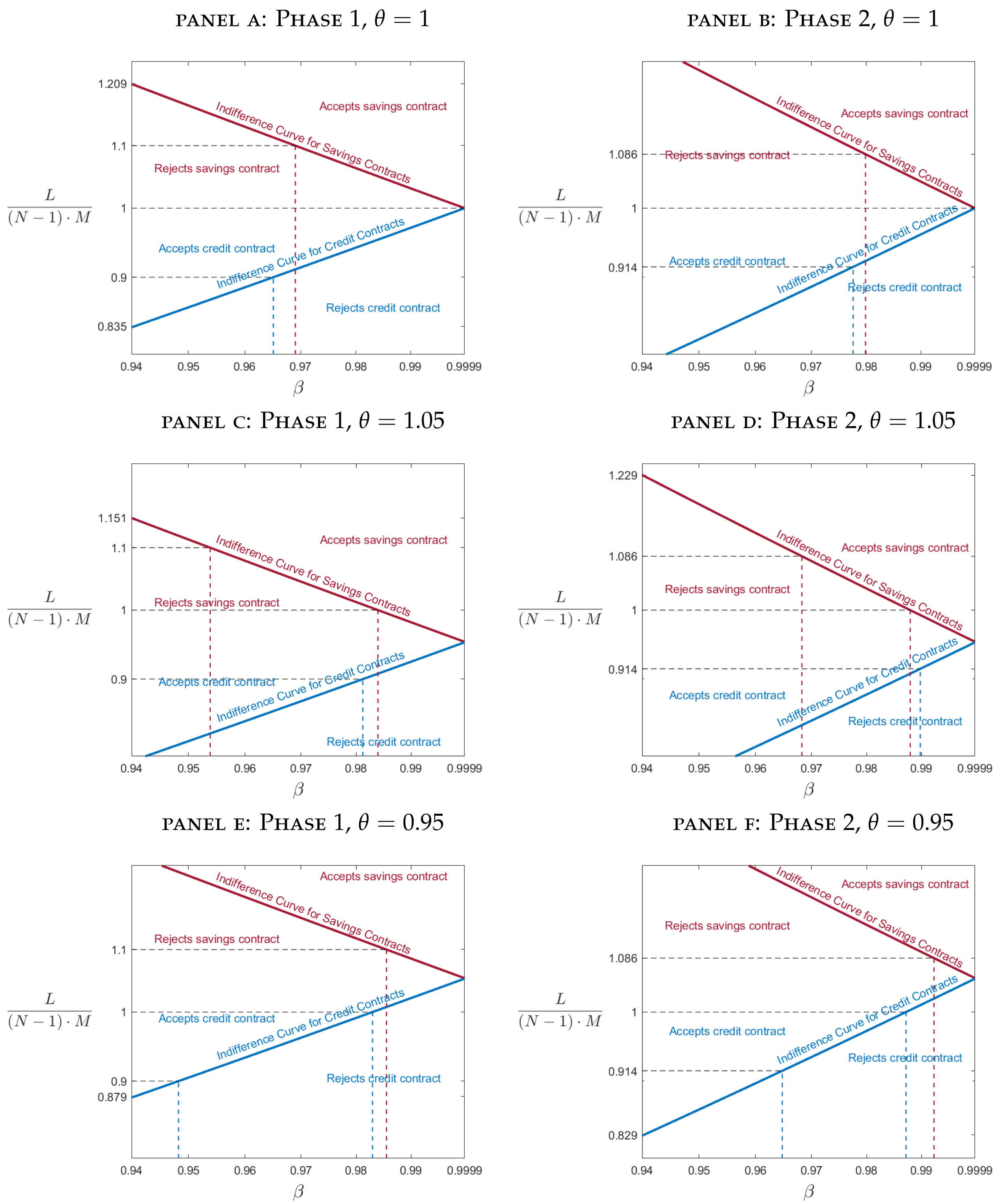
E Implied discount factors

Since contract offers are randomized across the same subjects within each wave, 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. For a model scenario to fit the data, the two distributions should therefore overlap. This is testable under the experimental design with both credit and saving products being offered to the same sample over waves. If β's overlap over the three waves 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 – for example, 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 instalments M varies with income and other shocks (i.e., $(N-1)(\mathbb{E}[U(y)] - \mathbb{E}[U(y-M)])$ is larger).

In this appendix, we implement this idea by computing, for each of the six graphs in Figure A10, 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. (That is, Figure A10 predicts a critical β value for offered contract; for behaviour generated by that model, the proportion of respondents having β less than the critical value can be inferred directly by the proportions taking up and rejecting the offered contracts.) 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, it means that this model scenario is rejected by the data.

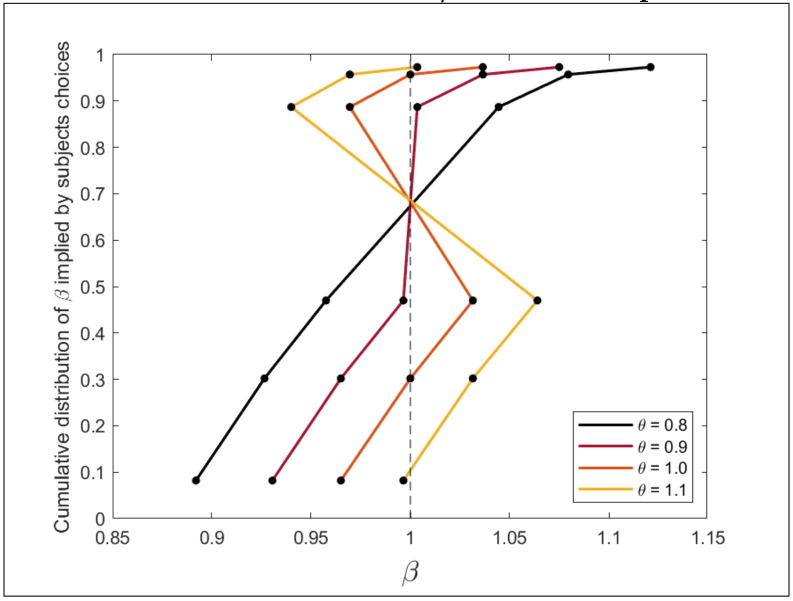
The results are presented in Figure A11, combining Phase 1 and Phase 2 subjects. Four sets of markers are shown, corresponding to four values of θ, i.e.: $\theta = 1$ (linear model); $\theta = 1.1$ (high demand for lumpsum accumulation); and $\theta = 0.9$ and $\theta = 0.8$ (a net preference for consumption smoothing). Each of the set of six points associated with a given value of theta map out an 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 experimental phases of Table 3. 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 waves, this θ is inconsistent with a demand for lumpsum accumulation.

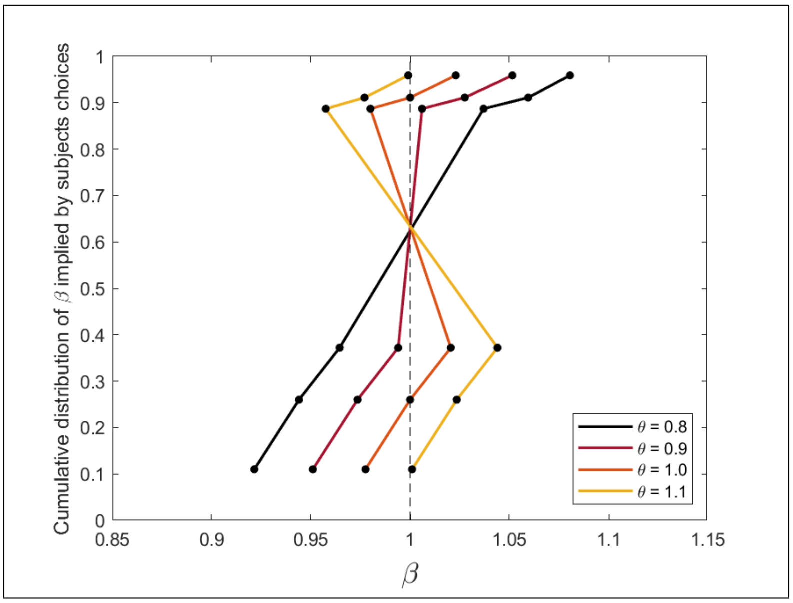
Figure A10: **Take-up predictions: Indifference curves**



Each graph shows the two indifference curves from equations A12 and A13. 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 $L/[(N-1) \cdot M]$. θ refers to the relative strength of preference for lumpsum accumulation. The graphs use a log transformation for β . 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 $L/(N-1)M$.

Figure A11: **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 waves. 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 A12 and A13, 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 implies rejection of a particular value of θ.

F Heterogeneity in take-up: A machine learning analysis

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

We implement the Chernozhukov et al. (2018) method as follows. First, we randomly split treated respondents into auxiliary and main samples. In the auxiliary sample, we use a machine learning method to estimate the probability of product adoption, conditional on a vector of 58 baseline covariates. Specifically, we use an elastic net with a logistic link function; for each random split of the data, we rescale the covariate vector, then tune and train the model using two-fold cross validation – choosing α (the mixing percentage) and λ (the regularization parameter) to minimize deviance. We then use the estimated parameters from this model to predict take-up in the main sample, for post-processing.[24] 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 A12 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 A12, are a direct analog to Figure 4 in Chernozhukov et al. (2018). They show that we have substantial heterogeneity in take-up rates across individuals with different covariates: for the lowest quintile, the average take-up rate is

[24] 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).

approximately 10%, and for the highest quintile, the rate is above 25%.[25]

We augment this analysis in two ways. First, in the top panel of Figure A12, 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.[26] 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 A12, 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 3 – 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 A25, we perform this comparison for all 58 of the baseline covariates used for our analysis. In Table A13, 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 A13 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

[25] 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).

[26] 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.

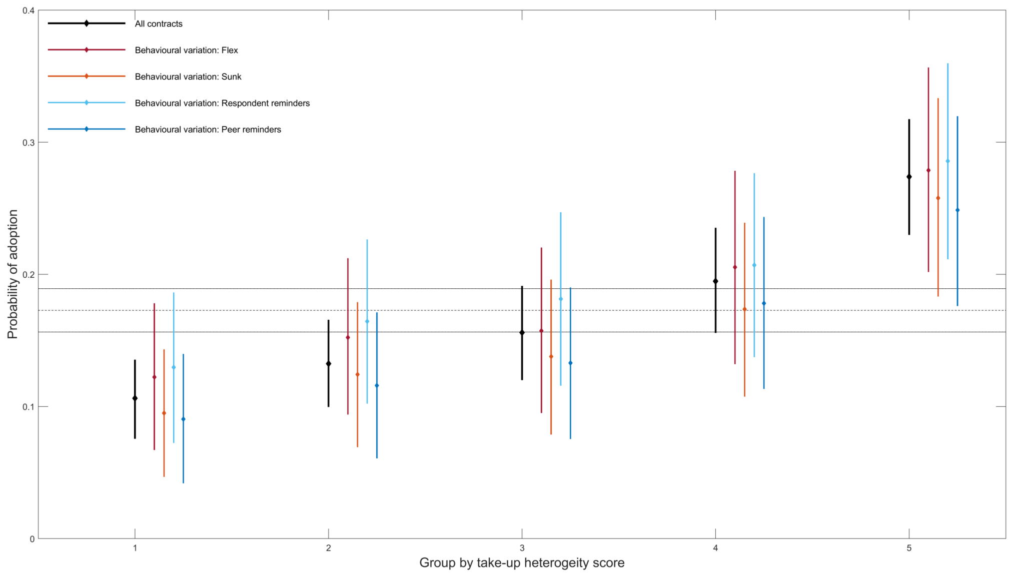
than twice as likely to agree that it is appropriate for a woman to invest in her business without consulting her husband and to go shopping for a personal item (specifically, a scarf).

The Phase 1 counterpart to Table A25 appears in the appendix as Table A26. 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.

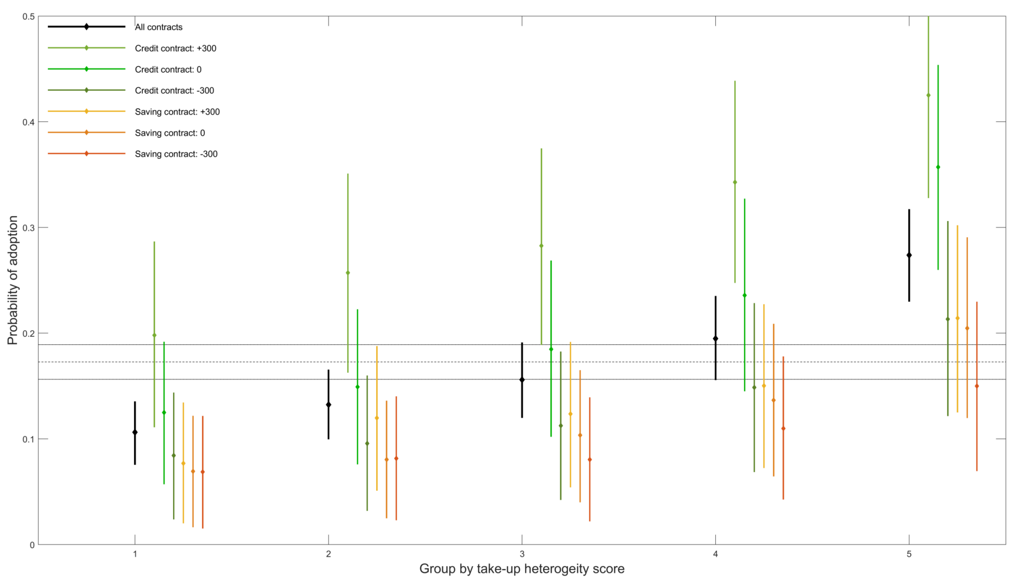
Finally, we show the Phase 1 counterpart to Figure A12 in the appendix, as Figure A23. The general patterns are the same, though Phase 1 respondents appear to have a greater sensitivity to the size of lumpsum payment.

Figure A12: **Group Average Treatment Effects (sorted by take-up propensity)**

PANEL A: TAKE-UP BY CONTRACTUAL ADD-ON



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 A13: **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).

G Heterogeneity in take-up: Respondent current loan and saving experience

In this appendix section, we use baseline covariates to characterise sample heterogeneity with respect to experience with the provider and current participation in loan or saving products. We explore variation in takeup of contractual addons among respondents who have had a long and successful relationship with the microfinance provider, among inexperienced borrowers and among respondents currently participating in an informal, saving committee. We focus on covariates that significantly correlate with take-up in the machine learning analysis summarizied in Table A25. That is, we explore whether the respondent is currently participating in a standard credit (loan) product offered by NRSP, and if the respondent is currently participating in a saving committee. In addition, to proxy a respondent's experience with NRSP, we explore heterogeneity by the number of years since the borrower first took out a loan with NRSP. We create a binary variable for if the respondent has been engaged with NRSP longer than the sample median (3 years).

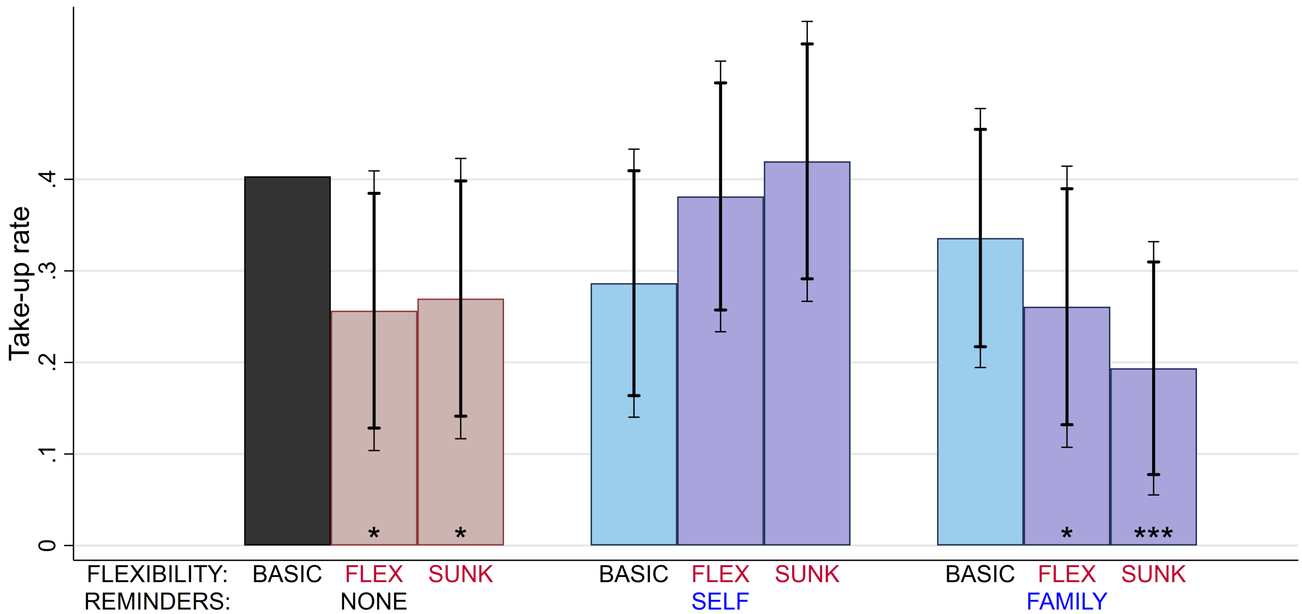
Figures A13 - A15 summarize the results for credit products. We find that respondents who are currently participating in a standard credit product, do not value the additional commitment incorporated in sunk contracts, especially when combined with reminders to family. Similar to the average trends discussed in Section 4.2 of the paper, contractual add-ons are not popular among older borrowers, who may be considered to be more experienced. The exception are flex contracts where respondents are also reminded of an upcoming payment, where take-up is higher than that of a basic contract but not significantly so. Notice that similar trends exist when we look at those with long and short engagements with NRSP (Figure A14), which is expected given the two measures are likely to be correlated. Current and new borrowers are less likely to value sunk products. Sunk contracts with family reminders are particularly disliked by the more experienced NRSP clients. Respondents with current saving product commitments particularly value the flex contract but only when it is combined with softer commitment in the form of reminders (Figure A15, Panel A). Respondents without existing commitments in the form of informal saving products dislike all forms of contractual add-ons, preferring the standard features of a basic product.

Figures A16 - A18 summarize a similar analysis for saving products. Statistical power to detect small effects is low due to overall low take-up of saving products. However, we note that current borrowers value soft commitment in the form of reminders to the respondent for an upcoming instalment payment. On the whole, the more experienced clients and borrowers do not value the add-ons. Current committee participants particularly dislike the family reminder feature (Figure

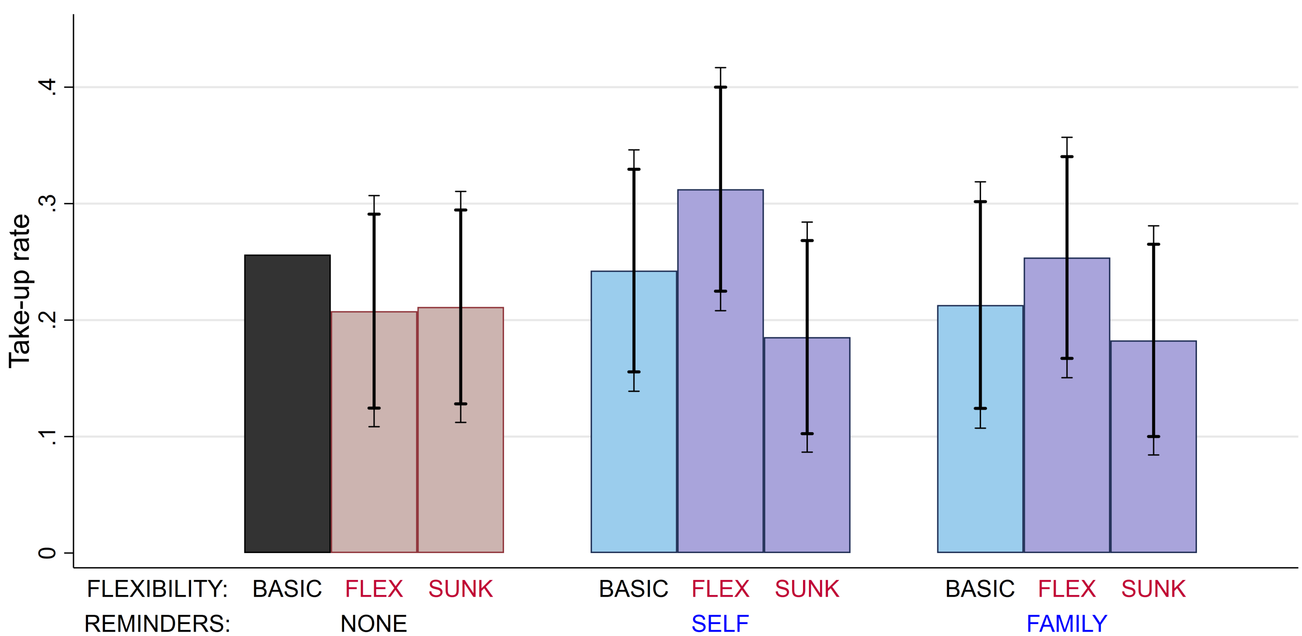
A18, Panel a). Conversely, those who likely lack experience with a saving product, value the additional commitment that a sunk contract can offer if it is combined with reminders to family members.

Figure A13: **Average take-up for credit products by contractual add-ons among individuals with current NRSP loans**

PANEL A: CURRENT BORROWER



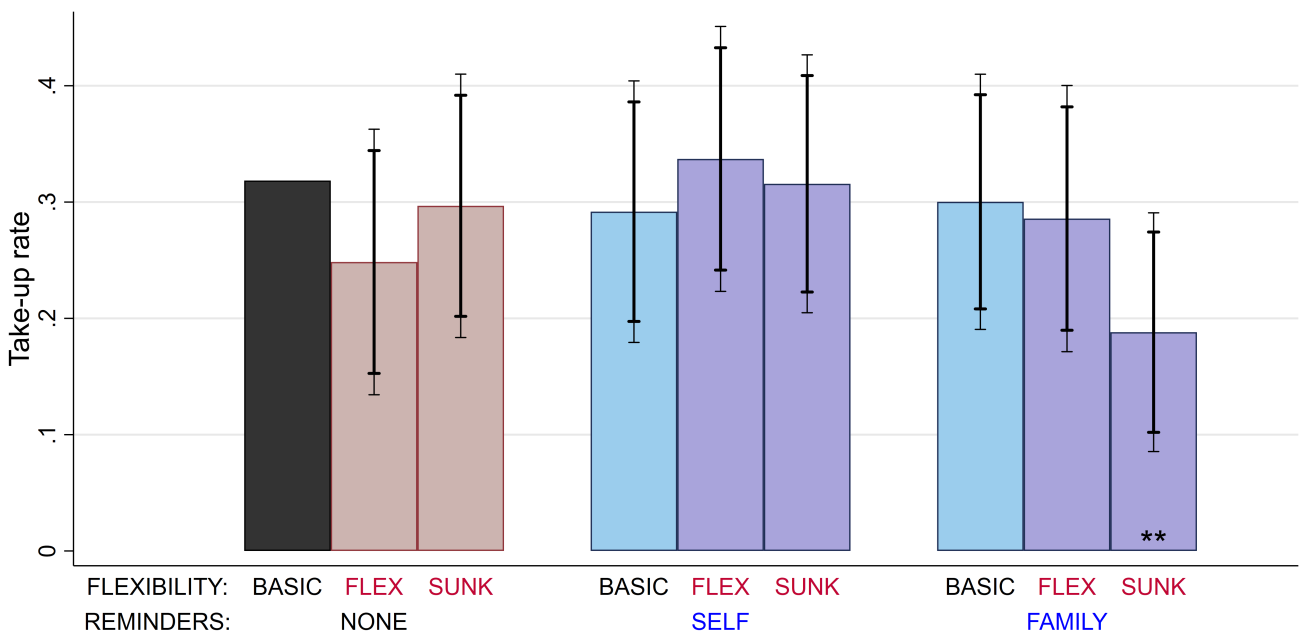
PANEL B: PAST (NOT CURRENT) BORROWER



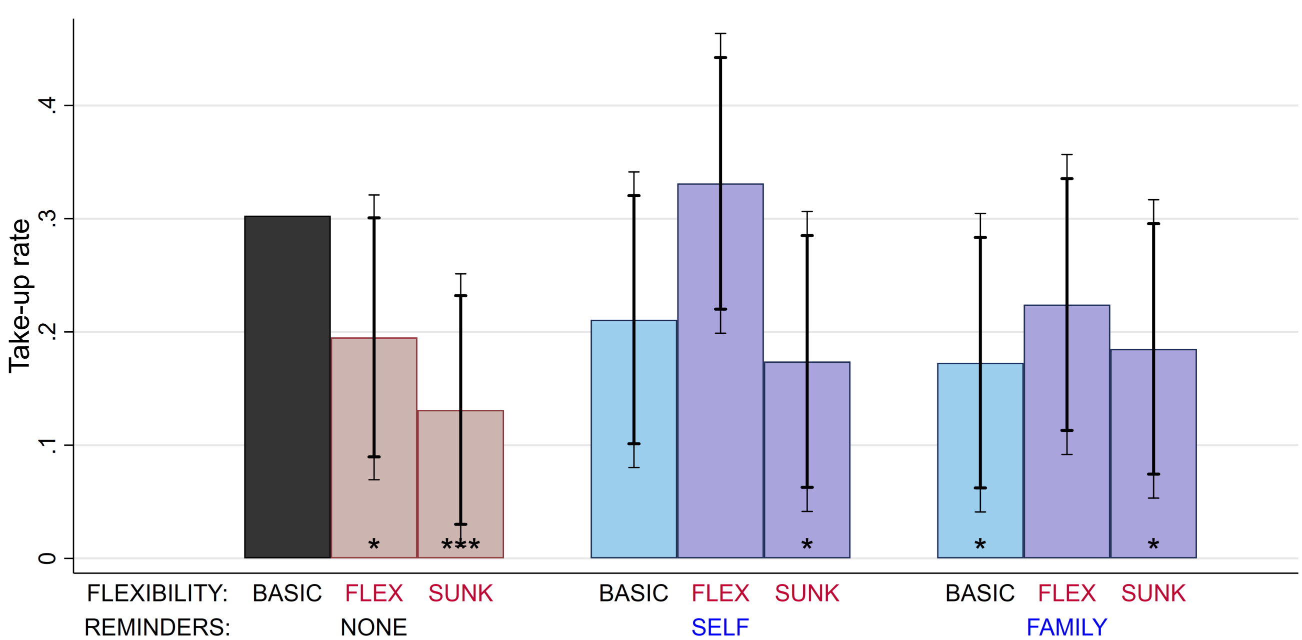
This figure shows the average take-up for the basic credit product (that is, the product with neither the 'sunk' or flexibility ('flex') variation nor the 'self'/'family' reminder variation), and take-up for each of the eight possible add-ons. Error bars show 90% and 95% confidence intervals on the difference 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 A14: **Average take-up for credit products by contractual add-ons among borrowers with long or short NRSP engagement**

panel a: Long engagement ($\geq$ 3 years)



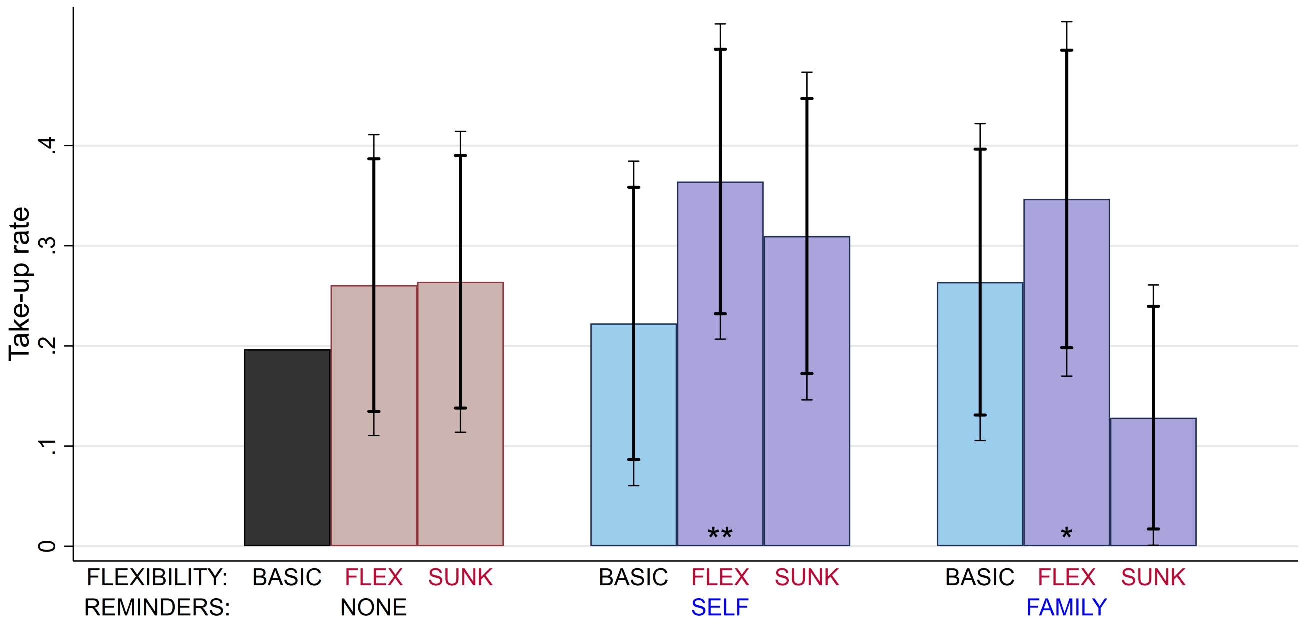
panel b: Short engagement ($<$ 3 years)



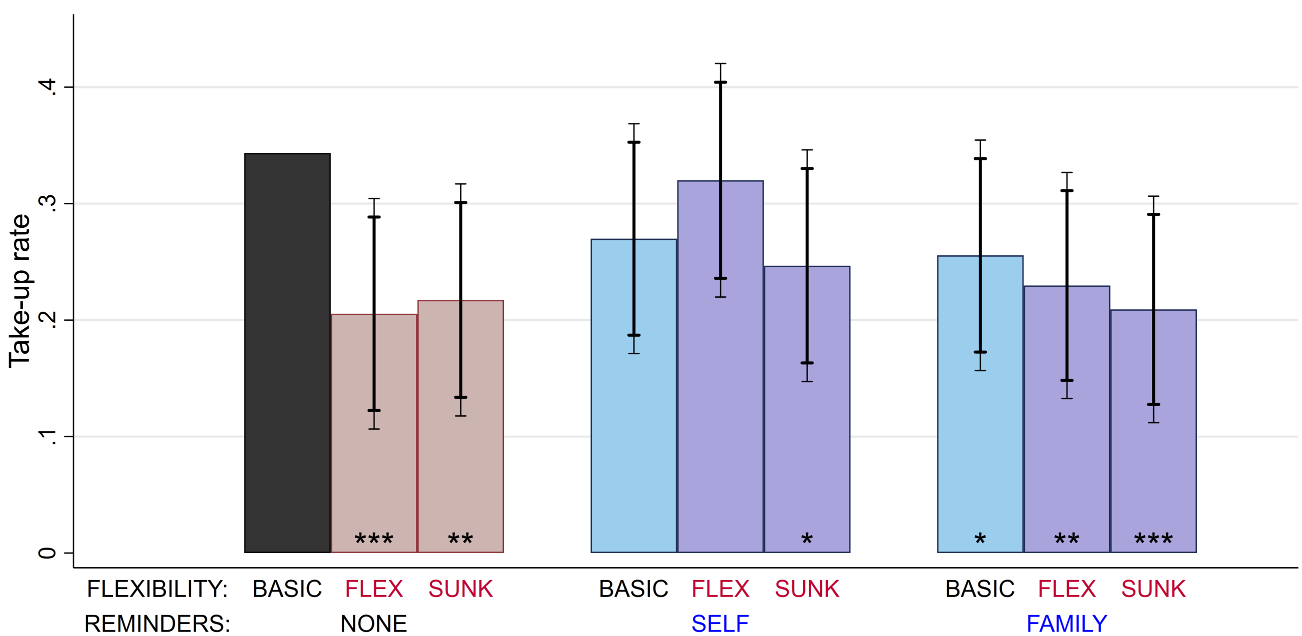
This figure shows the average take-up for the basic credit product (that is, the product with neither the 'sunk' or flexibility ('flex') variation nor the 'self'/'family' reminder variation), and take-up for each of the eight possible add-ons. Error bars show 90% and 95% confidence intervals on the difference 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 A15: **Average take-up for credit products by contractual add-ons among individuals currently participating in a saving committee**

PANEL A: CURRENTLY PARTICIPATING IN COMMITTEE



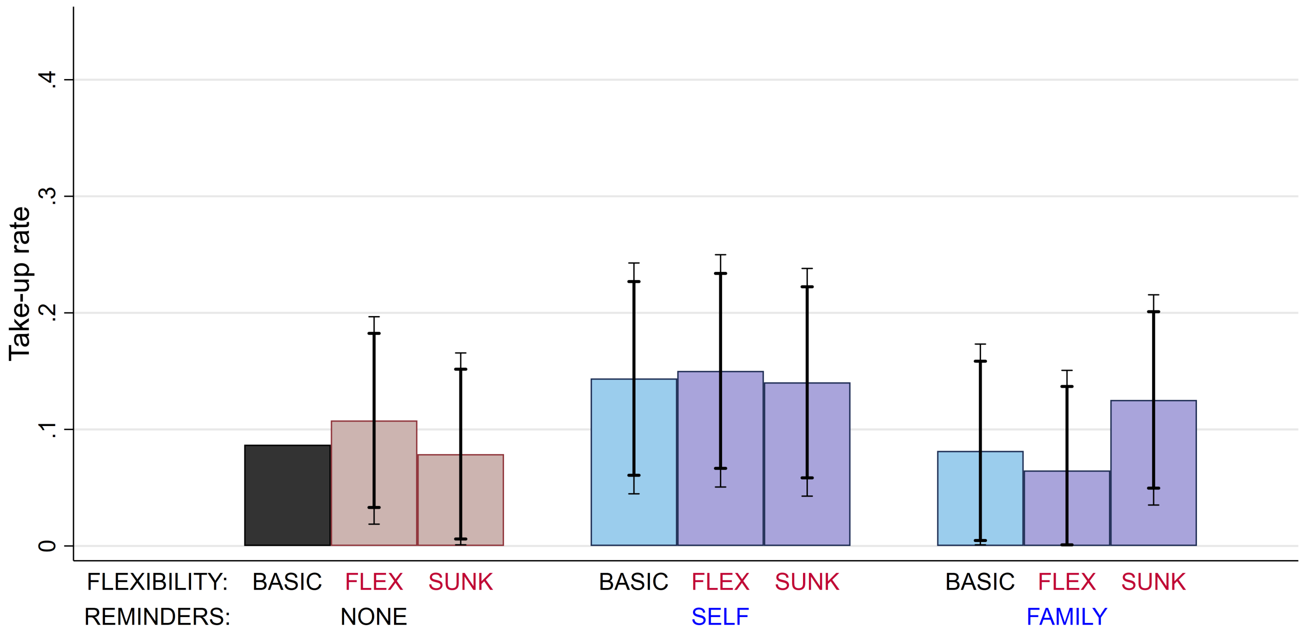
PANEL B: NOT CURRENTLY PARTICIPATING IN A COMMITTEE



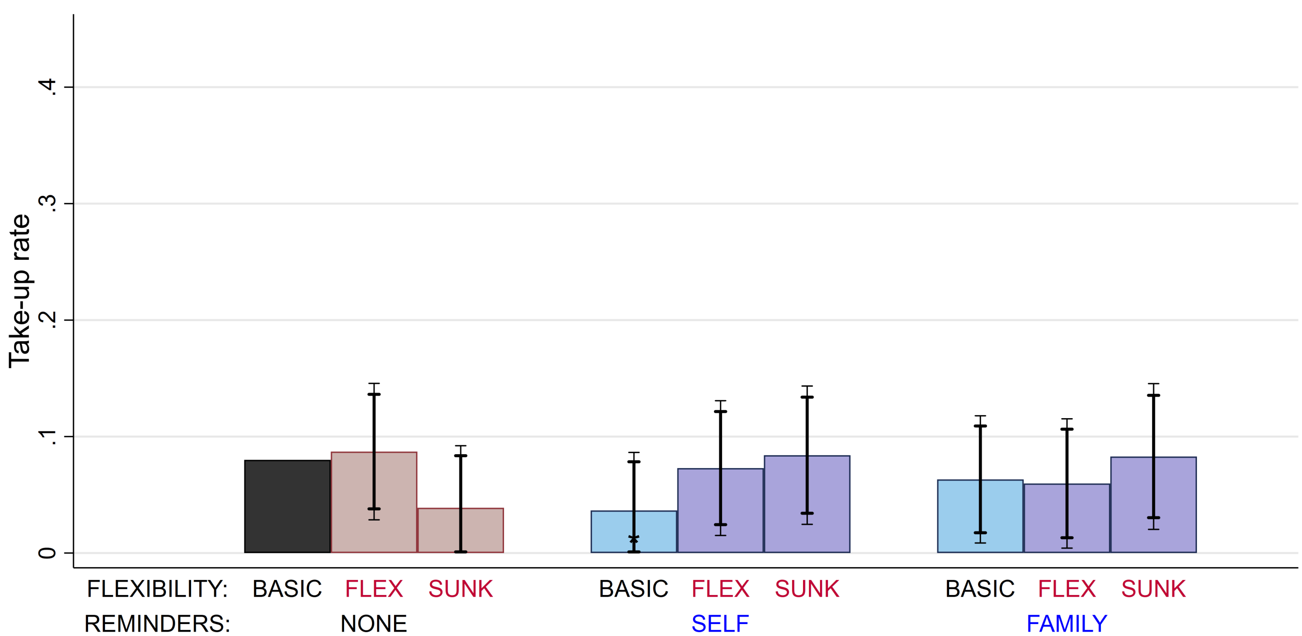
This figure shows the average take-up for the basic credit product (that is, the product with neither the 'sunk' or flexibility ('flex') variation nor the 'self'/'family' reminder variation), and take-up for each of the eight possible add-ons. Error bars show 90% and 95% confidence intervals on the difference 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 A16: **Average take-up for saving products by contractual add-ons among individuals with current NRSP loans**

PANEL A: CURRENT BORROWER



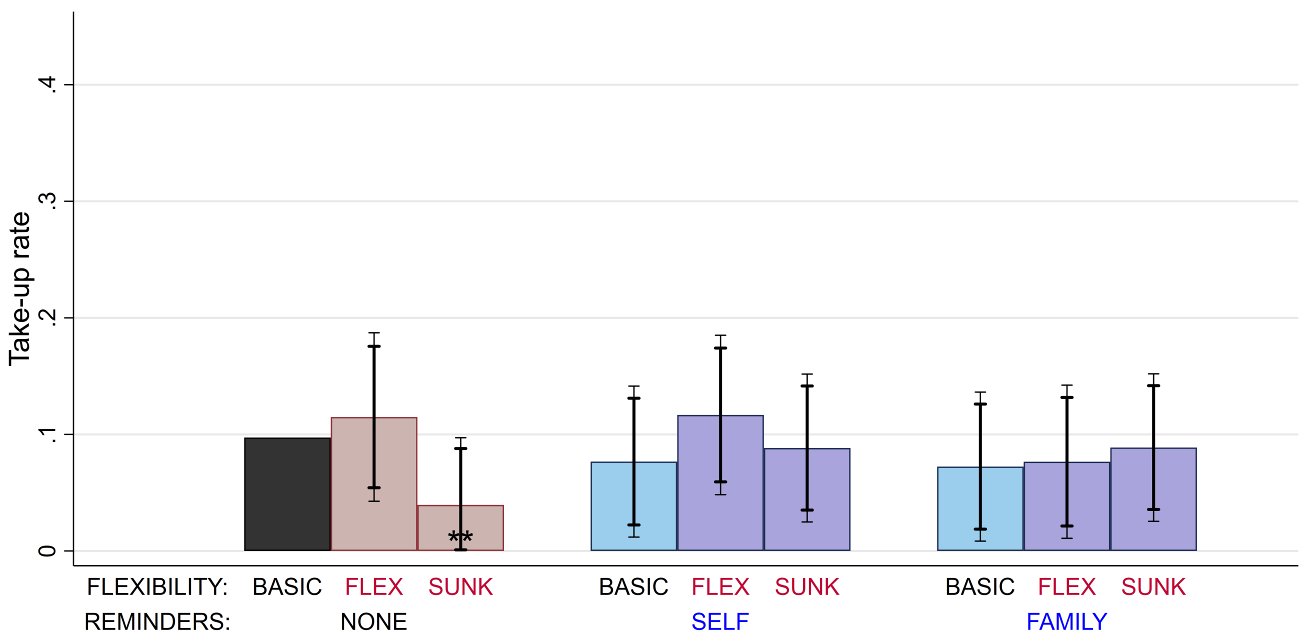
PANEL B: PAST (NOT CURRENT) BORROWER



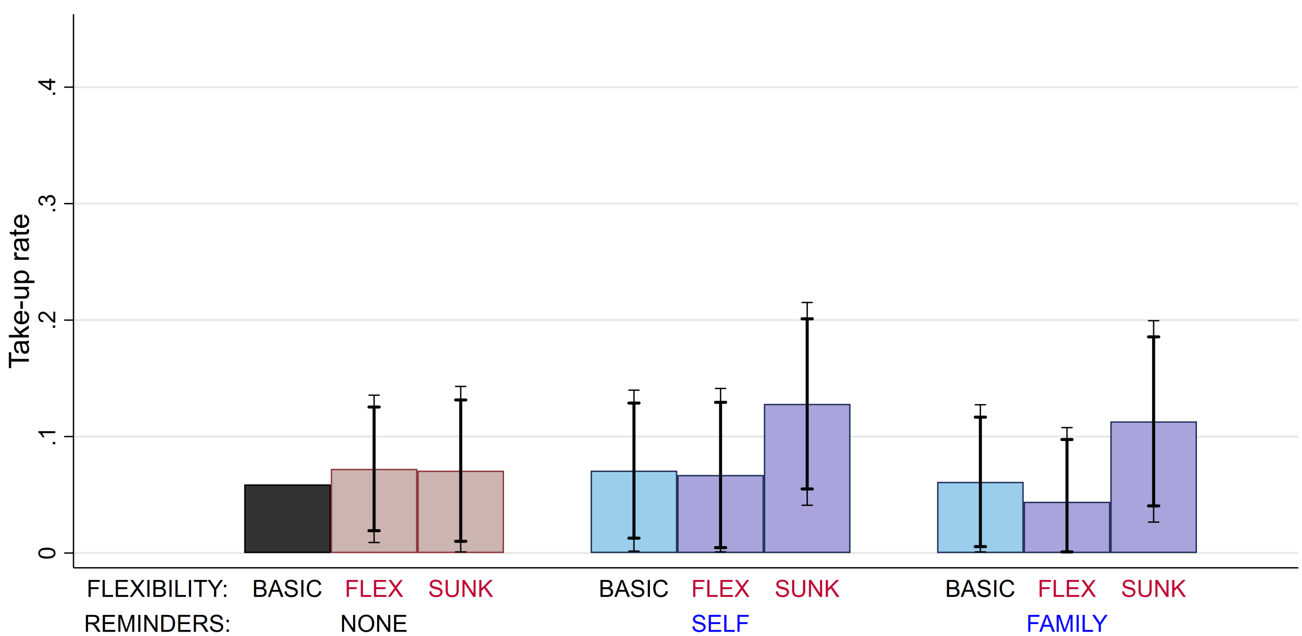
This figure shows the average take-up for the basic saving product (that is, the product with neither the 'sunk' or flexibility ('flex') variation nor the 'self'/'family' reminder variation), and take-up for each of the eight possible add-ons. Error bars show 90% and 95% confidence intervals on the difference 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 A17: **Average take-up for saving products by contractual add-ons among individuals with long or short NRSP engagement**

PANEL A: LONG ENGAGEMENT $\geq$ 3 YEARS



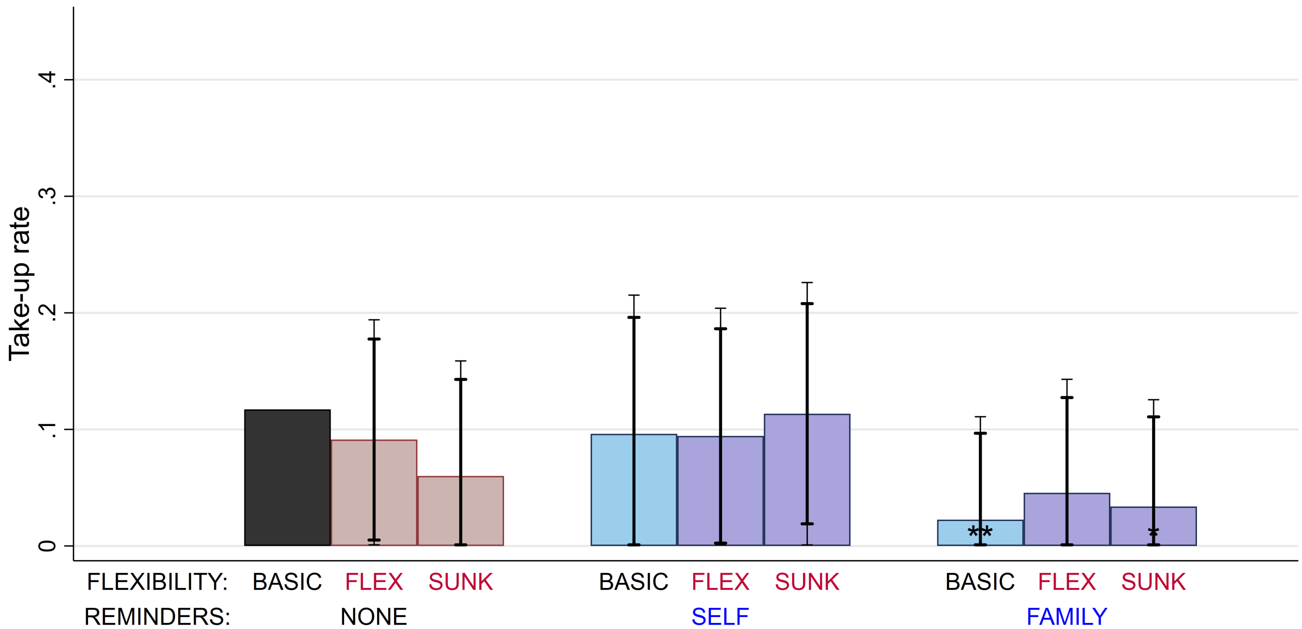
PANEL B: SHORT ENGAGEMENT (< 3 YEARS)



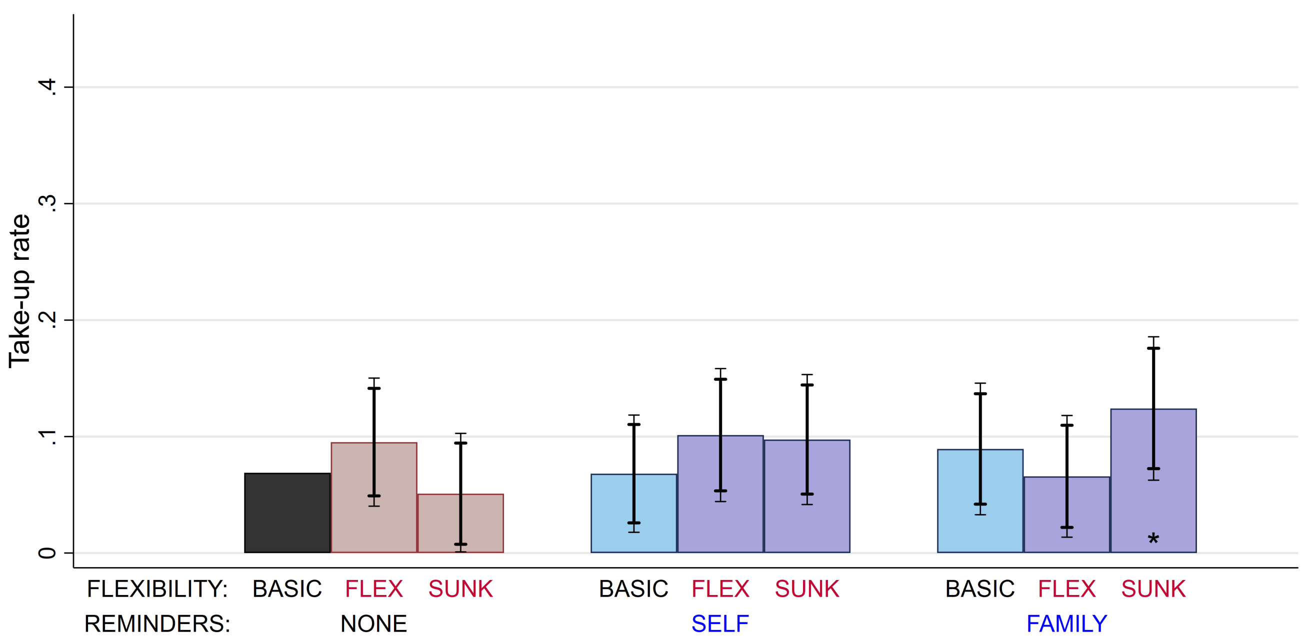
This figure shows the average take-up for the basic saving product (that is, the product with neither the 'sunk' or flexibility ('flex') variation nor the 'self'/'family' reminder variation), and take-up for each of the eight possible add-ons. Error bars show 90% and 95% confidence intervals on the difference 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 A18: **Average take-up for saving products by contractual add-ons among individuals currently participating in a saving committee**

PANEL A: CURRENTLY PARTICIPATING IN COMMITTEE



PANEL B: NOT CURRENTLY PARTICIPATING INN A COMMITTEE



This figure shows the average take-up for the basic saving product (that is, the product with neither the 'sunk' or flexibility ('flex') variation nor the 'self'/'family' reminder 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.

H Defining financial discipline

In this appendix section, we use baseline covariates to characterise sample heterogeneity in financial discipline. We define respondents to have high financial discipline if they face fewer challenges in saving, are able to keep track of finances and commitments, and can be considered to have 'empowered' attitudes. We use the covariates listed in Table A13 to create an index that is increasing in financial discipline' using Principal Component Analysis (PCA).[27] We then create two sub-groups using the first component from the PCA: respondents with scores less than the median index value can be considered to have 'low' financial discipline.

Table A14 summarizes average value of the covariates for the full sample and across the two subgroups. Individuals with high financial discipline are more likely to say that they do not find it hard to save and do not face pressures to share cash at hand, compared to individuals in the low financial discipline group. Individuals in the high financial discipline group are better at keeping track of time and finances, are more likely to follow a strict schedule on finance, act early to avoid forgetting and earmark their funds. Finally, they are also more likely to have a say in household decision making: their opinions are considered when the household makes decisions and similarly, they are less likely to find a woman making decisions without consulting others appropriate.

Table A15 summarizes the average and marginal treatment effects on takeup using a regression with a fully interacted sum of dummies for each treatment, separately for credit and savings contracts, for individuals with 'high' and 'low' financial discipline. Heterogeneity in treatment effects by financial discipline are discussed in detail in section 4.2 of the paper.

[27] Note, Table A13 lists both an indicator for present and future bias, where one is the converse of the other. We include one (future bias) but not both in the PCA estimation.

Table A14: **Financial Discipline: Descriptive characteristics of respondent groups**

Respondents:	All		High discipline		Low discipline	
	Mean	SD	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)	(5)	(6)
Saving challenges:						
Dummy: does not find it hard to save	0.255	(0.436)	0.289	(0.453)	0.222	(0.416)
Dummy: does not face pressure to share cash on hand	0.220	(0.414)	0.248	(0.432)	0.192	(0.394)
Keeping track:						
Dummy: Good at keeping track of time	0.686	(0.464)	0.918	(0.274)	0.454	(0.498)
Dummy: Good at keeping track of finances	0.591	(0.492)	0.869	(0.337)	0.312	(0.464)
Dummy: Follows a strict schedule on finances	0.603	(0.489)	0.873	(0.333)	0.333	(0.471)
Dummy: Follows a tight routine	0.493	(0.500)	0.666	(0.472)	0.320	(0.467)
Dummy: Acts early to avoid forgetting	0.501	(0.500)	0.734	(0.442)	0.268	(0.443)
Dummy: Acts early to avoid forgetting finances	0.464	(0.499)	0.692	(0.462)	0.235	(0.424)
Dummy: Keeps cash earmarked	0.543	(0.498)	0.714	(0.452)	0.373	(0.484)
Dummy: Keeps funds earmarked in accounts	0.130	(0.336)	0.190	(0.392)	0.070	(0.256)
Dummy: Future bias	0.098	(0.298)	0.135	(0.342)	0.061	(0.240)
Women empowerment:						
Share of examples where view always considered	0.671	(0.369)	0.728	(0.346)	0.614	(0.382)
Appropriate for a woman to invest in her business	0.263	(0.440)	0.207	(0.405)	0.319	(0.466)
Appropriate for a women to buy a scarf	0.323	(0.468)	0.255	(0.436)	0.391	(0.488)
PCA index (scores for component 1)	-0.000	(1.563)	1.285	(0.760)	-1.285	(1.001)
Observations	2,416		1,208		1,208	

This table provides the mean and standard deviations of baseline characteristics summarized in Table A13. Specifically, we describe the characteristics of all respondents (columns 1 and 2), respondents with high financial discipline, Group J (columns 3 and 4) and respondents with low financial discipline, Group U (columns 5 and 6). We use Principal Component Analysis to create an index out of listed characteristics and provide the average values of the first score in the last row.

Table A15: Treatment effects on take-up for credit and saving, for individuals with low and high financial discipline

Contract:	Credit		Saving	
Financial discipline	High	Low	High	Low
Panel (a) Average treatment effects				
Basic	0.316***	0.306***	0.096***	0.065***
	(0.043)	(0.051)	(0.027)	(0.024)
Respondent reminder	-0.105*	-0.003	-0.020	0.008
	(0.056)	(0.067)	(0.036)	(0.032)
Family reminder	-0.095*	-0.006	-0.025	-0.001
	(0.057)	(0.067)	(0.034)	(0.033)
Flex	-0.103*	-0.072	0.005	0.022
	(0.058)	(0.065)	(0.036)	(0.033)
Sunk	-0.113**	-0.041	-0.058*	0.008
	(0.055)	(0.068)	(0.031)	(0.036)
Respondent reminder*Flex	0.200**	0.127	0.015	0.005
	(0.079)	(0.089)	(0.051)	(0.047)
Respondent reminder*Sunk	0.128	0.037	0.044	0.049
	(0.078)	(0.091)	(0.044)	(0.051)
Family reminder*Flex	0.127	0.045	-0.029	-0.011
	(0.080)	(0.090)	(0.046)	(0.047)
Family reminder*Sunk	0.024	-0.019	0.035	0.070
	(0.073)	(0.091)	(0.044)	(0.052)
Panel (b) Aggregated average treatment effects				
Flex	0.004	0.052	-0.000	0.027
	(0.032)	(0.036)	(0.019)	(0.021)
Sunk	-0.044	0.002	-0.022	0.019
	(0.030)	(0.036)	(0.018)	(0.021)
Respondent reminder	0.000	-0.010	-0.000	0.020
	(0.032)	(0.036)	(0.020)	(0.019)
Family reminder	-0.066**	-0.033	-0.032*	0.050**
	(0.030)	(0.036)	(0.018)	(0.021)

*Average effects for each of the nine treatment cells are estimated using a regression of the dependent variable on a fully interacted set of dummies for each treatment type, shown in panel (a). The regression is estimated separately for credit and saving contracts, and for individuals with high financial discipline and low financial discipline. The aggregated average effects of the main treatments that are reported in panel (b) are obtained using the 'margins, dydx' command in Stata. They measure the average effect of each of the four treatments aggregated over all treatment cells. The reported number of observations is larger than actual because some subjects said they were not interested in any contract and are thus regarded as refusing 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 (in percentage points) are clustered at the individual level and in parentheses. *** p<0.01, ** p<0.05, * p<0.1.*

I Business and household outcomes

To estimate impacts on business and household outcomes, we use the following ANCOVA specification:

$$y_{i1} = \beta_0 + \beta_1 \cdot T_i + \beta_2 \cdot y_{i0} + \phi_s + \eta_d + \varepsilon_i, \tag{A14}$$

where y_{i1} denotes an outcome variable of interest measured at endline 1, y_{i0} is the baseline value of y_{i1}, ϕ_s are strata dummies, and η_d are district fixed effects. We cluster errors at the household level. (Note that equation A14 estimates the combined average impact of being assigned to treatment; our conclusions remain unchanged if we disaggregate treatments; we show this analysis in section 5 of the online appendix for the Phase 2 Pre-Analysis Plan.)

The variable T_i takes two interpretations, depending on the specification. First, we denote T_i as assignment to treatment; in that case, we estimate equation A14 using OLS, and interpret $\hat{\beta}_1$ as the ITT. Second, we denote T_i as take-up. This takes four possible values, which depend on whether the subjects takes up the contract in 0, 1, 2 or 3 waves. In this case, we calculate average take-up at the individual level, and instrument this using assignment to treatment and to contractual terms; we then interpret $\hat{\beta}_1$ as providing the LATE, normalized for a case where a respondent takes up in all three product waves.[28]

Outcomes are divided into two broad categories: business outcomes and household finance and consumption. In Table A16, we report business outcomes; we collect ITT and LATE estimates for both phase 1 and phase 2 samples. In principle, our MFI partner lends for business purposes: it is therefore of primary interest whether our commitment saving contract is able to improve business investment and performance.

[28] To instrument average take-up, we proceed as follows. First, for each wave s, we estimate the predicted take-up of individual i based on the different types of treatments i was exposed to in that wave – i.e., payment week, negative or positive interest, reminders, and 'flex' or 'sunk' treatment. This is achieved using the same regression that was used in generating Table 3 for the six combinations of payment week and interest rate – except that it is estimated separately for each wave. This generates a predicted take-up for each product wave. The sum over all three waves is then used as instrument for T_i when estimating. For automatic refusers, we do not have a specific payment week or interest rate on which to base our prediction – since these subjects refused the contract before cards were drawn. To circumvent this issue, we ascribe to each of these observations the average predicted take-up associated with their commitment and reminder treatment, assuming an average interest rate and payment week. In practice this is achieved, as before, by generating six observations for each refuser, one for each combination of payment week and interest rate, and ascribing a weight of 1/6 to each of these observation when estimating the predicting equation.

Table A16: **Summary of ITT and LATE estimates of business outcomes**

	Phase 1			Phase 2		
	Control mean	ITT	LATE	Control mean	ITT	LATE
Runs a business	0.606	0.016	0.047	0.125	-0.009	-0.009
		(0.019)	(0.059)		(0.013)	(0.036)
Number of businesses	1.116	0.087*	0.072	0.156	-0.002	-0.006
		(0.047)	(0.153)		(0.017)	(0.047)
Value of capital invested in business	7803	610	2310	2023	-301	-184
		(607)	(1723)		(371)	(1034)
Value of monthly sales	8184	709	3406*	1237	-42	8
		(519)	(1764)		(188)	(526)
Value of monthly expenses	6228	152	2571	502	-42	56
		(452)	(1627)		(82)	(230)
Monthly profit (sales - expenses)	1871	737**	716	665	25	-9.5
		(353)	(1134)		(112)	(314)
Monthly profit (self-reported)	2933	518	1079	869	-23	-78
		(329)	(1061)		(130)	(368)
Observations		789	789		1991	1991

This table reports regression estimates of equation A14. We report standard errors under each coefficient in parentheses. All values are in Pakistani rupees. Monthly self-reported profits include the imputed values of business goods consumed. Confidence: $* \leftrightarrow p < 0.1$; $** \leftrightarrow p < 0.05$; $*** \leftrightarrow p < 0.01$.

We find almost no significant effect on business and household outcomes of having been offered our treatment; this is consistent with a growing body of evidence on the effects of microfinance (see, for example, Meager (2022) and Meager (2019)). In the phase 1 sample, 60% of respondents have a business. Among these subjects, we find generally positive point estimates on business performance, as measured by investment, sales, or profit. But these point estimates are in general not statistically significant. Two of the ITT coefficients are above the 10% significance level, but only one of the LATE coefficients is significant, and it is for another dependent variable. In contrast, among the phase 2 sample, estimated treatment effects are small in magnitude and never significant. This may be because a much smaller proportion (12.5%) of these households have a business at baseline.

Results for household material outcomes are presented in Table A17. We find no significant effect on household consumption or household income (the latter being measured only in the phase 2 sample). In the phase 1 sample, we find a large and significant LATE coefficient on total household assets and total individual assets. This encouraging result is, however, negated in the phase 2 sample where we find a large but negative LATE effect on total household assets.

The bottom part of Table A17 relates to household finances. We see that 75% of control subjects in the phase 1 sample save in a 'committee'. The proportion is smaller in phase 2: 16.6%. We find a positive and significant LATE effect on participation in a committee, but given that the corresponding ITT coefficient is essentially 0, it is unclear how much faith to put in this result. We also find a positive LATE for participation in a committee among phase 2 respondents, but the effect is not statistically significant. The last row of Table A17 reports results for the total debt of the respondent. Our commitment saving product should have helped participants reduce their stock of debt. We find little evidence of this. Among phase 1 subjects, ITT and LATE coefficients are positive but not significant, while among phase 2 subjects the ITT is negative and significant but the LATE coefficient is not.

Further, we measure the impact on a short list of indicators using higher frequency information from phone surveys conducted at the end of each experiment wave. Table A27 (appendix) summarizes the results for business and household outcomes. We find generally insignificant effects. There are no significant effect on the likelihood of running a business, the number of businesses or on the value of capital invested in the business in the last one month. Treated participants have higher consumption and lower debt but this difference is never significant.

Finally, we check for heterogeneity in these effects, by the quintiles of take-up rates estimated earlier. Specifically, we estimate equation A14 separately for each of those quintiles, for all of the outcomes in Table A17 and Table A17. We use the bootstrap method of Chernozhukov et al. (2018), both for obtaining point estimates and for inference. We do not find heterogeneous effects (it is not the case, for example, that some quintiles are benefiting from being offered the treatment while others are not). In Figures A19 and A20, we show this for two outcomes of particular interest: business investment and household consumption (for Phase 1 and for Phase 2 respectively).

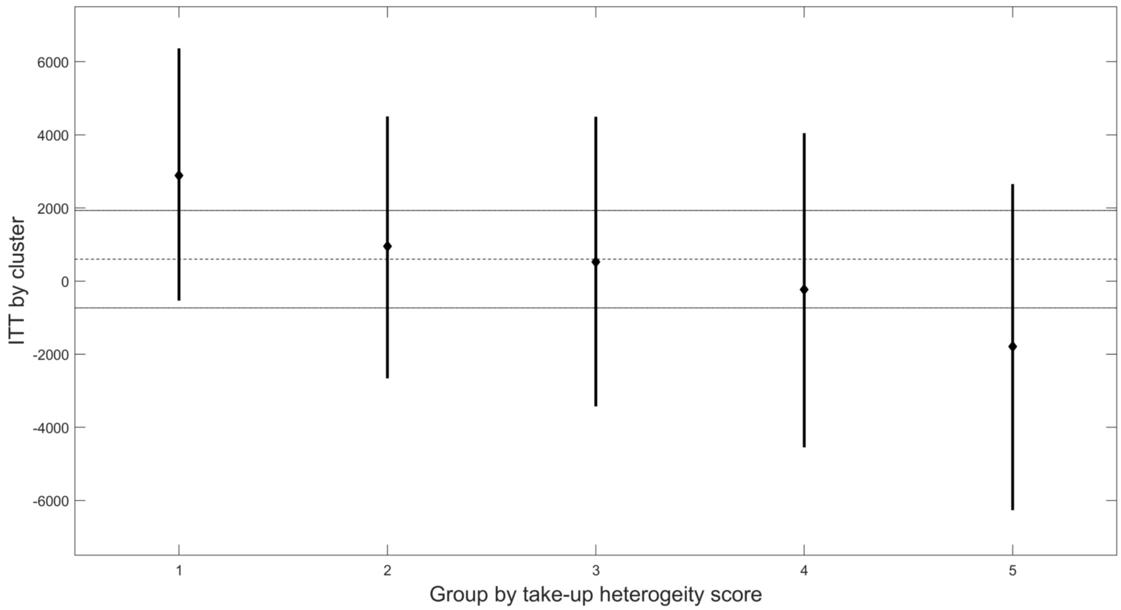
Table A17: **Summary of ITT and LATE estimates of household material outcomes**

	Phase 1			Phase 2		
	Control mean	ITT	LATE	Control mean	ITT	LATE
Household monthly consumption	24706	599	1270	18814	355	898
		(810)	(2599)		(582)	(1626)
Household monthly income	n.a.			21974	-165	2998
					(681)	(2021)
Value of household assets	46041	3106	33310**	40821	-3446	-12000*
		(3990)	(15906)		(2818)	(6948)
Value of subject's assets	23151	3007	18328*	n.a.		
		(2470)	(9615)			
Participates in a committee	0.750	-0.002	0.319**	0.166	0.005	0.102
		(0.046)	(0.148)		(0.018)	(0.056)
Total debt of respondent	13300	1987*	4911*	11587	-1670*	182
		(1030)	(2929)		(901)	(2366)
Observations		789	789		1991	1991

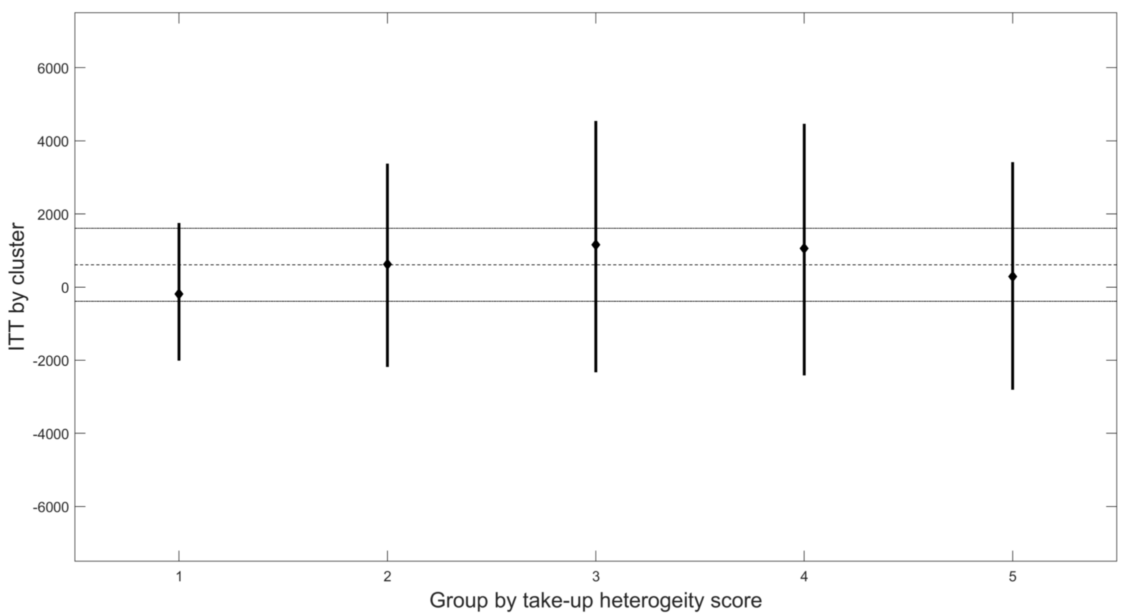
This table reports regression estimates of equation A14. We report standard errors under each coefficient in parentheses. All values are in Pakistani rupees. Confidence: $* \leftrightarrow p < 0.1$; $** \leftrightarrow p < 0.05$; $*** \leftrightarrow p < 0.01$.

Figure A19: **Group Average Treatment Effects (sorted by take-up propensity): Consumption and investment (Phase 1)**

HOUSEHOLD MONTHLY CONSUMPTION



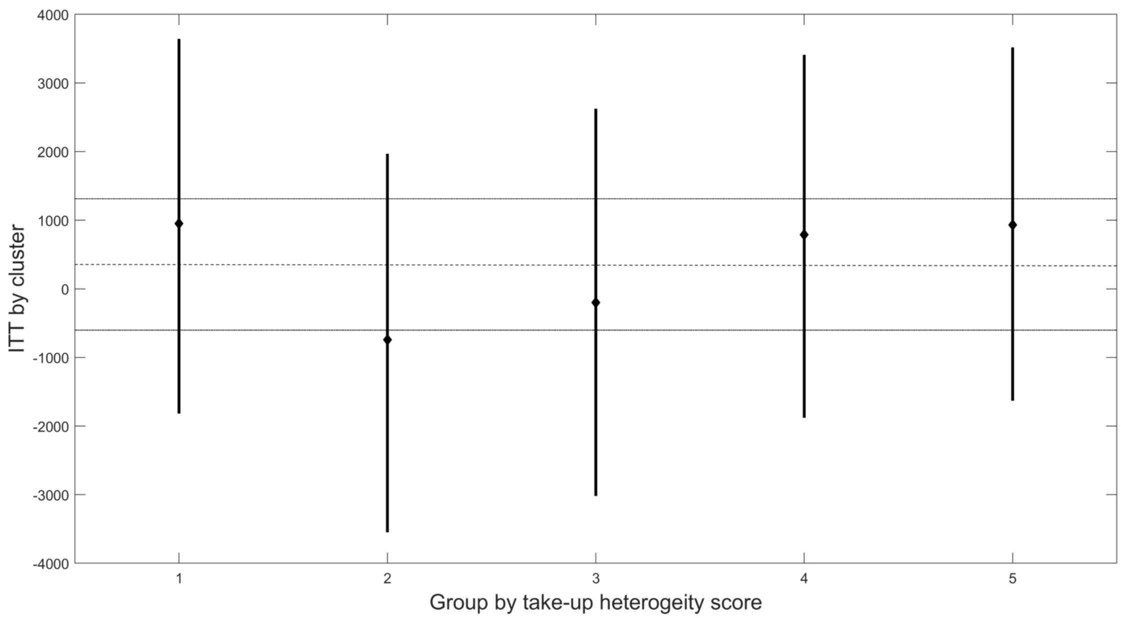
BUSINESS INVESTMENT



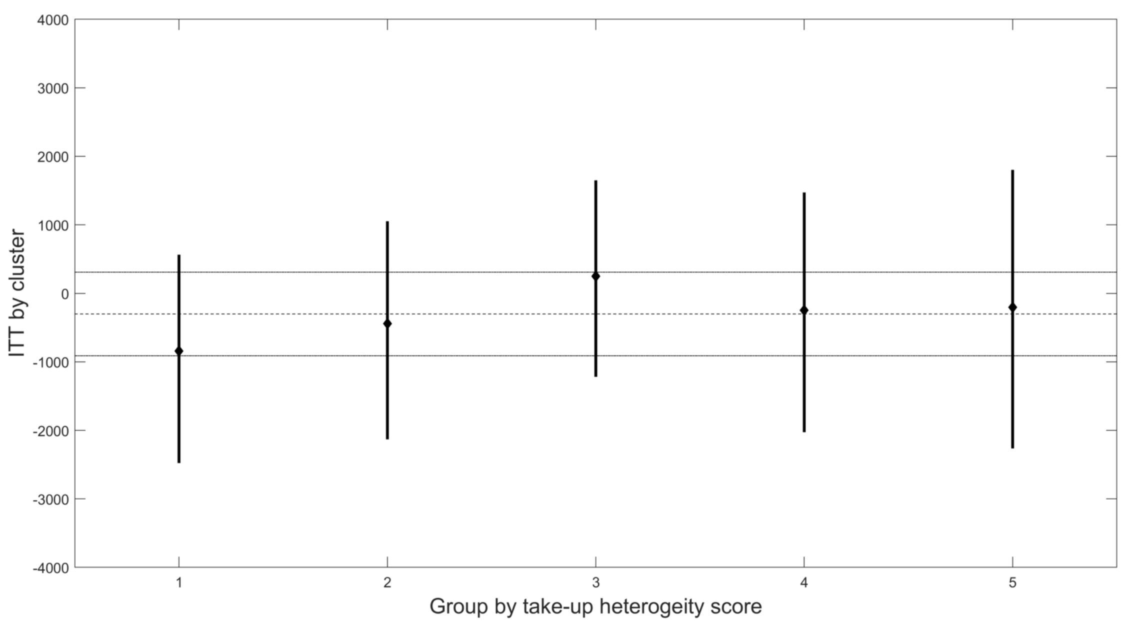
This figure shows the Group Average Treatment Effects, sorted by the take-up propensity estimated in the main text. For each of the five groups separately, we estimate equation A14; the graph shows the estimated ITT and 90% confidence intervals (both formed using the bootstrap methodology proposed by Chernozhukov et al. (2018)). The horizontal lines show the point estimates and 90% confidence intervals for the ITT across the sample (as earlier reported in Table A16 and Table A17).

Figure A20: **Group Average Treatment Effects (sorted by take-up propensity): Consumption and investment (Phase 2)**

HOUSEHOLD MONTHLY CONSUMPTION



BUSINESS INVESTMENT



This figure shows the Group Average Treatment Effects, sorted by the take-up propensity estimated in the main text. For each of the five groups separately, we estimate equation A14; the graph shows the estimated ITT and 90% confidence intervals (both formed using the bootstrap methodology proposed by Chernozhukov et al. (2018)). The horizontal lines show the point estimates and 90% confidence intervals for the ITT across the sample (as earlier reported in Table A16 and Table A17).

J Additional tables and figures

Table A18: **Description of the sample: Phase 1**

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

This table provides summary statistics for sample characteristics in Phase 1. We report p-values for balance of (i) assignment to treatment rather than control, and (ii) assignment to contract terms. For treatment balance, the test is conducted by regressing each variable on a treatment dummy. For contract terms, we regress each variable on dummies for interest rate and lumpsum week; we then test for joint significance. The p-value for joint significance of all variables against treatment status is 0.208. Profit (1) is calculated from disaggregated business income and expenditure information. Profit (2) is the difference of aggregated monthly business income and expense. Household assets include electrical appliances, vehicles, and livestock.

Table A19: **Description of the sample: Phase 2**

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

This table provides summary statistics for sample characteristics in Phase 2. We report p-values for balance of (i) assignment to treatment rather than control, and (ii) assignment to contract terms. For treatment balance, the test is conducted by regressing each variable on a treatment dummy. For contract terms, we regress each variable on dummies for interest rate and lumpsum week; we then test for joint significance. The p-value for joint significance of all variables against treatment status is 0.912. Profit (1) is calculated from disaggregated business income and expenditure information. Profit (2) is the difference of aggregated monthly business income and expense. Household assets include electrical appliances, vehicles, and livestock.

Table A20: **Description of the sample by treatment groups: Phase 2**

	Control	Basic-None	Basic-Family	Basic-Resp	Flex-None	Flex-Family	Flex-Resp	Sunk-None	Sunk-Family	Sunk-Resp
Dummy: participates in a committee	0.2	0.1	0.2	0.1	0.2	0.1	0.2	0.2	0.2	0.2
Total amount owed by individual (PKR)	12100.5	13506.1	10674.1	13138.9	12089.4	9502.3	13235.7	14143.7	11104.3	11059.3
Total household consumption last month (PKR)	19739.9	19820.5	19083.6	19013.3	18040.1	18947.8	19514.8	19506.1	18172.5	20468.6
Total monthly income (PKR)	19772.8	20313.2	19553.8	19613.2	19501.5	19469.0	20568.5	20748.5	20276.8	20117.4
Total value of assets owned by household (PKR)	35391.7	36084.8	33764.9	38051.8	33766.8	36944.9	35598.3	34261.0	37422.4	34420.5
Dummy: runs a business	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Total number of businesses owned	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.1
Total capital invested in business(es)	2084.7	2003.5	1859.7	2036.1	1580.5	2992.0	3200.1	2683.1	1827.2	1755.0
Total monthly sales of the business (PKR)	1210.9	1246.2	1148.7	1160.8	1185.6	1327.3	1587.7	1251.2	1241.5	841.6
Total monthly expense of the business (PKR)	559.0	598.5	484.9	413.0	599.9	606.3	751.6	566.8	484.1	435.1
Total monthly profit(1) of the business (PKR)	566.3	664.5	688.9	704.2	580.8	540.2	829.6	620.7	764.7	307.5
Total monthly profit(2) of the business (PKR)	790.4	788.6	816.8	742.2	780.5	886.9	937.4	764.8	815.2	545.1
Dummy: finds it hard to save	0.7	0.8	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.7
Index: opinions taken into account in household decisions	0.0	0.0	0.1	0.1	0.1	0.1	0.1	0.0	0.1	-0.1
Dummy: faces pressure to share cash on hand	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8
Index: needs to ask permission for **making decisions**	0.0	0.0	0.0	0.1	0.1	0.1	0.0	0.0	0.1	0.0
Age (years)	38.8	40.0	38.4	39.7	38.6	39.8	39.4	38.8	38.7	39.3
Dummy: is currently married	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.8	0.8	0.8
Level of education	4.2	3.7	4.7	3.5	5.1	4.2	4.3	4.7	4.3	4.4
Dummy: can read and write	0.5	0.5	0.5	0.4	0.6	0.5	0.5	0.6	0.5	0.5
Number of children	3.5	3.5	3.4	3.7	3.2	3.6	3.7	3.2	3.3	3.3
Dummy: is the household head	0.2	0.2	0.2	0.2	0.1	0.2	0.2	0.2	0.2	0.2
N	602	197	199	204	202	198	204	201	207	202

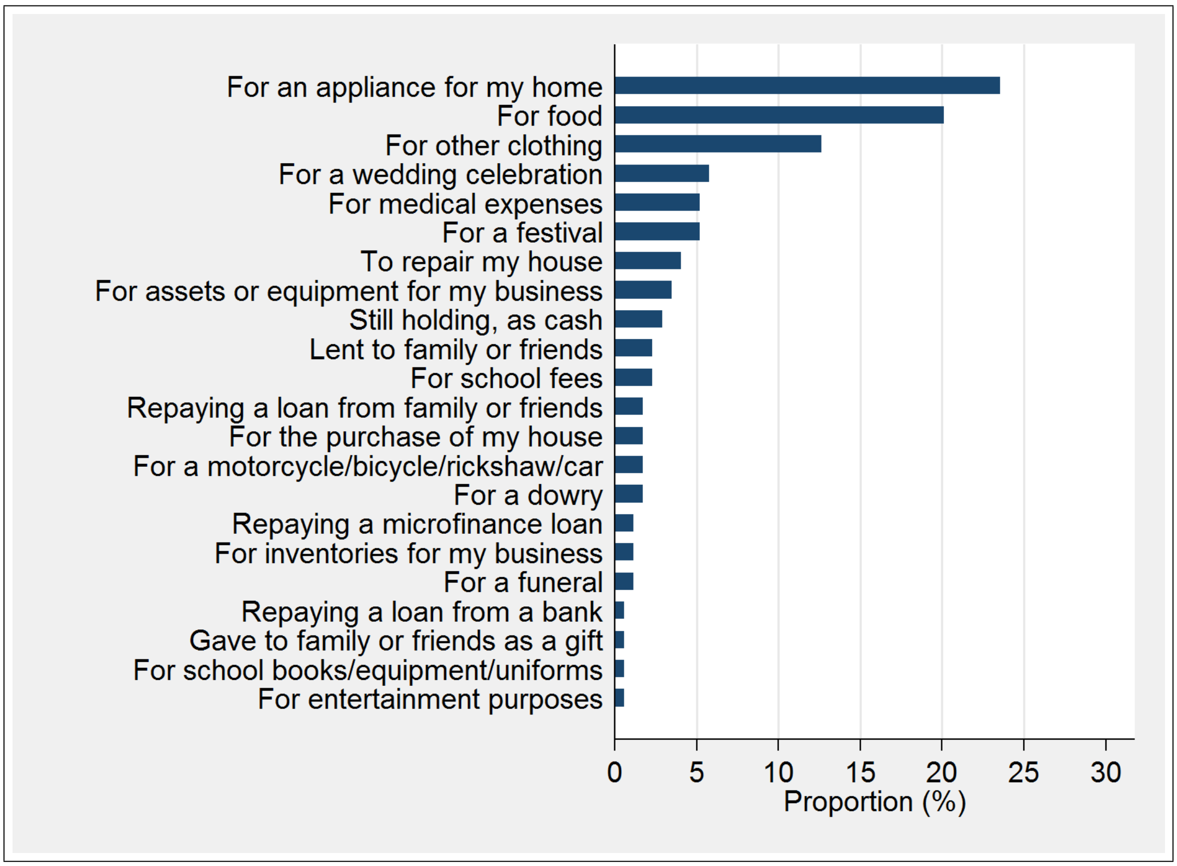
This table provides basic summary statistics for sample characteristics across the 3X3 factorial design with controls in Phase 2. For each variable, we report the mean value for control and each of the treatment group.

Table A21: **Proportion of treated who refused the product offered before the contract offer cards were drawn**

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

This table shows the proportion of the treated who refused the product offered before the cards determining the contract terms were drawn. 'All waves' refers to the proportion of the treated who refused the product before contractual terms were drawn in any one of the three waves. 'Three times' is the proportion of the treated who refused the product in all three waves.

Figure A21: **Spending of the lumpsum: Most common categories**



This figure shows the most common categories listed by respondents in describing how they used the lumpsum payments from Phase 2; reported proportions show the proportion of respondents listing each separate category.

Table A22: **Average take-up by contractual add-ons: Credit contracts**

Panel 1: Average take-up in each treatment combination

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

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

Panel 2: Pairwise take-up comparisons

A. Difference from basic contract with no reminders

	Flex	Basic	Sunk
No reminders	-8.9%**	*reference*	-8.1%*
Reminder to self	2.3%	-5.4%	-4.7%
Reminder to family	-5.5%	-5.3%	-12.5%***

B. Difference from basic contract

	Flex	Basic	Sunk
No reminders	-8.9%**	*reference*	-8.1%*
Reminder to self	7.7%*	*reference*	0.7%
Reminder to family	-0.2%	*reference*	-7.2%*

C. Difference from no reminder contract

	Flex	Basic	Sunk
No reminders	*reference*	*reference*	*reference*
Reminder to self	11.2%***	-5.4%	3.4%
Reminder to family	3.4%	-5.3%	-4.4%

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

Table A23: **Average take-up by contractual add-ons: Savings contracts**

Panel 1: Average take-up in each treatment combination

	Flex	Basic	Sunk
No reminders	9.4%	8.2%	5.4%
Reminder to self	9.9%	7.4%	10.1%
Reminder to family	6.1%	6.8%	9.9%

Joint equality test (p-value): 0.321

Panel 2: Pairwise take-up comparisons

A. Difference from basic contract with no reminders

	Flex	Basic	Sunk
No reminders	1.2%	*reference*	-2.9%
Reminder to self	1.7%	-0.8%	1.9%
Reminder to family	-2.1%	-1.4%	1.7%

B. Difference from basic contract

	Flex	Basic	Sunk
No reminders	1.2%	*reference*	-2.9%
Reminder to self	2.5%	*reference*	2.7%
Reminder to family	-0.7%	*reference*	3.1%

C. Difference from no reminder contract

	Flex	Basic	Sunk
No reminders	*reference*	*reference*	*reference*
Reminder to self	0.5%	-0.8%	4.7%
Reminder to family	-3.3%	-1.4%	4.5%

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

Table A24: **Average take-up by contractual add-ons: Pooled across credit and saving**

Panel 1: Average take-up in each treatment combination

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

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

Panel 2: Pairwise take-up comparisons

A. Difference from basic contract with no reminders

	Flex	Basic	Sunk
No reminders	-4.1%	*reference*	-5.6%**
Reminder to self	2.1%	-3.2%	-1.6%
Reminder to family	-3.9%	-3.5%	-5.5%**

B. Difference from basic contract

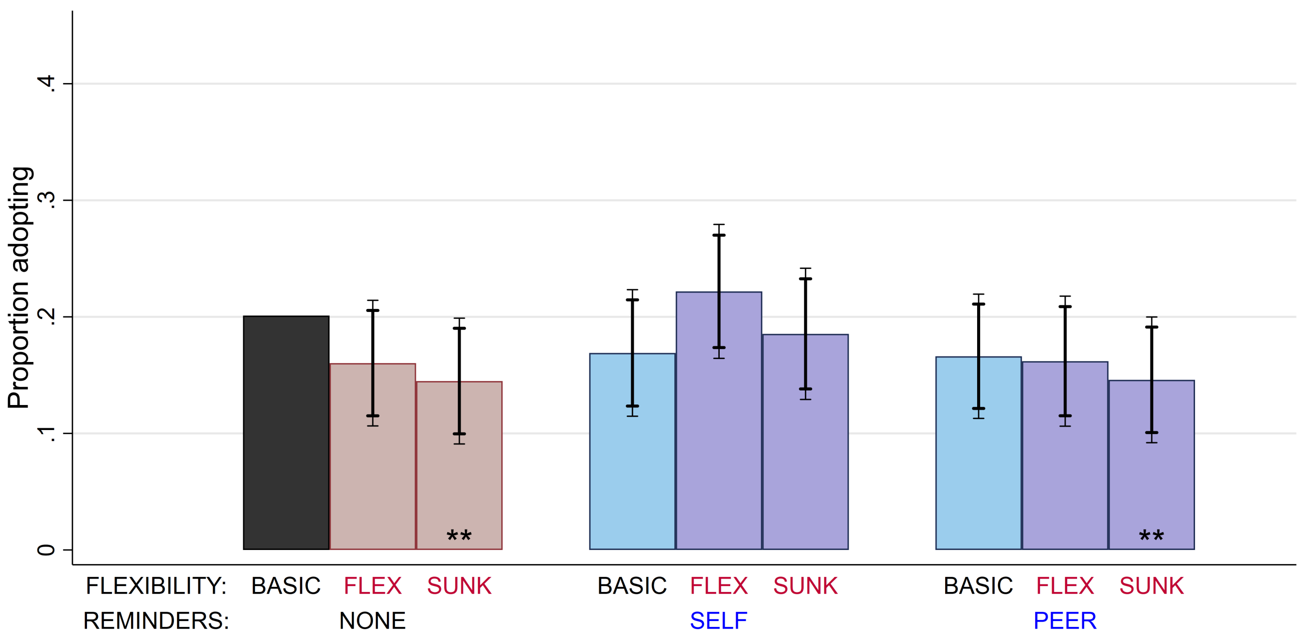
	Flex	Basic	Sunk
No reminders	-4.1%	*reference*	-5.6%**
Reminder to self	5.3%**	*reference*	1.6%
Reminder to family	0.4%	*reference*	-2.0%

C. Difference from no reminder contract

	Flex	Basic	Sunk
No reminders	*reference*	*reference*	*reference*
Reminder to self	6.2%**	-3.2%	4.0%
Reminder to family	0.2%	-3.5%	0.1%

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

Figure A22: **Average take-up by contractual add-ons: Pooled across credit and saving**



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

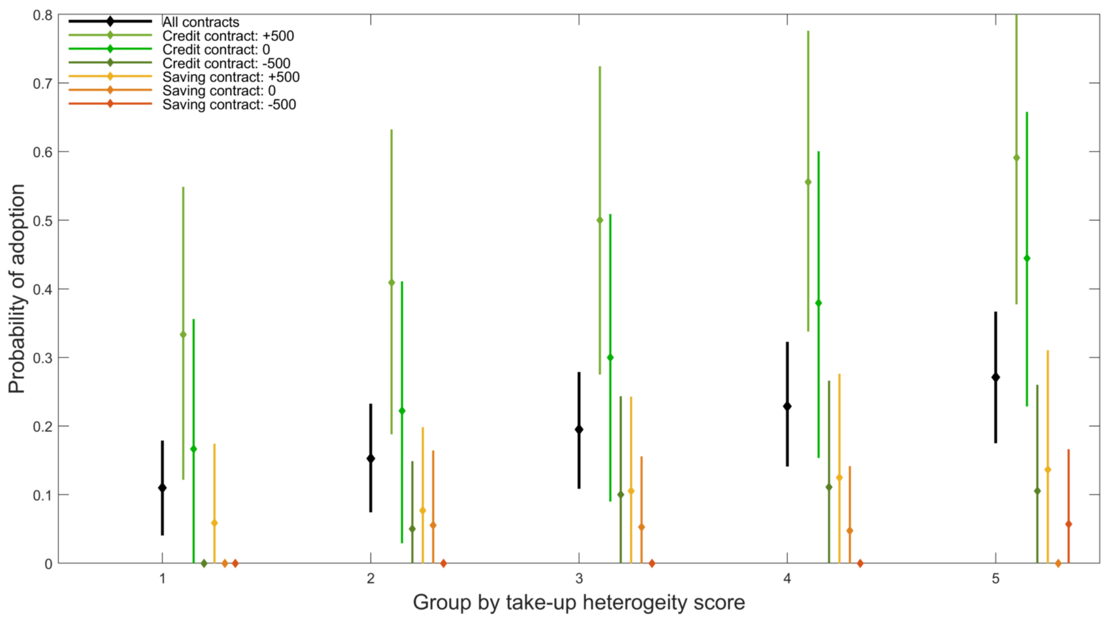
Table A25: **Cluster Analysis: Description of extreme groups (all covariates: Phase 2)**

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

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

Figure A23: **Group Average Treatment Effects (sorted by take-up propensity)**

TAKE-UP BY CONTRACTUAL TERMS (PAYMENT AND TIMING)



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

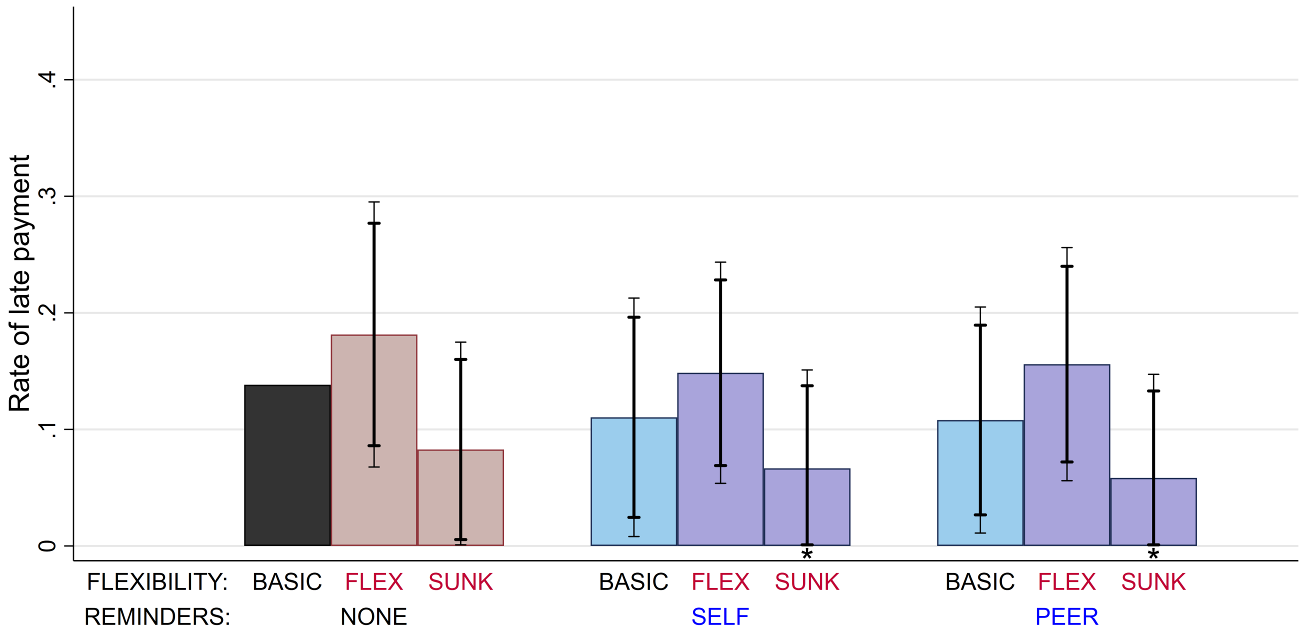
Table A26: **Cluster Analysis: Description of extreme groups (all covariates): Phase 1**

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

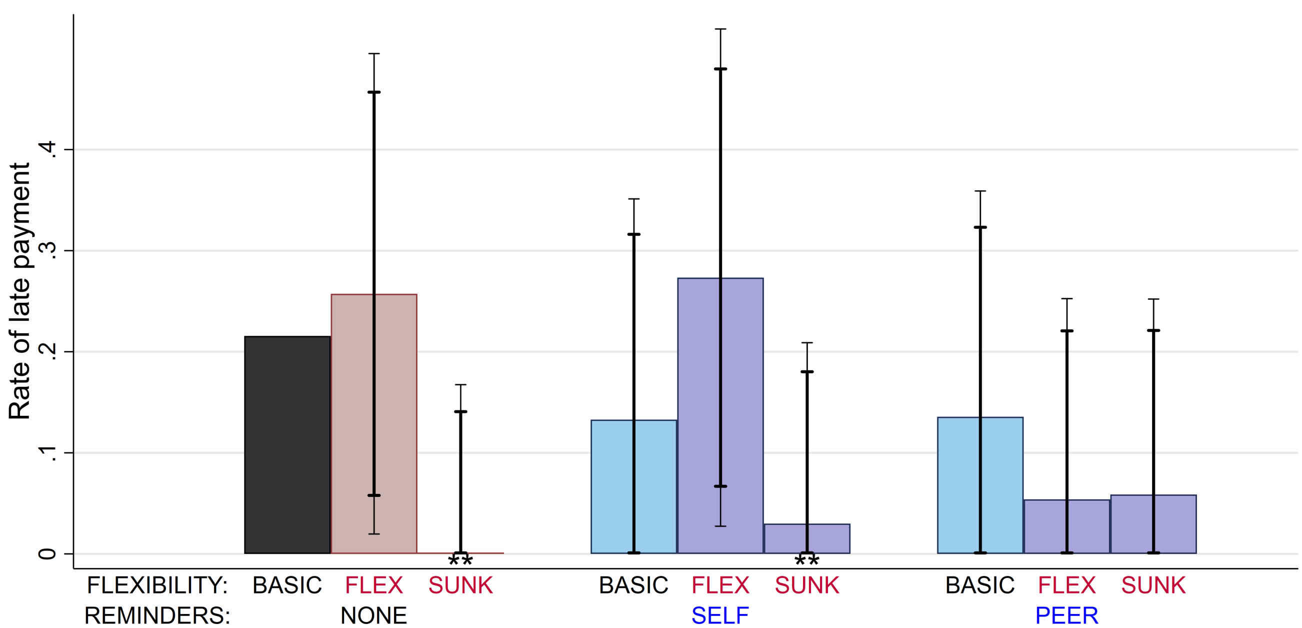
This table provides a cluster analysis of all the baseline covariates used for the machine learning analysis for Phase 1. Specifically, we describe the characteristics of those respondents in the 'most affected' and 'least affected' groups, defined in terms of estimated probability of adopting. We provide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

Figure A24: **Late payment by contractual add-ons (including respondent using the 'flex' option)**

PANEL A: CREDIT CONTRACTS



PANEL B: SAVING CONTRACTS



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. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic contract. Stars indicate a significant difference from the basic contract.

Table A27: **Summary of ITT and LATE estimates of mobile phone data: Phase 2 experiment**

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

This table reports regression estimates of equation A14. We report standard errors under each coefficient in parentheses. All values are in Pakistani rupees. Confidence: $* \leftrightarrow p < 0.1$; $** \leftrightarrow p < 0.05$; $*** \leftrightarrow p < 0.01$.

References

Attanasio, O. and E. Pastorino (2020). Nonlinear pricing in village economies. *Econometrica 88*(1), 207–263.

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.

Bruhn, M. and D. McKenzie (2009). In Pursuit of Balance: Randomization in Practice in Development Field Experiments. *American Economic Journal: Applied Economics 1*(4), 200–232.

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.

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

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*.

Dean, M. and A. Sautmann (2021). Credit constraints and the measurement of time preferences. *Review of Economics and Statistics 103*(1), 119–135.

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

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.

Meager, R. (2019). Understanding the average impact of microcredit expansions: A bayesian hierarchical analysis of seven randomized experiments. *American Economic Journal: Applied Economics 11*(1), 57–91.

Meager, R. (2022). Aggregating distributional treatment effects: A bayesian hierarchical analysis of the microcredit literature. *American Economic Review 112*(6), 1818–47.

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