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

Appendix for Online Publication

Uzma Afzal
University of Nottingham[*]

Giovanna d'Adda
University of Milan[†]

Marcel Fafchamps
Stanford University[‡]

Simon Quinn
University of Oxford[§]

Farah Said
Lahore School of Economics[¶]

January 2021

Contents

[*] **School of Economics**: Uzma.Afzal@nottingham.ac.uk.
[†] **Department of Economics, Management and Quantitative Methods**: giovanna.dadda@unimi.it.
[‡] **Freeman Spogli Institute for International Studies**: fafchamp@stanford.edu.
[§] **Department of Economics and Centre for the Study of African Economies**: simon.quinn@economics.ox.ac.uk.
[¶] **Department of Economics and Centre for Research in Economics and Business**: farahs@lahoreschool.edu.pk.

List of Figures

List of Tables

A Further implementation details

A.1 Randomization

In both phases, we used two separate mechanisms to assign 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. Within each stratum, we then randomly assigned two respondents to the treated group and two respondents to the control group, as described in Table 3. 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 3. 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 3. 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 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). This assignment was implemented by inviting participants to draw a card at random at the beginning of each cycle. 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 cycle. 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 cycle, then week 0 of the second cycle and, finally, week 0 of the third and final cycle in each Phase of the experiment.

A.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 cycle of phase 1 and 2 are given in Table A18.

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 A1 for phase 1 and Table A2 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 A1: **Describing the characteristics of automatic refusers in Phase 1**

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

	(1)	(2)	(3)
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 cycle, 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 A2: **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 cycle, 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 4, we show average take-up frequencies for all six combinations of lumpsum payment timing and net balance, for both phases. In Table A3, 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 A3: **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.

B Robustness checks

B.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.

B.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 A4: **Phase 1: Are you familiar with the concept of a savings committee?**

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

Table A5: **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 A6: **Phase 1: Have you ever participated in any committee?**

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

Table A7: **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).

B.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 A8: **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 A9: **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 A10: **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.

B.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.

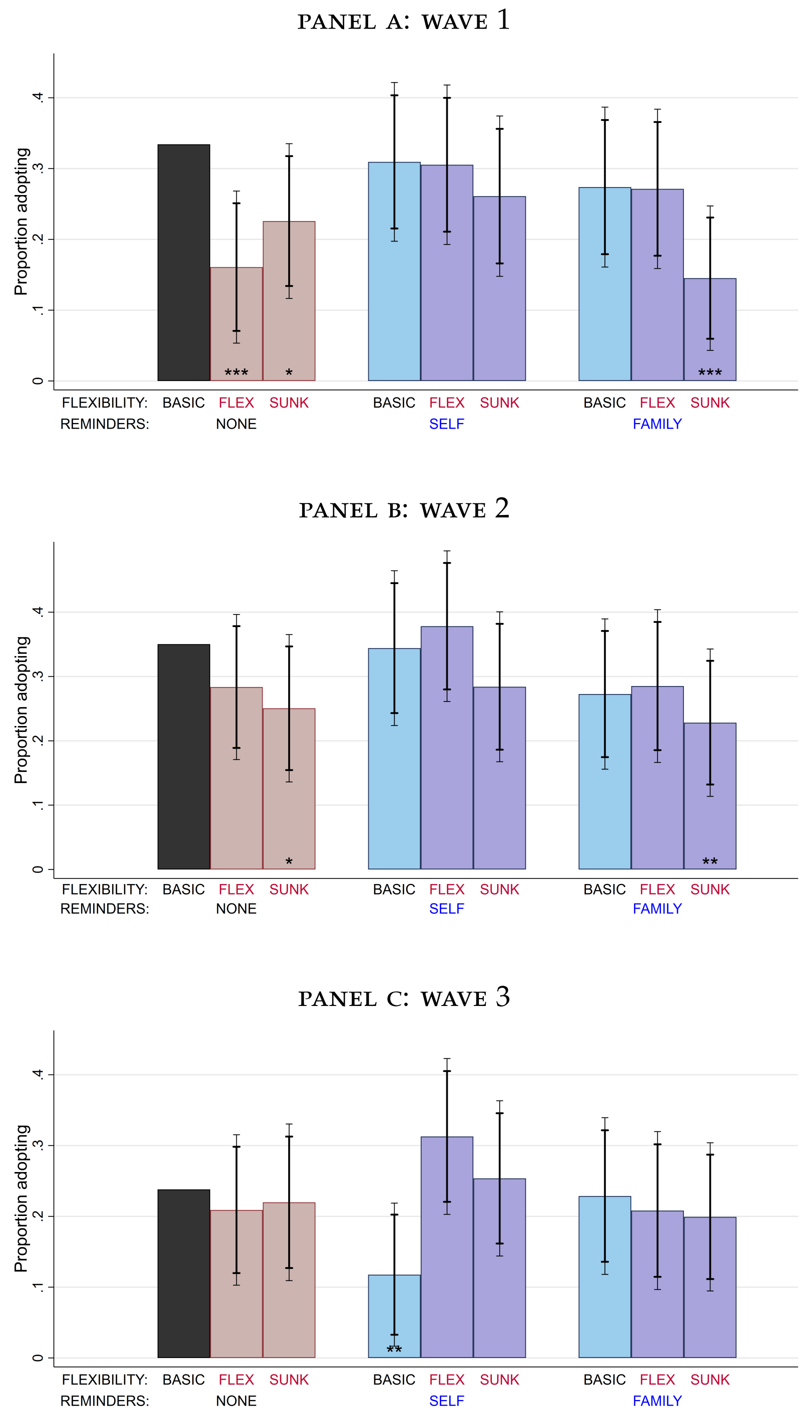
B.2 Additional analysis on dynamics

B.2.1 Disaggregating by wave

In this section, we provide additional analysis on take-up dynamics. In the original Figure 3, 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 3 (at the top) shows these results for credit contracts (that is, contracts where the lumpsum is offered to be paid in the first period); Panel B (at the bottom) shows the results for savings contracts (where the lumpsum is to be paid in the final period).

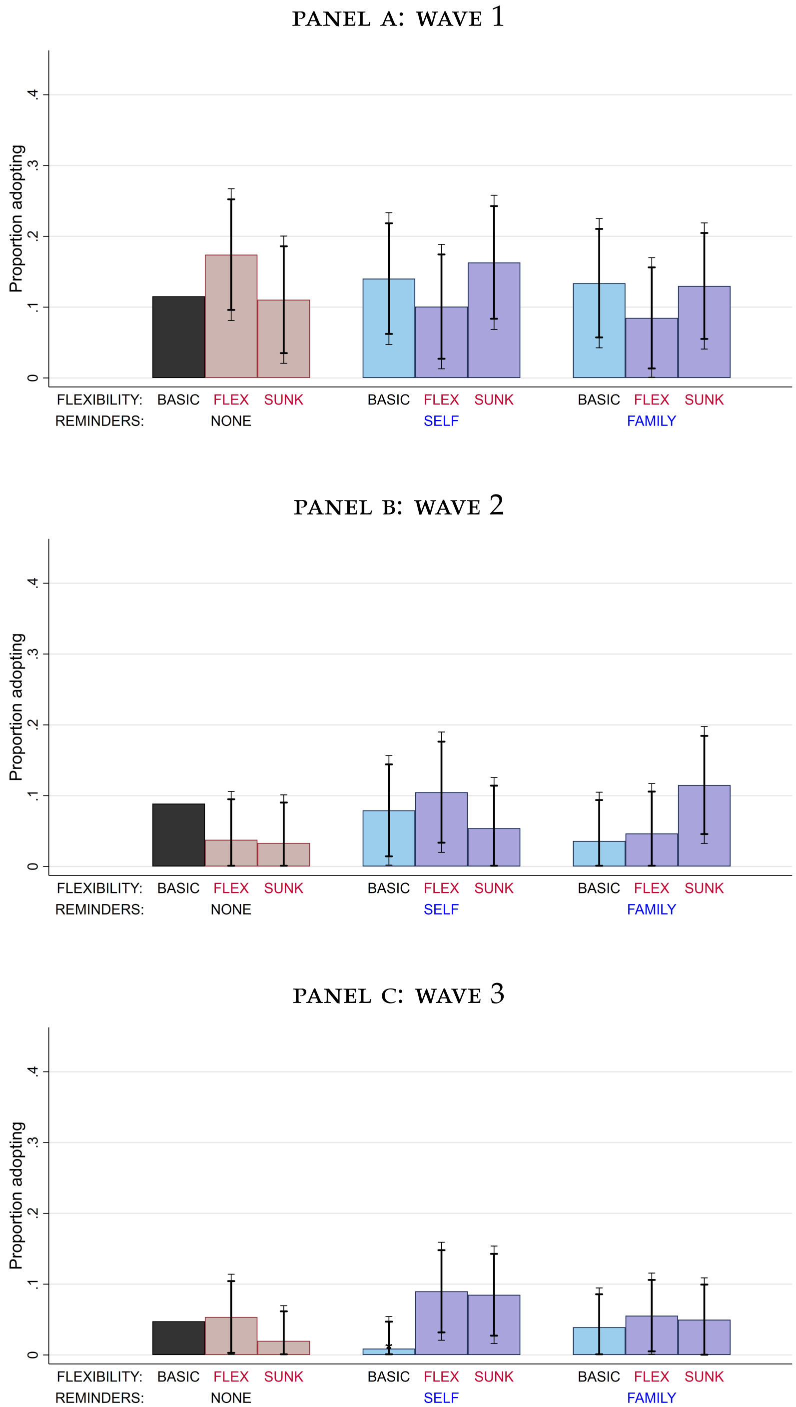
Figure A1 then disaggregates Panel A of Figure 3 by experiment wave; that is, it repeats the credit analysis, splitting the same into wave 1, wave 2 and wave 3. Figure A2 shows the same disaggregating for savings (that is, Panel B of Figure 3). While there is (inevitably) some variation in take-up patterns 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.

Figure A1: **Average credit take-up by contractual add-ons: Disaggregating by wave**



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 A2: **Average savings take-up by contractual add-ons: Disaggregating by wave**



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.

B.2.2 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 A11. 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 A11: **Phase 2: Effect of lagged take-up (instrumented with lagged constractual terms)**

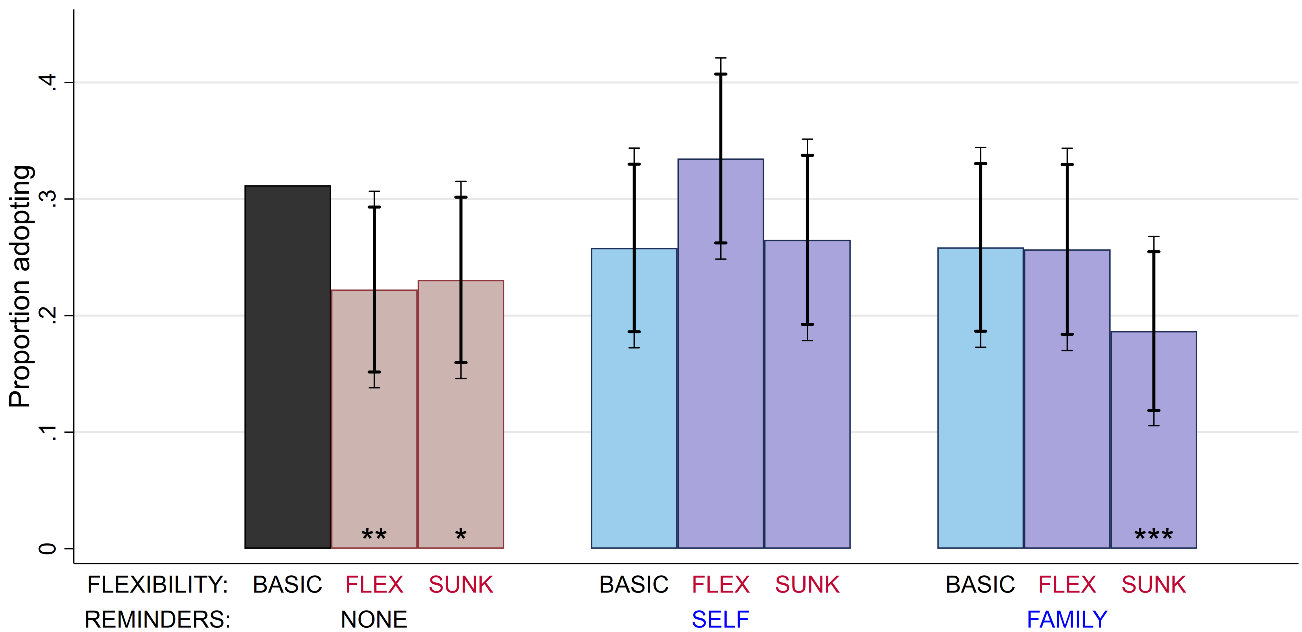
	(1) Waves 2 & 3	(2) Wave 2	(3) Wave 3	(4) Contract terms (all waves)	(5) Contract terms (waves 2 and 3)	(6) Contract terms adding lag
Lagged take-up	0.527	0.547	0.529			0.452
	(0.043)***	(0.074)***	(0.050)***			(0.041)***
Constant	0.061	0.079	0.039			
	(0.009)***	(0.016)***	(0.010)***			
Take-up by contract 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%

B.3 Sensitivity to including defaulters

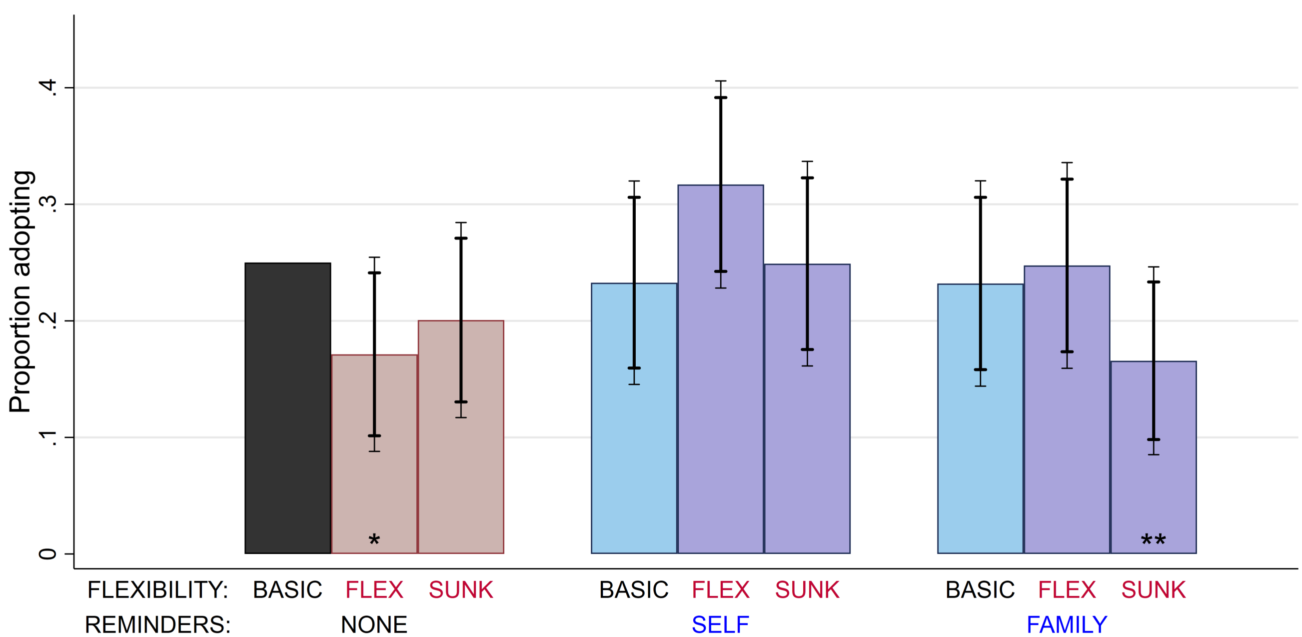
On the following two pages, we show the effect of dropping individuals who ever defaulted. (Thus, for example, if an individual defaulted in wave 3, we drop her observations in waves 1 and 2.) On each page, we show the original figure at the top (as reported in the paper), and the amended figure at the 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 A3: **Average take-up by contractual add-ons: Credit domain**

PANEL A: CREDIT CONTRACTS



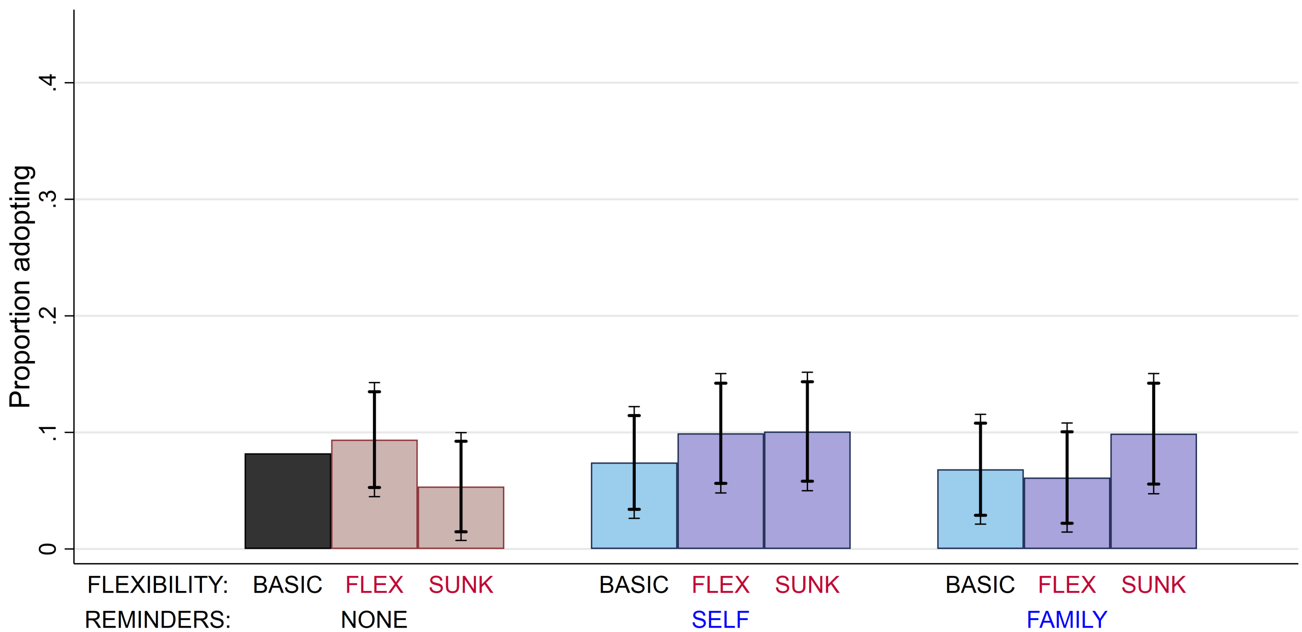
PANEL B: CREDIT CONTRACTS, DROPPING 'EVER DEFAULTERS'



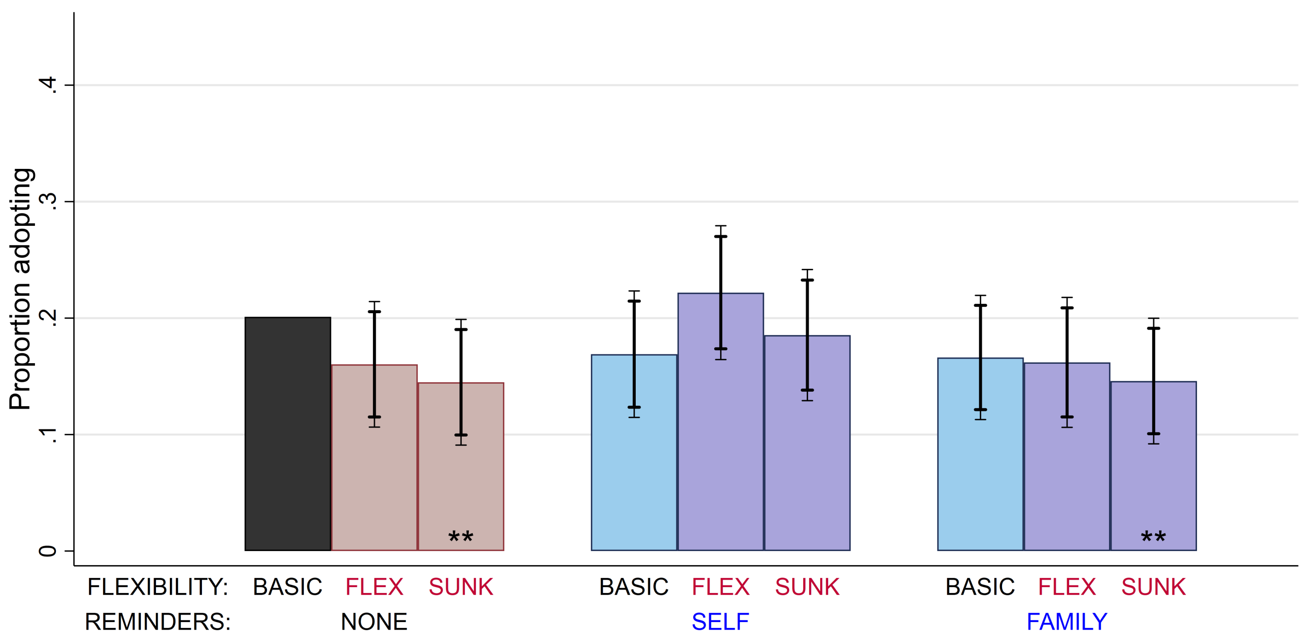
This figure shows the average take-up for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'family' variation), and take-up for each of the eight possible variations. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic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

Figure A4: **Average take-up by contractual add-ons: Saving domain**

PANEL A: SAVING CONTRACTS



PANEL B: SAVING CONTRACTS, DROPPING 'EVER DEFAULTERS'



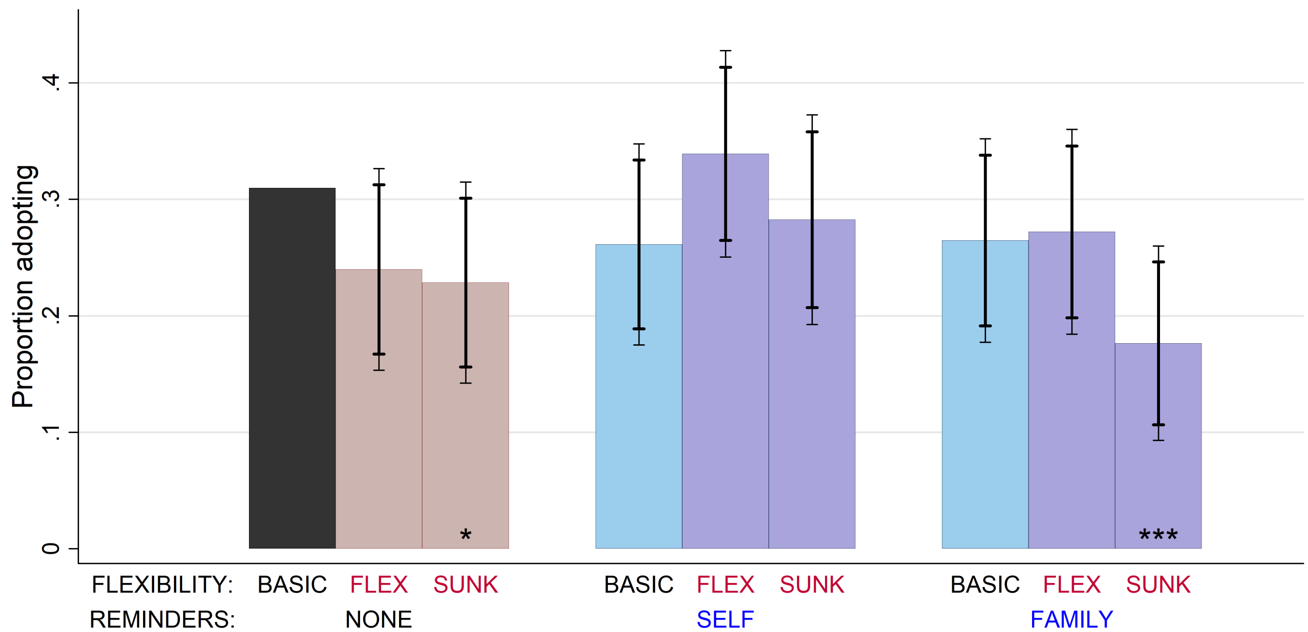
This figure shows the average take-up for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'family' variation), and take-up for each of the eight possible variations. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic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

B.4 Sensitivity to including baseline characteristics as controls

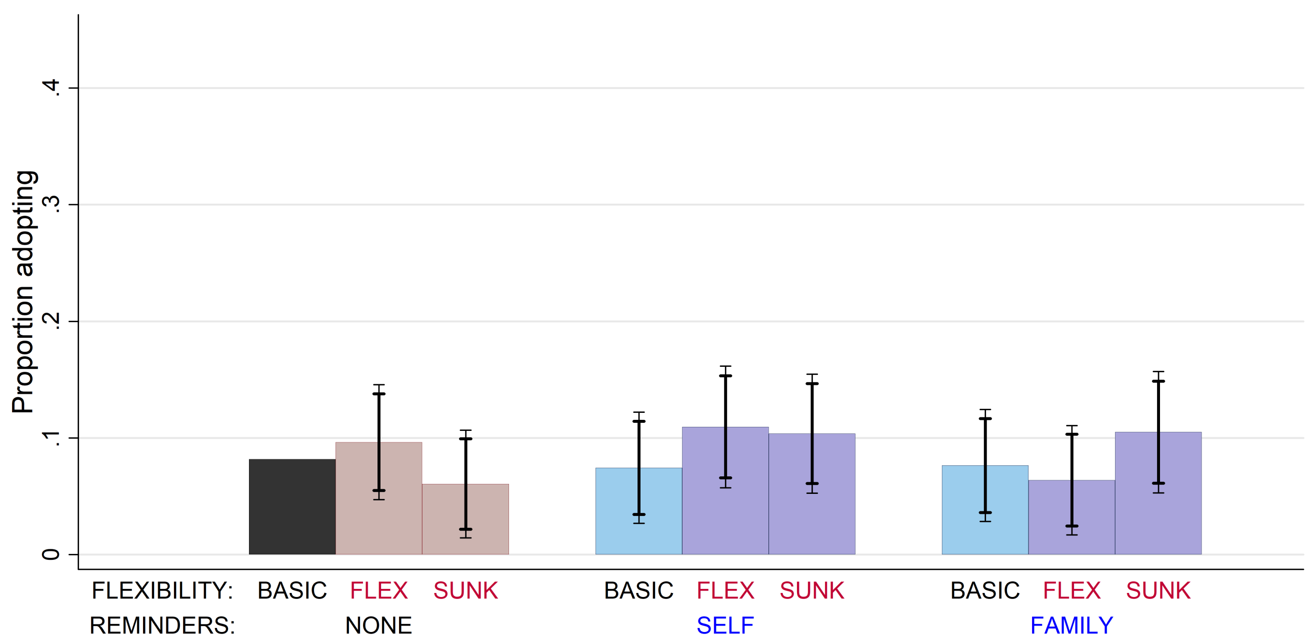
Next, we test the robustness of our estimation results to the inclusion of baseline controls and find that our main estimation results are unaffected. We show (on the far left) of Figure A5, take-up rates for the basic contract (that is, the product with neither the 'flex'/'sunk' variation nor the 'self 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 A6, we similarly show the effects on rate of late payment.

Figure A5: **Average take-up by contractual add-ons: including baseline controls**

PANEL A: CREDIT CONTRACTS



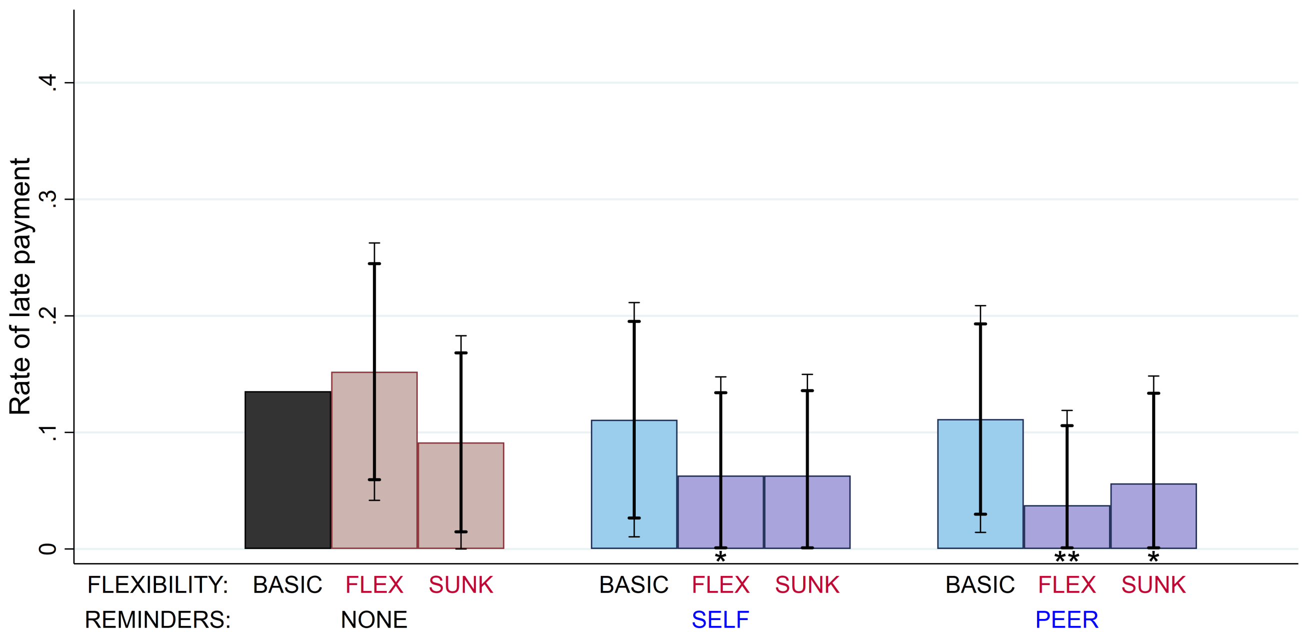
PANEL B: SAVING CONTRACTS



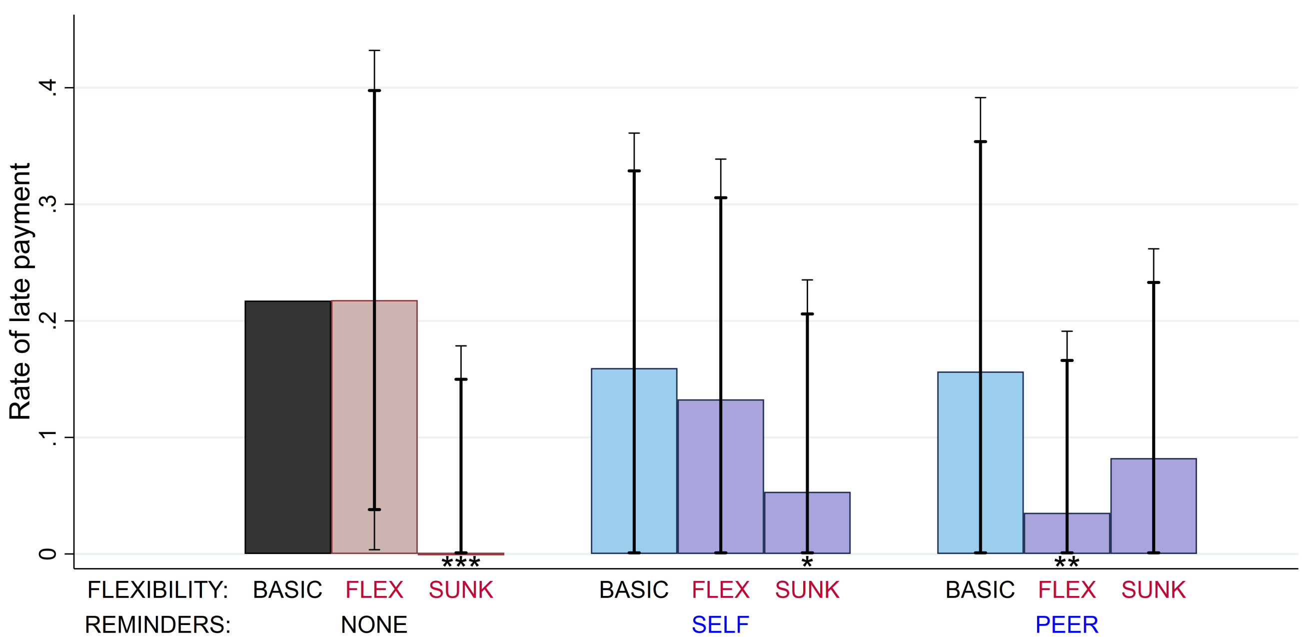
This figure shows the average take-up for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'family' variation), and take-up for each of the eight possible add-ons. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic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

Figure A6: **Rate of late payment by contractual add-ons: including baseline controls**

PANEL A: CREDIT CONTRACTS



PANEL B: SAVING CONTRACTS



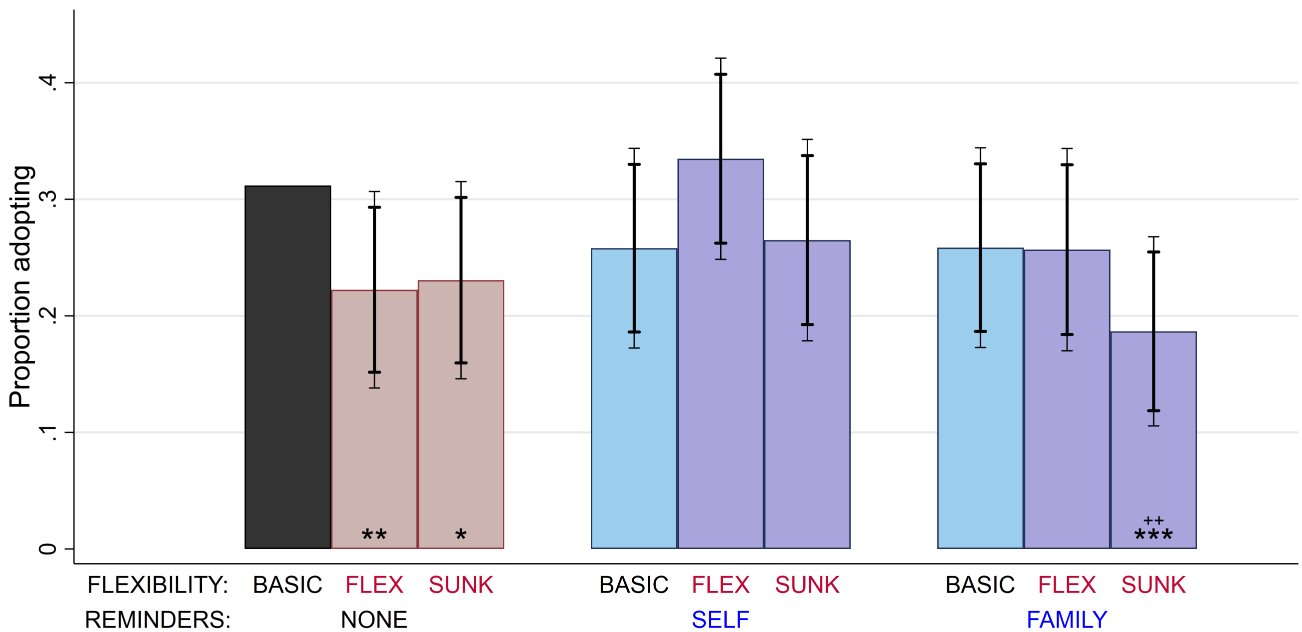
This figure shows the rate of late payment for the basic product (that is, the product with neither the 'flex'/'sunk' variation nor the 'self'/'family' variation), and for each of the eight possible variations, where the minimum rate of late payment is set to 0. Error bars show 90% and 95% confidence intervals on the difference in rate of late payment to the basic contract. Stars indicate a significant difference from the basic contract.

B.5 Correcting for multiple hypothesis testing

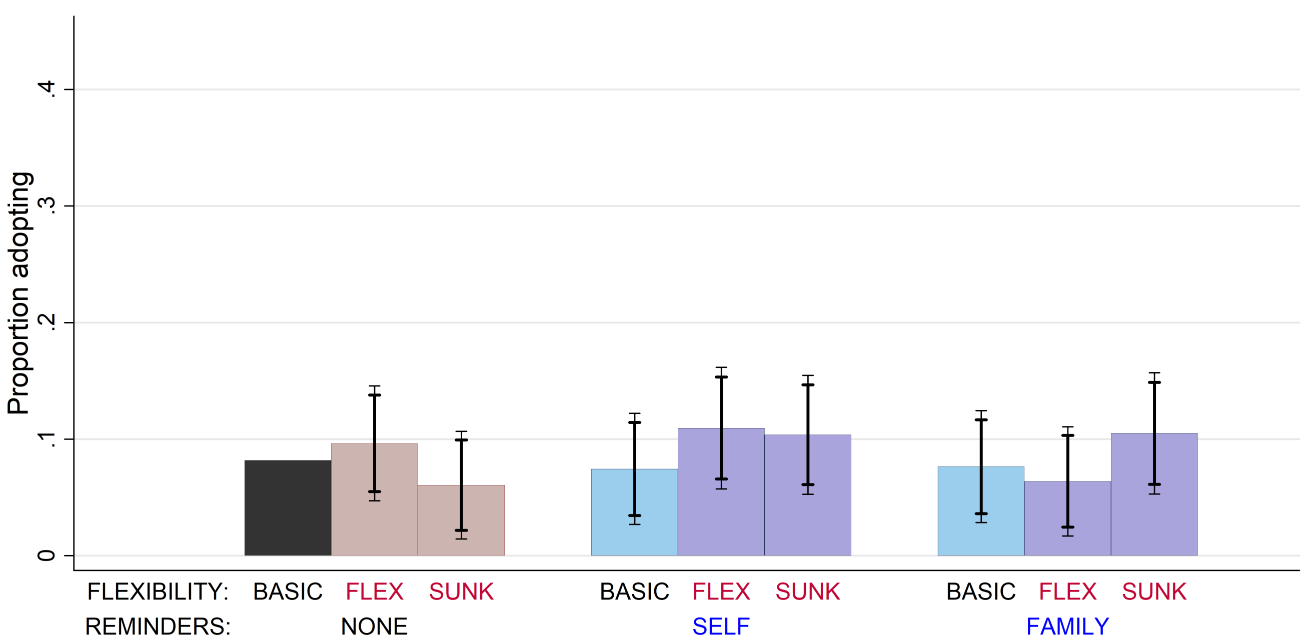
Next, we correct for multiple hypothesis testing. For each bar representing take-up rates in Figure 3 and rates of late payment in Figure 4 in the main text, we add sharpened False Discovery Rates (FDR) q-values and indicate when they are significant using '+' when the difference from base contract is significant at the 10% level and '++' when it is different at the 5% level.

Figure A7: **Average take-up by contractual add-ons: correcting for multiple hypothesis testing**

PANEL A: CREDIT CONTRACTS



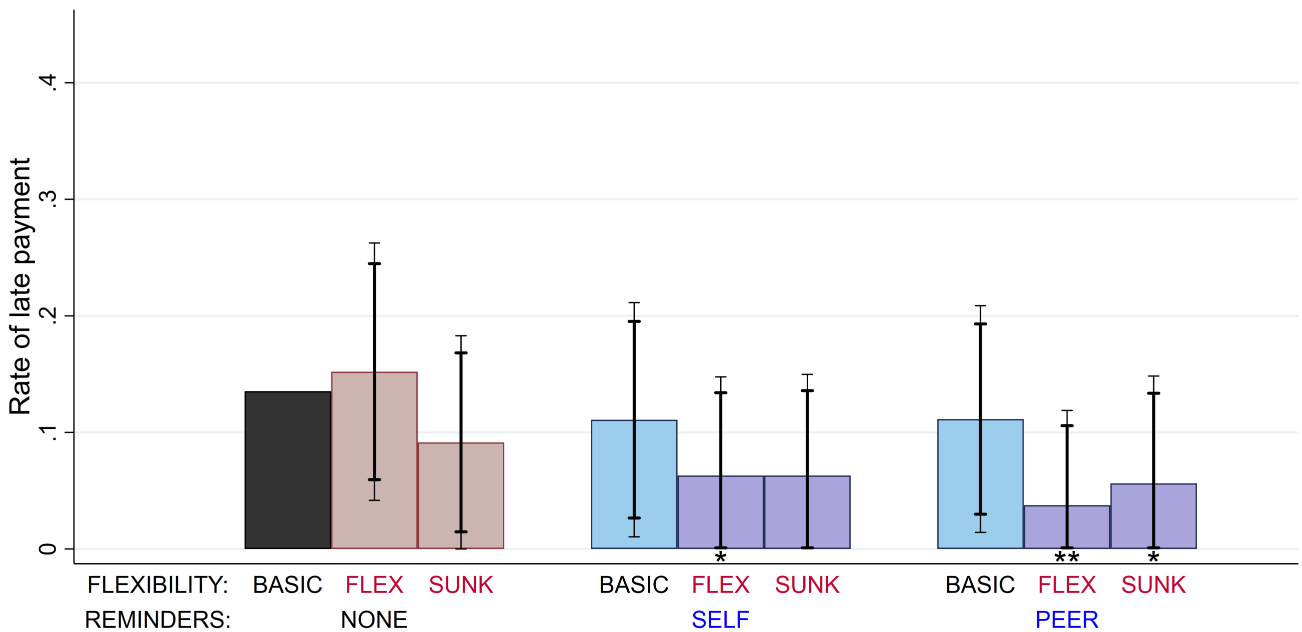
PANEL B: SAVING CONTRACTS



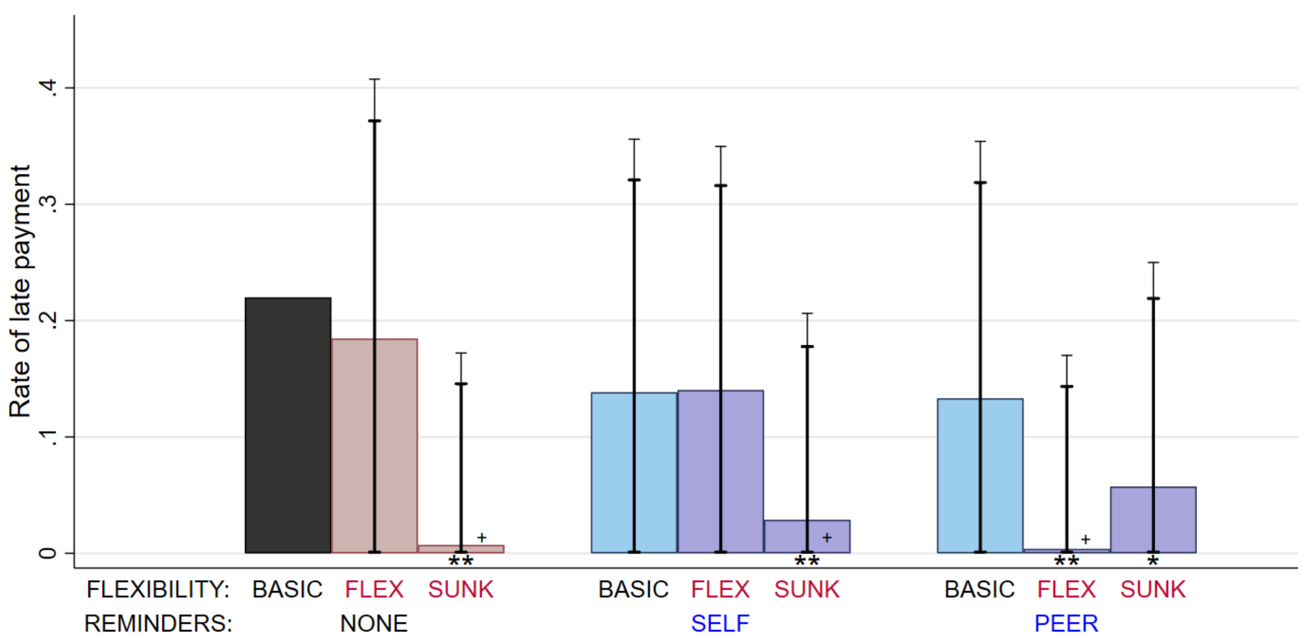
This figure shows the average take-up for the basic product, and take-up for each of the eight possible add-ons. Error bars show 90% and 95% confidence intervals on the difference in take-up to the basic contract. Stars indicate a significant difference from take-up of the basic contract. '+'s indicate significance levels corrected for Multiple Hypothesis Testing.

Figure A8: **Rate of late payment by contractual add-ons: correcting for multiple hypothesis testing**

PANEL A: CREDIT CONTRACTS



PANEL B: SAVING CONTRACTS



This figure shows the rate of late payment for the basic product and for each of the eight possible variations. 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. '+'s indicate significance levels corrected for Multiple Hypothesis Testing.

C 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.[1] 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 A9 shows Group Average Treatment Effects ('GATES'), sorted by take-up propen-

[1] 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).

sity. 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 A9, are a direct analog to Figure 4 in Chernozhukov et al. (2018). They show that we have substantial heterogeneity in take-up rates across individuals with different covariates: for the lowest quintile, the average take-up rate is approximately 10%, and for the highest quintile, the rate is above 25%.[2]

We augment this analysis in two ways. First, in the top panel of Figure A9, 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.[3] 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 A9, we repeat the analysis for the contract terms: that is, for variations in the lumpsum amount and in the time of payment. Again, the basic pattern – and all of the stylised facts noted in Table 4 – holds across all quintiles.

Who, then, are the respondents who fall into these quintiles? Following Chernozhukov et al. (2018), we answer this question by describing the characteristics of those respondents in the 'most affected' and 'least affected' groups – that is, the 20% with the highest adoption rate (which we term the 'highest adopters') and the 20% with the lowest adoption rate (the 'lowest adopters'). In Table A22, we perform this comparison for all 58 of the baseline covariates used for our analysis. In Table A12, 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

[2] 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).

[3] 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.

keep track of tasks and finances.

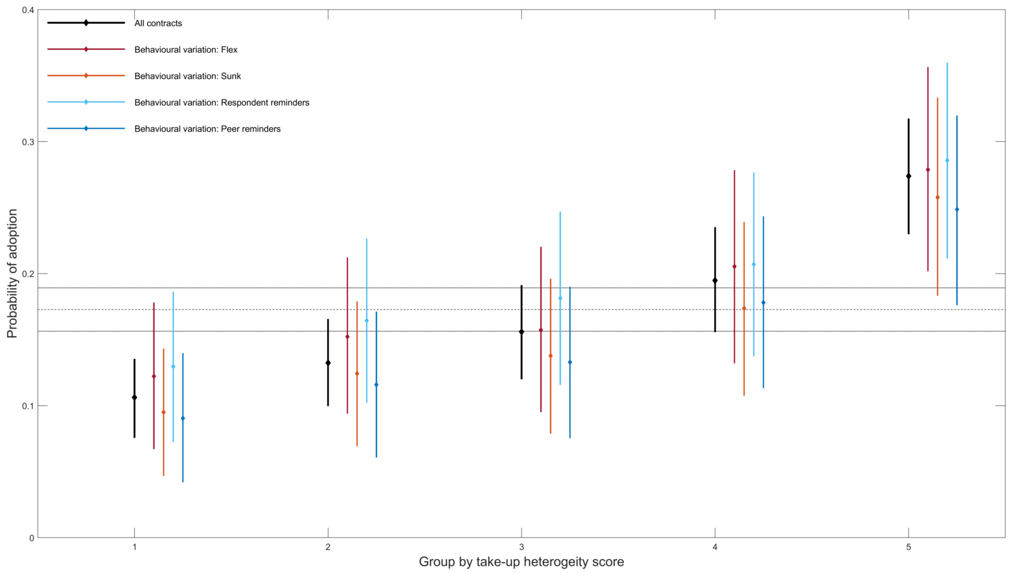
Table A12 shows large and highly significant differences in respondent characteristics for almost all of the 'behavioral' characteristics in Phase 2. It is particularly noteworthy that, of the highest adopters, 89% said at baseline that they find it hard to save, and 94% said they face pressure to share; the equivalent figures for the lowest adopters are just 54% and 55% respectively. Further – and consistent with our interpretation that the basic contract provides a useful commitment device – the highest adopters are significantly less likely to have described themselves at baseline as 'good at keeping track of time', 'good at keeping track of finances', to follow a strict schedule on finances, to follow a tight routine, and less likely to act early to avoid forgetting (either generally or with respect to finances). Finally, as one might expect, the highest adopters report significantly higher intra-household empowerment at baseline: they report a significantly higher share of household decisions in which the woman's view is always considered, and are more than twice as likely to 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 A22 appears in the appendix as Table A23. 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.

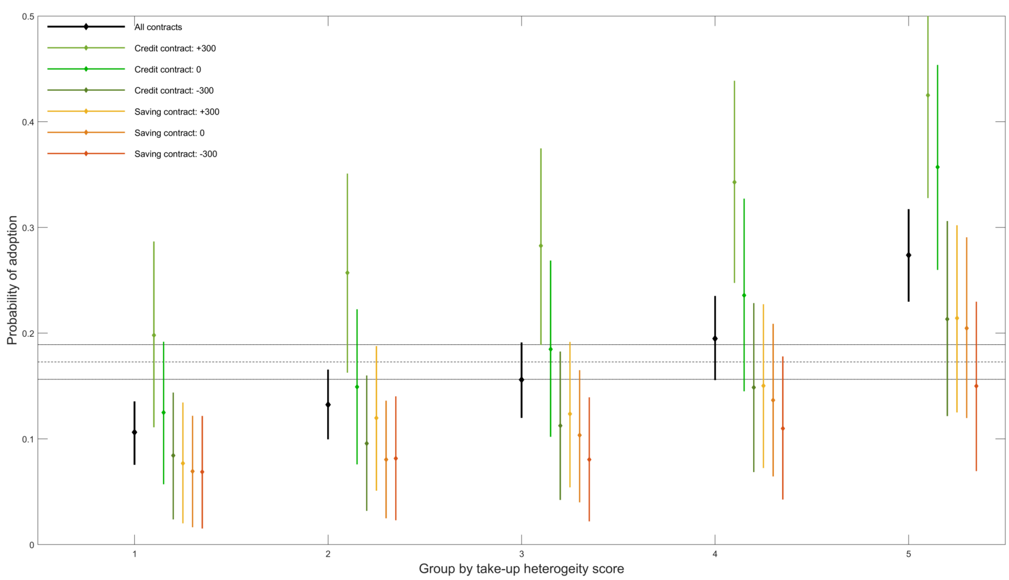
Finally, we show the Phase 1 counterpart to Figure A9 in the appendix, as Figure A14. The general patterns are the same, though Phase 1 respondents appear to have a greater sensitivity to the size of lumpsum payment.

Figure A9: **Group Average Treatment Effects (sorted by take-up propensity)**

PANEL A: TAKE-UP BY BEHAVIORAL VARIATION



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



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

Table A12: **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).

D Business and household outcomes

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

$$y_{i1} = \beta_0 + \beta_1 \cdot T_i + \beta_2 \cdot y_{i0} + \phi_s + \eta_d + \varepsilon_i, \tag{A1}$$

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 A1 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 A1 using OLS, and interpret $\hat{\beta}_1$ as the ITT. Second, we denote T_i as take-up. This takes four possible values, which depend on whether the subjects takes up the contract in 0, 1, 2 or 3 cycles. In this case, we calculate average take-up at the individual level, and instrument this using assignment to treatment and to contractual terms; we then interpret $\hat{\beta}_1$ as providing the LATE, normalized for a case where a respondent takes up in all three product cycles.[4]

Outcomes are divided into two broad categories: business outcomes and household finance and consumption. In Table A13, 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.

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

Table A13: **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 A1. 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 (2018a) and Meager (2018b)). 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 A14. 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 A14 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 A14 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 A24 (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 A1 separately for each of those quintiles, for all of the outcomes in Table A14 and Table A14. 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 A10 and A11, we show this for two outcomes of particular interest: business investment and household consumption (for Phase 1 and for Phase 2 respectively).

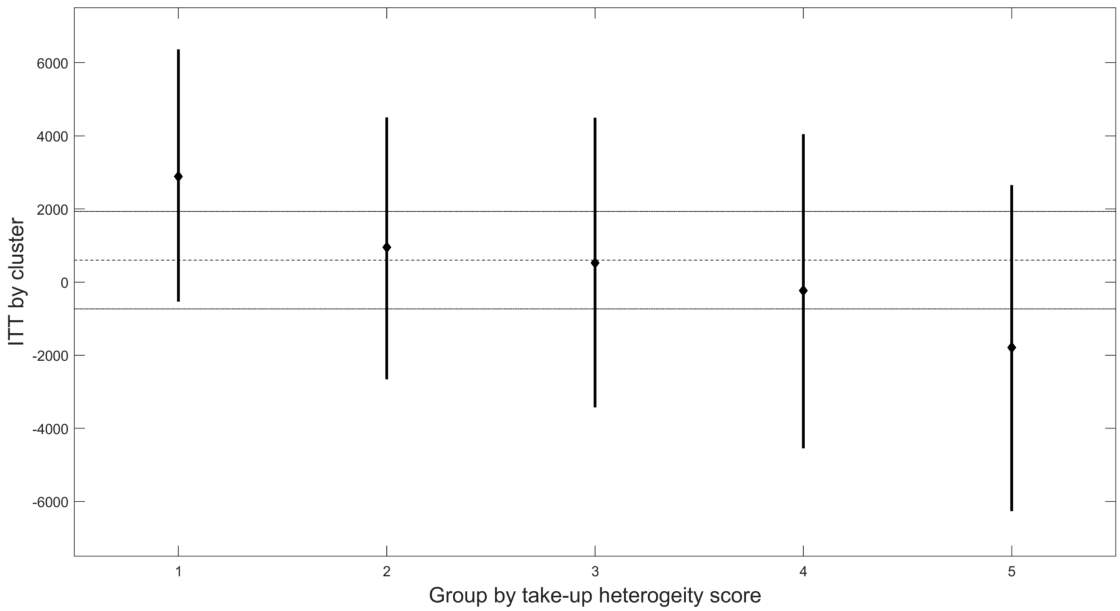
Table A14: **Summary of ITT and LATE estimates of household material outcomes**

	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

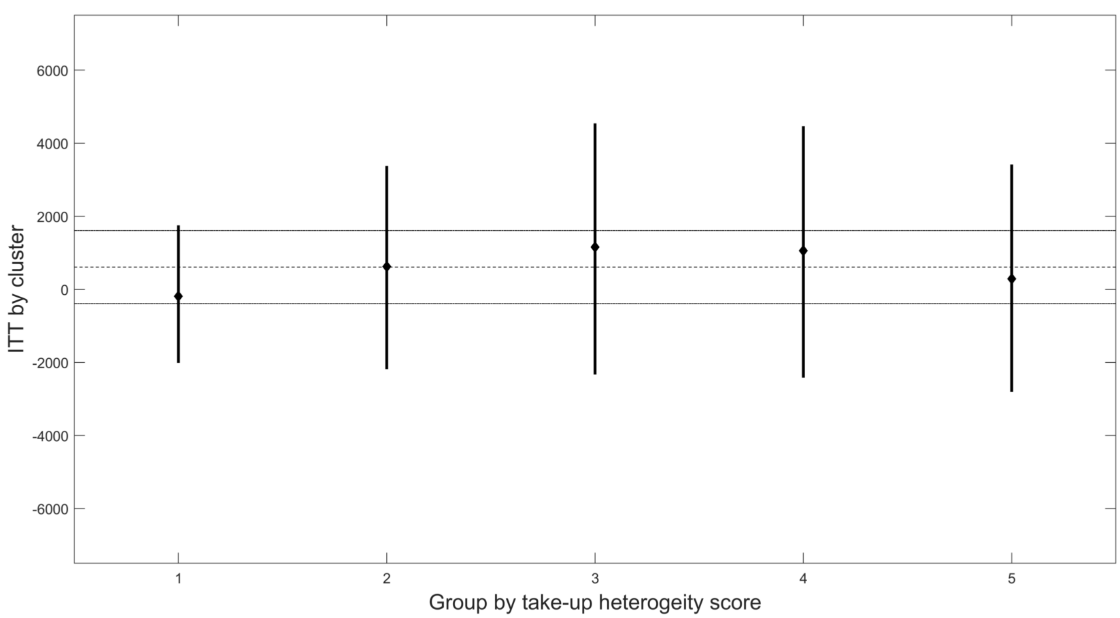
This table reports regression estimates of equation A1. We report standard errors under each coefficient in parentheses. All values are in Pakistani rupees. Confidence: $* \leftrightarrow p < 0.1$; $** \leftrightarrow p < 0.05$; $*** \leftrightarrow p < 0.01$.

Figure A10: **Group Average Treatment Effects (sorted by take-up propensity): Consumption and investment (Phase 1)**

HOUSEHOLD MONTHLY CONSUMPTION



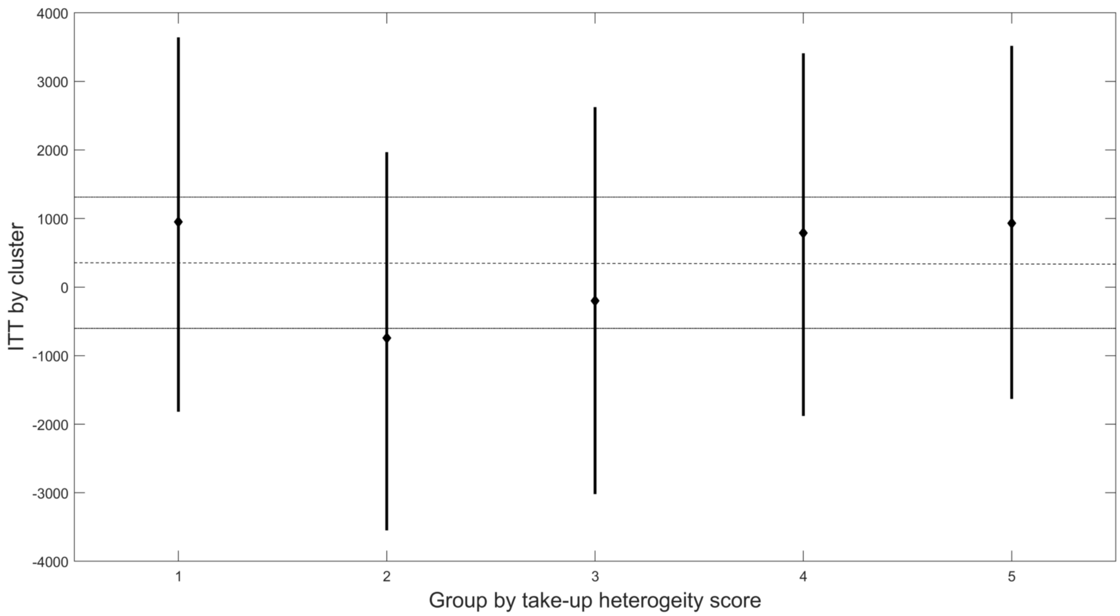
BUSINESS INVESTMENT



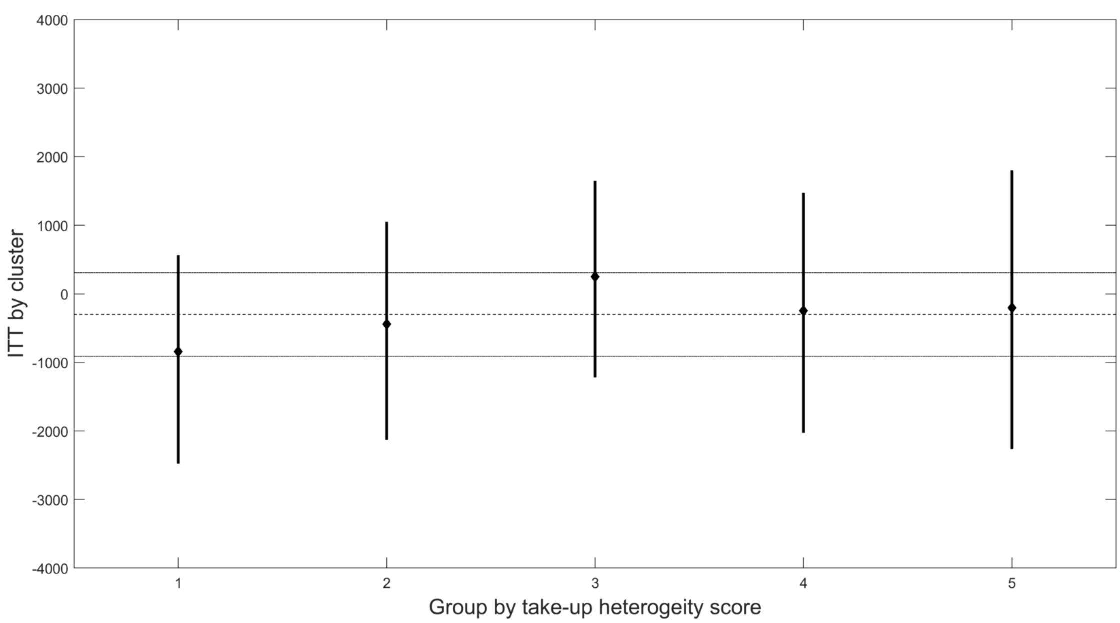
This figure shows the Group Average Treatment Effects, sorted by the take-up propensity estimated in the main text. For each of the five groups separately, we estimate equation A1; the graph shows the estimated ITT and 90% confidence intervals (both formed using the bootstrap methodology proposed by 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 A13 and Table A14).

Figure A11: **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 A1; 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 A13 and Table A14).

E Additional tables and figures

Table A15: **Description of the sample: Phase 1**

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

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

Table A16: **Description of the sample: Phase 2**

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

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

Table A17: **Description of the sample by treatment groups: Phase 2**

	Control	Basic-None	Basic-Peer	Basic-Resp	Flex-None	Flex-Peer	Flex-Resp	Sunk-None	Sunk-Peer	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

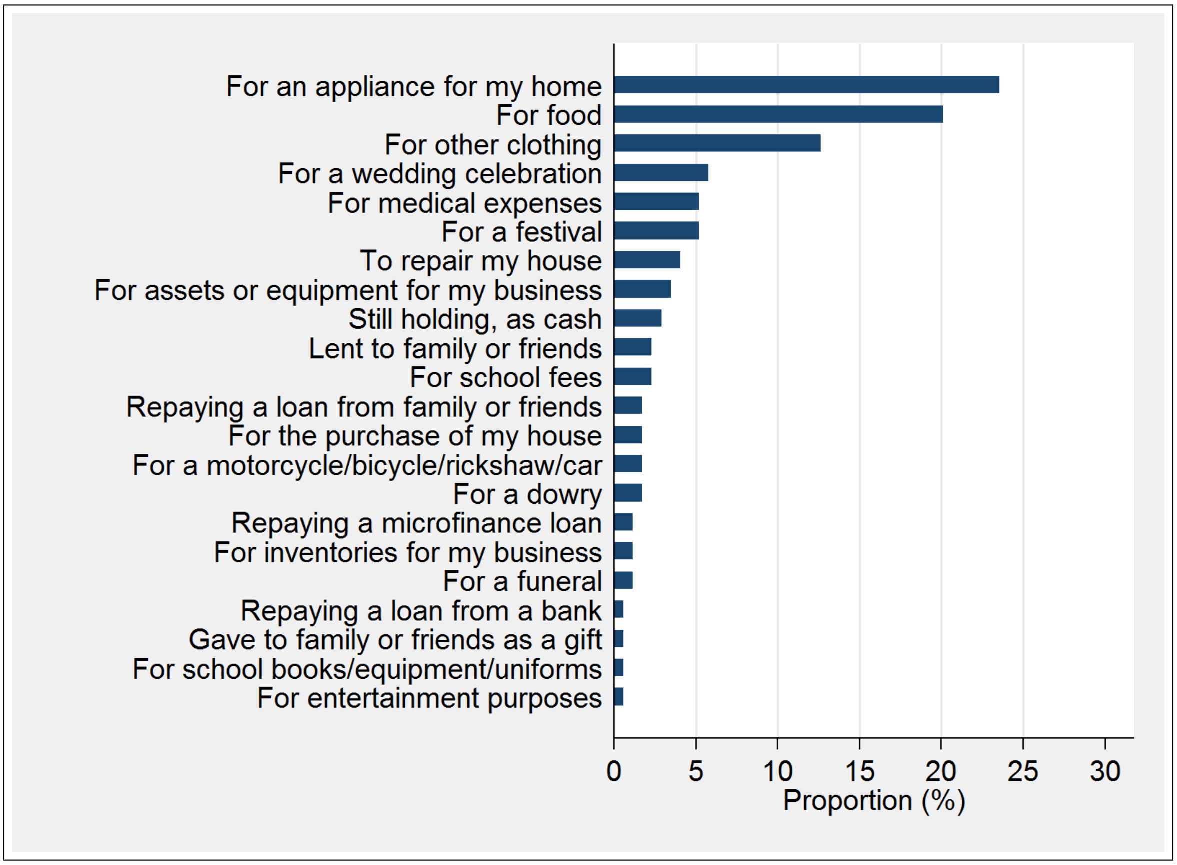
This table provides basic summary statistics for sample characteristics across the 3X3 factorial design with controls in Phase 2. For each variable, we report the mean value for control and each of the treatment group.

Table A18: **Proportion of treated who refused the product offered before the contract offer cards were drawn**

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

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

Figure A12: **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 A19: **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 A20: **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 A21: **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

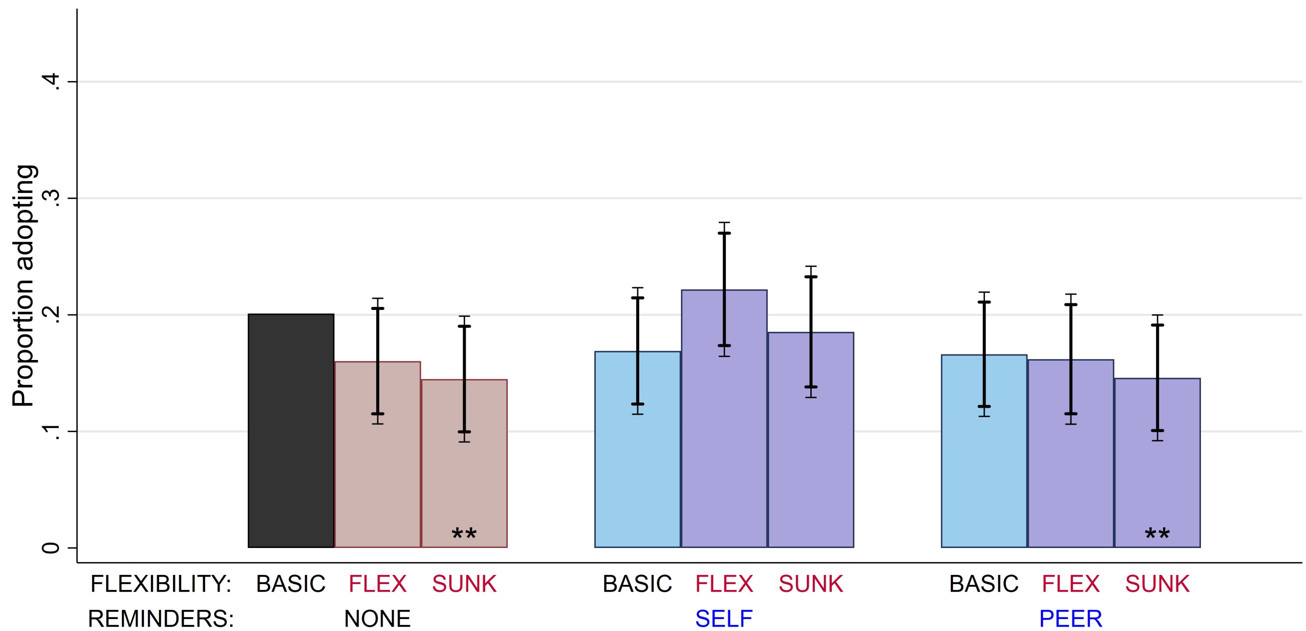
	Flex	Basic	Sunk
No reminders	-4.1%	*reference*	-5.6%**
Reminder to self	5.3%**	*reference*	1.6%
Reminder to family	0.4%	*reference*	-2.0%

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 A13: **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.

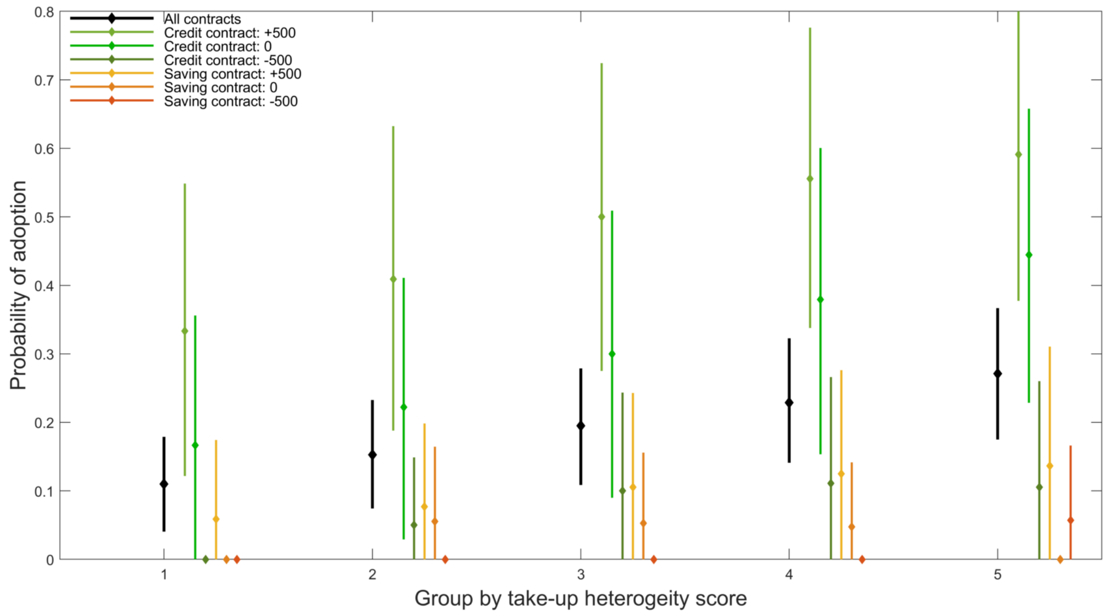
Table A22: **Cluster Analysis: Description of extreme groups (all covariates: Phase 2)**

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

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

Figure A14: **Group Average Treatment Effects (sorted by take-up propensity)**

TAKE-UP BY CONTRACTUAL TERMS (PAYMENT AND TIMING)



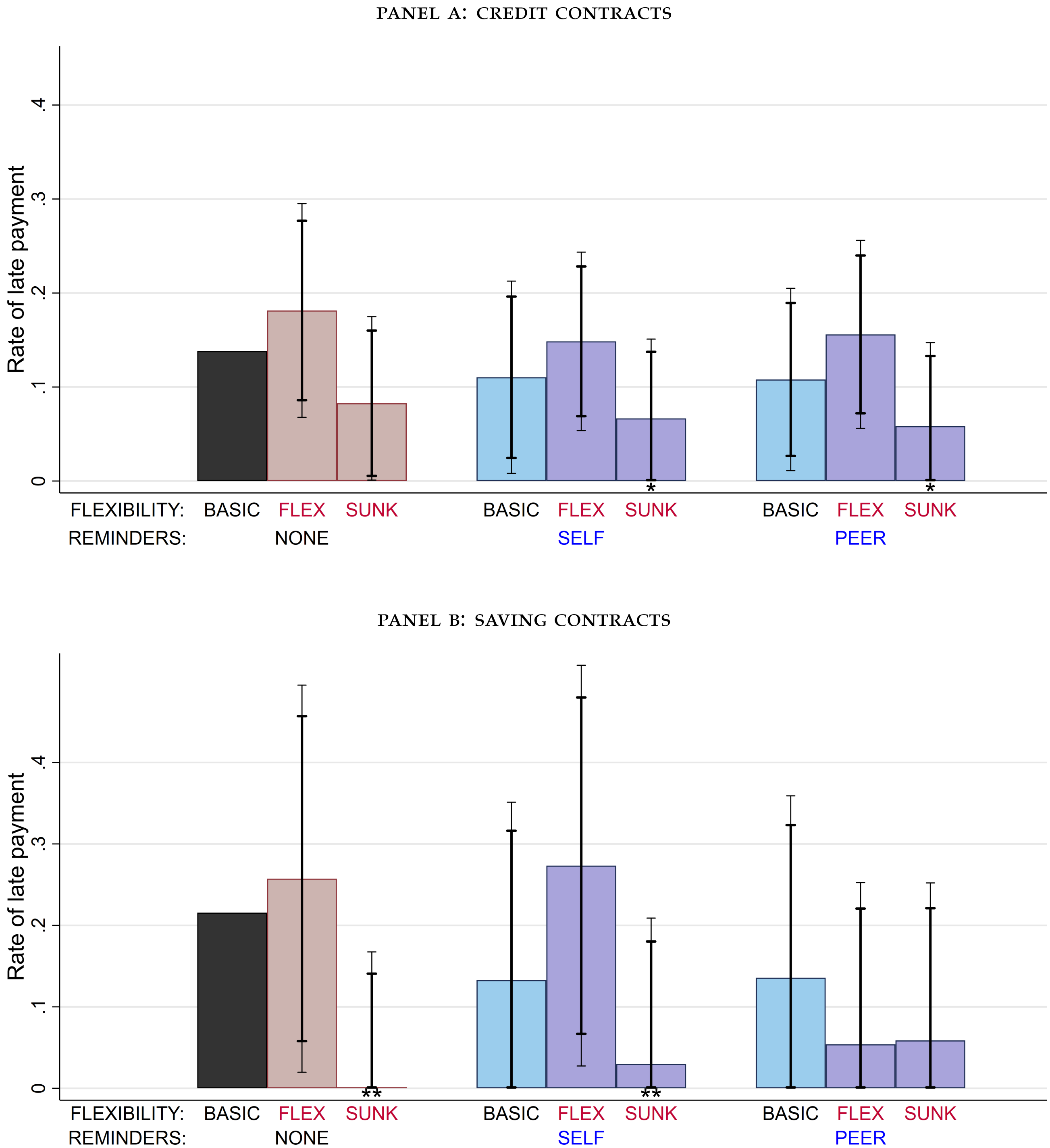
This figure shows the Group Average Treatment Effects, sorted by the take-up propensity estimated in the main text. In each figure, the leftmost (black) lines for each group show the average probability of take-up across all contract types; note that these leftmost lines are identical across figures (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 A23: **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 A15: **Late payment by contractual add-ons (including respondent using the 'flex' option)**



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 A24: **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 A1. We report standard errors under each coefficient in parentheses. All values are in Pakistani rupees. Confidence: $* \leftrightarrow p < 0.1$; $** \leftrightarrow p < 0.05$; $*** \leftrightarrow p < 0.01$.

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.

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*.

Meager, R. (2018a). Aggreating Distributional Treatment Effects: A Bayesian Hierarchical Analysis of the Microcredit Literature. *Working paper*.

Meager, R. (2018b). Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments. *American Economic Journal: Applied Economics*.