CESifo WORKING PAPERS

8322
2020

Original Version: May 2020
This Version: July 2023

Flexible Contract, Flexible Morale? Microcredit Design and Repayment Discipline

Kristina Czura, Anett John, Lisa Spantig

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

Flexible Contract, Flexible Morale? Microcredit Design and Repayment Discipline

Abstract

Recent evidence suggests that more flexible microloan repayment benefits borrowers, but lenders fear diminished repayment morale. We study repayment choices in rigid and flexible loan contracts with discretion in repayment timing. Using a lab-in-the-field experiment with 645 microcredit borrowers in the Philippines, we identify moral hazard and quantify social pressure. Payoff maximization predicts low repayment in our rigid benchmark contract, and increased repayment with flexibility. Results suggest the opposite: Repayment in the rigid contract is high, and drops substantially under flexible repayment. Social pressure decreases. Our results are consistent with a strong social norm for repayment, which is weakened by introducing flexibility. Norms, which may be inculcated by the lender, may help explain several recent puzzles in microfinance research, including high and equal repayment rates across individual and joint-liability contracts, and excessive peer pressure. Importantly, norm-driven behaviour may erode with the introduction of flexibility.

JEL-Codes: O160, D900, G210.

Keywords: microfinance, flexible repayment, ex-post moral hazard, social norms, peer punishment.

Kristina Czura
University of Groningen / The Netherlands
k.czura@rug.nl

Anett John
University of Birmingham / United Kingdom
A.John@bham.ac.uk

Lisa Spantig
RWTH Aachen University / Germany
Lisa.Spantig@rwth-aachen.de

July 2023
The authors would like to thank Giorgia Barboni, Toman Barsbai, Björn Bartling, Lasse Brune, Xavier Giné, Selim Gulesci, Christa Hainz, Sanjay Jain, Melanie Koch, Martin Kocher, Friederike Lenel, Robert Lensink, Andreas Madestam, Pushkar Maitra, Muhammad Meki, Simon Quinn, Pedro Rey Biel, Frédéric Schneider, Simeon Schudy, Klaus Schmidt, Roberto Weber, Ro'i Zultan, and various seminar audiences for helpful feedback and discussions. We are indebted to Antonia Delius for excellent research assistance. The authors are grateful for financial support from oikocredit International. Kristina Czura acknowledges funding from the LMUexcellence Junior Research Fund. Anett John acknowledges funding from Labex Ecodec (Investissements d'Avenir, ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047). Lisa Spantig acknowledges funding through the International Doctoral Program 'Evidence-Based Economics' of the Elite Network of Bavaria. The experiment was registered at the AEA RCT registry, ID 0002355. It has been approved by the Ethics Commission of the Department of Economics, University of Munich (Project 2016-04, approved 03.02.2016). All errors and omissions are our own. An earlier version of the paper has circulated under the title "Flexible Microcredit: Effects on Loan Repayment and Social Pressure".

"We pledge to attend regularly the weekly Center meetings,
to utilize our loans for the purpose approved, to save and pay our installments weekly,
to use our increased incomes for the benefit of our families, to ensure that other members
of our group and Center do likewise and to take collective responsibility if they do not."

Official weekly pledge, recited at each center meeting
Grameen Foundation

1 Introduction

Microcredit has long been praised for its innovative product design, which extended access to finance to a new segment of the market: low-income borrowers without sufficient collateral. Joint-liability lending uses social capital to overcome screening and monitoring problems, and to provide mutual insurance (Armendáriz 1999; Besley and Coate 1995; Iyer et al. 2016). Rigid repayment schedules with high-frequency, same-sized installments are thought to instill repayment routine and morale (Armendáriz and Morduch 2010; Bauer et al. 2012; Labie et al. 2017; Meyer 2002). Indeed, microcredit providers emphasize that repayment rigidity plays an important role in preventing strategic default, i.e., the choice to default despite having the capacity to repay. However, both of these building blocks have been criticized as not fit for purpose: Joint liability may induce excessive peer pressure (Czura 2015b; Karim 2008; Montgomery 1996). Rigid repayments make it challenging to match cash flows and lead to underinvestment (Karlan and Mullainathan 2007).

While microlenders have responded by offering more individual-liability contracts, they remain reluctant to offer more flexible repayment (Labie et al. 2017). This is surprising for two reasons. First, providing borrowers with flexibility in when to repay, and thus enabling them to condition repayments on shock realizations, may increase investments in high-return projects. A series of recent field experiments confirms that repayment flexibility increases income by fostering investment in more profitable projects (Barboni and Agarwal 2021; Battaglia et al. 2021; Czura 2015a; Field et al. 2013).[1] Given increasing commercialization and competition in the sector (De Quidt et al. 2018), lenders have incentives to engage in product innovations that benefit their customers, including flexible repayment schemes. Second, the concern that flexible repayment timing may increase strategic default contradicts standard economic predictions: Flexibility *increases* the monetary

[1]'Flexible repayment' has been used to refer to various repayment structures that diverge from rigid weekly repayments starting immediately after loan disbursal. Field et al. (2013) study a two-month grace period at the beginning of the loan, while Barboni and Agarwal (2021) study a three-month repayment holiday that requires a one-month notice period. Throughout this paper, we use 'repayment flexibility' for contracts which allow for *discretion* in when to repay, and thus enable borrowers to condition repayments on shock realizations (see e.g. Bari et al. (2021), Battaglia et al. (2021), Brune et al. (2022), and Czura (2015a)).

incentives to repay, as the possibility to postpone repayments reduces the risk of a shock-induced default, and thus increases the expected benefits of repayment. It also lowers the effective interest rate. Strategic default should thus *decrease* with flexibility. Field experiments show both increases (Brune et al. 2022; Czura 2015a; Field et al. 2013) and decreases (Barboni and Agarwal 2021; Battaglia et al. 2021) in overall default rates under flexible repayment. However, field studies cannot address practitioners' concern about the decision to repay the loan or strategically default (ex-post moral hazard): This requires a separate observation of repayment capacity and repayment decisions, which is typically not feasible in the field. This paper fills this gap, and proposes an explanation why flexible repayment contracts are largely absent in practice despite the observed benefits for borrowers.

We present the first causal evidence on ex-post moral hazard under flexible repayment conditions, and compare it to that under rigid repayment. Using a lab-in-the-field experiment with 645 microcredit borrowers in the Philippines, we study how flexibility (the option to defer repayments and make up for them later) affects individual loan repayment choices. To provide insights into mechanisms, we randomize individual and group loans, which differ in their liability structure and the possibility to show disapproval through peer punishment. We present four main results. First, repayment in rigid contracts is high and constant across individual and joint-liability contracts: 66 percent of participants choose to fully repay, despite the fact that repayment is monetarily unprofitable. Second, peer punishment is excessive and sequentially irrational: 51 percent punish non-repayment caused by observable shocks, i.e., when punishment cannot discipline group members. Third, individuals choose to repay less often when they have discretion in when to repay: Flexibility *increases* strategic default by 50 percent (16 percentage points). This result holds irrespective of the liability structure, and despite a setting that controls for present bias. Fourth, flexibility reduces peer punishment by around half. Our results are consistent with a strong social norm for repayment, whose effect is weakened by introducing flexibility. We illustrate our hypothesis using a theory framework of microcredit repayment in the presence of social norms. In addition, we provide evidence from an incentivized norm elicitation experiment.

Our lab-in-the-field setting has a number of unique advantages: First, we can disentangle repayment capacity from the choice to repay, and thus cleanly identify ex-post moral hazard. Second, we can measure social pressure in an incentive-compatible way through costly punishment choices. The ability to observe punishment when shocks are fully visible to peers allows us to speak to recent concerns about excessive social pressure in microcredit. Third, we minimize the distance to borrowers' natural environment: Experimental sessions are run with borrowing peers in existing microcredit centers in their weekly meeting locations. Repayment decisions are framed using terminology from real lending contracts. This field context allows us to build upon the experience and the existing social capital which prevail in the centers.

In light of our results, we discuss an understudied driving force in microcredit repayment: social norms. Through the client induction process as well as meeting and reciting pledges every week (see quote at the beginning of this paper), clients internalize what it means to be a 'good borrower': to pay installments every week without fail, and to take collective responsibility for the behavior of their peers (Grameen Foundation 2010). This image of a good borrower, as shaped by the lender, closely relates to the concept of identity in organizations (see the seminal work by G. A. Akerlof and Kranton 2000, 2002, 2005). Microfinance organizations have strong incentives to inculcate a sense of identity in their borrowers, with associated norms for good behavior. If they deviate from the prescribed behavior, borrowers suffer penalties which can be psychological (e.g., cognitive dissonance, guilt, shame) or social (reputation). Identity models can hence serve as a microfoundation for social norms – shared perceptions of what behavior is appropriate for whom and in which situation (G. A. Akerlof and Kranton 2005; P. Fischer and Huddart 2008).[2] Specifically, norms for 'good borrowers' may compel individuals to pay installments, even if this is not strategically optimal in a monetary sense. Similarly, these norms may induce peers to punish excessively, e.g., for non-repayment in case of observable shocks.[3] Social norms could help explain two recent puzzles in microfinance research: First, why repayment rates do not differ between individual- and joint-liability contracts, especially when weekly group meetings are held constant (Attanasio et al. 2015; Giné and Karlan 2014). Second, why peer pressure appears to be excessive and sequentially irrational (Czura 2015b; Karim 2008; Montgomery 1996). In addition, and most relevant for our findings, applying the norm may no longer be straightforward when repayment flexibility gives borrowers discretion whether or not to repay at a given point in time. As a result, uncertainty in socially prescribed behavior may lower incentives for repayment.

In our mechanism section, we present suggestive evidence for installment-based social norms. We start by showing theoretically how an exogenous norm affects repayment incentives. To illustrate the basic mechanism, we focus on the case of individual liability. As in the experiment, we model flexibility as the option to postpone individual repayments. We derive theoretical predictions for the timing of repayment and the use of flexibility. These are used to re-examine our empirical findings in more detail. In linking theory and experiment, we interpret the peer punishment we observe

[2] Bicchieri 2016 distinguishes between moral norms – doing something because it is *right* – and social norms – doing something because others expect it. As these concepts interact, and one may emerge from the other, we do not believe this distinction to be meaningful in our setting. For the remainder, we use "norm" and "social norm" interchangeably.

[3] Norms may induce punishment directly, if disciplining non-repaying borrowers is part of the norm (see pledge). Alternatively, peers may be angered by repayment norm violations, and express this anger in the form of punishment (R. Akerlof 2016). Psychological game theory offers several rationales for anger-based punishment, which we discuss in Section 4.4.

in our joint-liability treatments as a reflection of the prevailing social norms.[4] To measure social norms and investigate their alignment with punishment behavior in our lab setting, we conduct an incentivized norm elicitation experiment following Krupka and Weber (2013). In a small sample of borrowers from the same lender (N=44), we find that social norms for repayment mirror the punishment patterns we observe in our experiment: Default is rated less socially inappropriate, and with more dispersion in the ratings, if borrowers use flexibility to defer payments before they default. Our results suggest that flexibility may decrease repayment by creating uncertainty in the socially prescribed behavior for good borrowers.

Our study builds on and contributes to the literature in four ways. To our knowledge, we are the first to identify ex-post moral hazard in a flexible repayment setting, and the first to identify the effect of flexibility on ex-post moral hazard. With this, we speak to two strands of literature. On the one hand, a few studies isolate the role of ex-post moral hazard in lending, and argue that it plays an important role in rigid repayment contracts (Breza 2014; Gertler et al. 2021; Karlan and Zinman 2009). We document that ex-post moral hazard is even more important when repayment timing is flexible. The other strand is a growing literature on flexible repayment schedules. This literature documents positive effects on investments, and mixed evidence on overall default rates (which are affected by both shocks and strategic behavior): Field et al. (2013) study the effects of a grace period between loan disbursement and the start of repayment and find increases in business profits at the expense of higher default. Barboni and Agarwal (2021) study self-selection into flexible microloans, which allow for three-months repayment holidays (with one-month advance notice), but carry higher interest rates than rigid loans. Offering this contract leads to increased repayment rates and business revenues. Neither of these studies give borrowers discretion on whether to repay at a given moment, and thus to condition repayments on shock realizations. Closer to our definition of flexibility, Czura (2015a) examines repayment that allows for occasional skips. She finds suggestive evidence of increased investments, higher income, and higher defaults, though these are obfuscated by a crisis of the lender. Battaglia et al. (2021) offer borrowers with a good credit history to delay up to two monthly repayments at any time. They find improved business outcomes and lower defaults, and argue that the insurance value of flexibility facilitates increased entrepreneurial risk-taking. Most recently, Brune et al. (2022) offer up to three skips in a 12-months loan and find no effect on revenue and profits, but an increase in defaults. The most flexible credit is assessed by Aragón et al. (2020) who find large effects of fully flexible credit lines on short-term profits, but cannot assess effects on defaults. We

[4] See also footnote 3. In our setting, punishments are unlikely to influence repayments as they are designed as costly incredible threats that are small in magnitude relative to the stakes of the repayment choices. Given equal repayment rates between individual and joint liability, our evidence is not consistent with punishment as a direct driver of repayment.

complement this work by isolating the strategic incentives under flexible repayment, independent of borrower selection, project choices, and shock-induced defaults. We argue that flexibility has a hidden cost, in that it may create uncertainty in the socially appropriate repayment behavior. This effect may be masked in the existing field evidence by the observed positive effects on investment choices.

Second, we provide theoretical and empirical evidence on social norms in lending. The importance of social norms has been well-documented outside of credit markets, including norms for risk sharing in village economies (Jakiela and Ozier 2016), productivity in firms (Huck et al. 2012), or xenophobia (Bursztyn et al. 2020a). In the context of lending, attention to the role of social norms for repayment has been limited. Suggestive evidence includes Guiso et al. (2013), who show that survey measures of 'morality' predict defaults on mortgage loans. In a field experiment in Islamic banking, Bursztyn et al. (2020a) find that a moral appeal reduces credit card delinquency. Bu and Liao (2022) and Dhami et al. (2022) point to the the role of guilt and shame for loan repayment in microfinance, consistent with the presence of social norms. We argue that high-repayment equilibria in microcredit may be sustained by social norms. We provide a theoretical framework showing that social norms can be an important determinant for repayment when dynamic incentives are weak. While in a small sample, our incentivized norm elicitation is the first of its kind in microfinance.

Third, we add to the literature on mechanisms to overcome credit market failures. Similar to Giné et al. (2010) and Dhami et al. (2022), we take a systematic approach to unpacking such mechanisms in the context of microfinance.[5] One prominent mechanism in this literature is the role of social capital in maintaining high repayment rates (Armendáriz 1999; Besley and Coate 1995; Ghatak and Guinnane 1999; Karlan 2007; Wydick 1999). Social capital predicts differences in repayment between individual- and joint-liability contracts (De Quidt et al. 2016). However, recent empirical studies do not find different repayment rates between individual- and joint-liability contracts, especially when weekly group meetings are held constant (Attanasio et al. 2015; Giné and Karlan 2014). Our evidence for social norms as a driver of loan repayment helps to explain this puzzle. While we are ultimately agnostic about the source of these norms, a possible microfoundation comes from the work on identity in organizations (G. A. Akerlof and Kranton 2005). Identity models can also explain the comprehensive measures taken by lenders to instill a culture of good borrowing. With the exogenous variation in the liability structure, we further help to consolidate findings from previous studies on flexible loan repayment, which use both individual- and joint-liability contracts.[6]

[5] While Giné et al. (2010) focus on ex-ante moral hazard and project choice as determinants of loan repayment in rigid contracts, we focus on ex-post moral hazard by separating repayment capacity from the choice to repay. Dhami et al. (2022) investigate psychological mechanisms in the context of ex-ante moral hazard.

[6] Field and Pande (2008) study joint-liability group loans; Field et al. (2013) study individual-liability group loans; Battaglia et al. (2021) study individual-liability group loans but only offer flexibility to clients with a good repayment

Fourth, as one of the few papers to measure and quantify peer punishment in a microfinance context, we are the first to study its interaction with the contract structure. While social capital is considered critical to high repayment rates in microcredit, standard models of group lending universally predict zero punishment in equilibrium: The credible threat of social sanctions is enough (Armendáriz 1999; Besley and Coate 1995). By contrast, a rich literature on coordination games documents that people frequently engage in costly and non-credible punishment (Fehr and Gächter 2000, 2002; Henrich et al. 2010, 2006; Masclet et al. 2003). Evidence on peer punishment in microfinance has been largely qualitative or anecdotal.[7] Czura (2015b) is the first to document and quantify excessive punishment in microfinance, relative to both game-theoretical and fairness-based benchmarks. We confirm and complement her findings by showing how excessive peer pressure reacts to changes in repayment structure.

The paper proceeds as follows: Section 2 describes the experimental design, the procedures, and the study setting. Section 3 outlines the empirical strategy. In Section 4, we present experimental results on repayment, flexibility use, and punishment. Section 5 presents theoretical and empirical evidence for social norms. Section 6 discusses limitations of the study and concludes.

2 Experiment

2.1 Design

We design a microcredit repayment game to analyze ex-post moral hazard. The experiment is set in a world of asymmetric information, where borrowers but not lenders observe project returns. We take a partial equilibrium approach, and focus entirely on the borrower's decision whether or not to repay the loan once project returns are realized. The borrower's monetary incentives are determined by a stochastic income process, the repayment installments, the future benefits from repaying the loan, and any discounting. At this point, the loan amount and the investment choice are sunk. Further, we do not model the lender - the market interest rate enters exogenously (through the required repayment amounts), and thus our results apply independently of the lender market structure.

The experiment exogenously varies two features of the loan contract: First, the repayment schedule is either rigid or flexible. Rigid schedules impose repayments every period, while flexible

record; Barboni and Agarwal (2021) study former borrowers in joint-liability group loans that have accumulated a good repayment history and are now promoted to receiving an individual-liability loan.

[7] A large body of qualitative evidence comes from anthropological case studies: Montgomery (1996), Rahman (1999), and Karim (2008) report cases of drastic social pressure on defaulting borrowers, such as verbal harassment, shaming in public, raiding of houses to confiscate assets for sale to cover the loan installments, or stripping down the defaulter's house completely. Baland et al. (2017) point out that punishment may take the form of social exclusion.

schedules allow for discretion in when to repay. Second, borrowers are either individually or jointly liable for their loans. Joint-liability loans include the possibility of punishment from borrowing peers. We describe the resulting four treatments in detail below and discuss major design choices and their implications in Section 2.3.

To facilitate clean comparisons across different liability structures, we focus on individual choices in all treatments. Choices are elicited using the strategy method to cover the full action space: We elicit decisions for different states of the world – here, for different realizations of income.[8] The ability to separately observe repayment capacity and repayment choices allows us to identify moral hazard. To build on social capital that participants have accumulated outside the lab during their microcredit borrowing routine, we use loaded framing and explain the setup and all choices in the context of microfinance.

Individual liability (IL)

Setup The standard game models a simple credit repayment choice under risk over three periods. An individual takes out a loan which is automatically invested into a risky project and generates a per-period income of $y_t = 2R$ with probability $1-\theta$, and $y_t = 0$ with probability $\theta = 0.25$. The loan requires a repayment R in periods $t=1,2,3$, where the total repayment of $3R$ covers both loan principal and interest. In the experiment, each R is represented by one income token and the shock is framed as a thief who steals the entire income of a given period. Each period, the individual makes a choice between two actions: make the required repayment R (and consume her remaining income R), or consume her entire income $2R$. The individual cannot save. When $y_t=0$, neither repayment nor consumption is possible. A loan is considered to be in default after the first non-repayment, whether due to choice or bad luck, for the rest of the game. Repayment in all periods yields a 'continuation value' V, a future benefit that reflects dynamic incentives, such as the utility from access to future loans or improved loan conditions (Giné et al. 2012).

Payoffs In the experiment, V is a payment of 100 pesos, paid in cash one month after the experimental sessions. In contrast, all experimental income allotted to consumption (income not spent on repayment nor lost to a shock) can be spent right after the session on a vast selection of consumption items. In the spirit of Jackson and Yariv (2014) and the shrinking pies in bargaining experiments (see

[8] The strategy method was first introduced by Selten (1967). Brandts and Charness (2011) review 29 comparisons of the strategy method with the direct-response method, and find that treatment effects observed with the strategy method are in all cases also observed with the direct method. In contrast, results from the strategy method constitute a conservative lower bound for emotionally-motivated outcomes, such as punishment.

Roth (1995) for a review), we induce discounting across periods by reducing the consumption value of income tokens: One token R is worth 40, 30, and 20 pesos in period 1, 2, and 3, respectively, implying that future repayments are discounted. Consequently, $3R$ from one period each are worth 90 pesos.[9]

Strategic considerations We consider two main strategies to be important in *IL*: first, choosing to repay when not encountering shocks (*RRR*) and second, defaulting on all three installments (*DDD*). In the presence of income shocks, the expected payout from always repaying (and receiving V if no shocks arrive) is 122 pesos (see Appendix D for an overview of expected payouts). The expected payout from *DDD* is 135 pesos. A payoff-maximizing and risk-neutral individual should therefore choose to default.[10] A person who places value on being a good borrower (we discuss reasons for this below), might sacrifice some expected payoff and choose *RRR*. Our payoff calibration is not intended to exactly match default rates observed in the field (see Section 2.3). Instead, it provides a theoretical benchmark of no repayment, and thus allows us to discriminate between payoff-maximizers and participants who are motivated by additional factors.

Elicitation of choices Throughout the experiment, we use the strategy method (see footnote 8) to elicit decisions. In *IL*, we ask participants for their repayment choice in each period if no shock arrives in the given period, that is whether they would like to repay *R* and consume *R*, or not repay and consume their entire income *2R*. We omit the repayment question for the case that a shock arrives, as participants have no choice to make in this case (this changes in the flexibility treatments). Further, we ask for the repayment choice if a shock arrived in the previous period, i.e., repayment after a loan is in default. Shocks are only realized at the end, as explained in Section 2.2, which implies that there is no feedback between periods.

Individual liability and flexibility (IL-flex)

Setup We design flexibility as the option to defer a repayment installment to the next period. This option is framed as a pass token that sets the repayment obligation for the current period to zero, but requires a double repayment in the subsequent period. By using the pass token when an income shock arrives, borrowers can prevent defaulting on the current installment. Each borrower receives one pass token, which can be used in either period 1, in period 2, or not at all (see Figure 1). Flexibility cannot be used in period 3, which serves as a catch-up period for repayments from period 2. Failure

[9] The exchange rate in March 2016 was 51 PHP per EUR. The average daily income in our sample was about 200 PHP.

[10] Adding risk aversion as well as any temporal discounting between the session and the payment of V one month later would further increase the appeal of default.

to make a double repayment results in default, as do shocks once the pass token has been used. We assume that it is not possible (or not cost-effective) for the lender to observe shocks, which means borrowers can use flexibility independent of shock arrival. Since flexibility is intended as a shock-coping mechanism, we label the use of flexibility to increase consumption as 'misuse' for the analysis. Experimental instructions were neutral and non-judgmental.

Strategic considerations Rather than reserving flexibility use for the self-insurance of idiosyncratic shocks, the borrower may choose to misuse flexibility to increase early consumption by delaying repayment until the next period. It is tempting to do so: Immediate consumption increases by R, while the future loss is $\delta(1-\theta)R$, where δ captures the experimentally induced reduction of R's purchasing power over time.[11] This creates a trade-off in period 1: Using the pass token in period 1 means that it cannot be used to insure shocks in period 2. The probability of a shock-induced default increases from θ (period 3 shock) to $\theta+(1-\theta)\theta$ (shock in periods 2 or 3). The trade-off between early consumption and default risk is reflected in the expected payouts (Table D.16): Misuse in period 1 becomes more attractive as the (delayed) continuation value V is more heavily discounted. There is no trade-off to flexibility use in period 2. Irrespective of the timing of flexibility use, the risk of default due to shocks is lower in *IL-flex* than in *IL*, as shocks are always insurable in period 1. As a result of this insurance provision, as well as the reduced effective interest rate from delayed payments, the overall monetary incentive to repay the loan is increased in *IL-flex* relative to *IL* (Table D.16).

Elicitation of choices The pass token enlarges the action space, so we elicit more choices in *IL-flex*. Borrowers state their repayment choices and use of flexibility conditional on the arrival of income shocks. Figure F.4 illustrates the decisions in IL-flex when income shocks are possible ex ante but do not arrive ex post.

Joint liability (JL)

Setup We model joint liability as a two-person borrowing group that is jointly responsible for repaying $2R$ in each period. Participants make individual choices knowing that they will be randomly and anonymously matched with a partner from the same session, i.e., a real-life borrowing peer, to calculate their payoffs. Borrowers thus simultaneously choose whether to repay or not and cannot condition their choices on the partner's choice or shock realization. Joint liability is enforced

[11] Note that our setting controls for present bias, in the sense that periods are too close together in real time to induce natural discounting (and consumption occurs at the end of the session). Present bias would not affect choices between periods (and thus decisions about flexibility use), but rather rescale the value of V (which is paid one month later). Empirically, we see no difference in repayment by measures of present bias (Table A.15).

automatically: If they choose to repay, but their partner does not, they automatically repay for their partner as well. The bank does not distinguish between the source of repayment: As long as $2R$ is repaid in each period, both borrowers will receive V.

A measure of peer pressure is introduced via the possibility to send punishment points (framed as 'dislike' tokens) to one's partner. Participants can choose between allocating zero, one or two punishment points for a given action of their partner. Each point costs the sender five pesos of her 70 pesos show-up fee, and reduces the receiver's show-up fee by 15 pesos.

Strategic considerations As in *IL*, the two main strategies in *JL* are *RRR* and *DDD*. While free-riding in only one or two periods is possible, our setup does not allow borrowers to condition own choices on partner's choices. Given simultaneous repayment choices and no feedback, the key difference to *IL* is in payoffs: A decision to repay in a given period costs either one or two income tokens, depending on the unknown repayment choice of an anonymous partner in the session. The repayment decision can thus be understood as a signal of repayment capacity, in which case the borrower is held liable for her partner's repayment. Choosing to repay might result in no consumption in the given period in case the partner does not repay. In terms of punishment, all choices are incredible threats as partners only learn whether they were punished after making their own repayment decisions.

Elicitation of choices For repayments, the elicitation of choices proceeds as in *IL*. Punishment is implemented for one randomly selected period. We elicit three choices conditional on the arrival of shocks: punishment in case the partner 1) repays, 2) does not repay and faces no shock, and 3) does not repay and faces a shock (see upper panel of Figure F.6). In addition to these incentivized measures for repayment and punishment, we ask for (non-incentivized) beliefs of the partner's repayment and punishment choices.

Joint liability and flexibility (JL-flex)

Setup We examine the interaction of joint liability and flexibility in a two-person borrowing group, where both partners have one pass token and can defer one repayment installment to the next period (see Figure F.5). The same rules apply as in *JL* and *IL-flex*.

Strategic considerations Borrowers can now choose between self-insurance and mutual insurance when a shock arrives. Mutual insurance may be associated with peer punishment, even when the borrower is mechanically unable to repay (Czura 2015b). If peers punish when they have to repay for their partner, self-insurance through repayment flexibility may avoid punishment, but comes at the cost of making a double repayment in the next period.

By design, mutual insurance and self-insurance through flexibility crowd each other out: In a period when a borrower uses the pass token, her repayment obligation is reduced to zero. She cannot simultaneously insure her partner's repayment (for instance, because this would reveal to the lender that she does not have a shock). In the next period, the borrower needs her full income for her own double repayment, which again leaves no scope for insuring her partner. If she faces a shock when the double repayment is due, her partner cannot insure her, since the group repayment obligation $3R$ exceeds the group income $2R$.

Elicitation of choices The elicitation of repayment choices proceeds as in *IL-flex* (compare Figures F.4 and F.5). As in the *JL* treatment, incentivized punishment decisions are elicited for each single-period action of the partner (which now include flexibility use and its repayment), and conditional on the arrival of shocks (see Figure F.6). In addition, we ask for (non-incentivized) beliefs about the partner's use of flexibility.

2.2 Procedures

Randomization We use a mixture of a within- and between-subject design. We first randomize the liability structure at the session level, and then vary flexibility treatments within individuals: IL-Sessions run *IL* and *IL-flex* treatments, while JL-Sessions run *IL*, *JL* and *JL-flex* treatments. While *IL* is run in both session types to facilitate comparisons, time constraints made it impractical to run all four treatments in the same session. For similar reasons and to facilitate comprehension, we do not vary the order of treatments, but allow them to naturally build upon each other. Appendix C discusses the consistency of our findings with the presence of order effects between the treatments.

Realization of payments At the end of the session, we randomly select one of the treatments to be paid out. Participants realize the shocks for each period by drawing chips from a black bag, which contains one shock chip and three non-shock chips (capturing $\theta = 0.25$). In *JL* treatments, we then match participants randomly and anonymously with a partner from the same session to calculate payoffs. Punishment is implemented for one randomly selected period, based on repayment choices and the shock realization. Average earnings amounted to 202 pesos (roughly four euros), which equals approximately a daily wage for our sample population. There were three types of payments: First, the show-up fee of 70 pesos was paid in cash at the end of the session. It was reduced by any punishment activity (five [ten] pesos for sending one [two] punishment tokens and 15 pesos for each punishment token received; so a maximum of 40 pesos could be deducted). Second, the continuation value V was paid as 100 pesos in cash, handed out by a research assistant in the borrowing center one month after the session.[12] Third, the income tokens earned in the microcredit game could be traded for items from a consumption table (see Figure G.7), containing a variety of products such as sweets, food staples, household items and beauty products, offered at typical market prices. All consumption was paid out in kind, which captures the temptation of immediate consumption and prevents the use of experimental payouts for non-consumption purposes, such as loan repayment.

Session organization Sessions lasted on average about three hours. After registration, participants completed a small individual survey covering incentivized measures of risk and time preferences over money, as well as survey questions regarding their borrowing group. The general setup of the microfinance repayment game was explained extensively using flip chart graphics, test questions, and three practice rounds. Unlike in the main experiments, shocks were realized immediately in the practice rounds to allow for experimentation with different actions and immediate feedback. Each of the treatments was explained in the same manner and test questions were asked. If more than five participants failed a specific question, the explanation was repeated before final choices were made. Choices were noted in private by local research assistants using paper and pen.

2.3 Discussion of Design Choices

To focus on the interaction of repayment flexibility and joint liability, as well as on repayment choices of individual borrowers, the experimental design entails several simplifications that merit discussion.

Rigidity of Repayment Schedule Our baseline repayment schedule is rigid in that it does not allow for incomplete or late repayments. The severity of the shock excludes partial repayment within a period, and savings constraints prevent borrowers from making repayments using past income.[13] The dynamic incentive V is lost as soon as one installment is missed. These simplifying assumptions lead to a binary and stylized classification of default, and default rates which are higher than those observed in the field. Yet, there is a direct correspondence with real-life borrowing settings: Our default measure captures any form of delinquency that is punished by the lender by withholding

[12] At the end of the session, participants received a voucher to confirm the payment. Trust is unlikely to be a significant concern given the long-standing reputation of the lender and the lender's regular weekly interactions with the borrowers.

[13] Savings constraints are a standard assumption in microfinance games (Abbink et al. 2006; Giné et al. 2010) and have been well-documented empirically (see e.g. Bauer et al. (2012) on present bias and Baland et al. (2011) on financial pressure from relatives or friends).

future benefits to the borrower. Delinquencies are important for the lender: First, it is costly for the lender to chase after delinquent borrowers (Aleem 1990). Second, they indicate the quality of the loan portfolio. Measures of delinquency, including the portfolio-at-risk, guide decisions on loan restructuring and loan loss allowances. Last, they also directly affect the lender's profitability, since the loan portfolio is revolving more slowly (Athreya et al. 2018; Rosenberg 1999). As a result, lenders frequently condition dynamic repayment incentives on reliable on-time repayment, rather than simply on non-default. Our lender, for example, offers larger loan sizes and improved conditions as well as different loan types only to borrowers who always repay their weekly installments. Delinquent borrowers forego these dynamic repayment incentives, even if they end up completely repaying their loan. In our experiment, V captures such dynamic repayment incentives.

Implementation of *JL* In order to study the effects of the liability structure while keeping the setup simple and comparable across *IL* and *JL* treatments, we take the following measures: We have two-person borrowing groups, repayment choices are taken simultaneously such that they cannot be conditioned on the partner's choices, and joint liability is automatically enforced. These design choices are commonly used simplifications in microfinance lab experiments (Abbink et al. 2006; Cason et al. 2012; Cassar et al. 2007; Giné et al. 2010). First, the reduction of the usual five- to two-person groups makes strategic considerations regarding partner's choices easier. Compared to individual liability, a two-person group retains the behavioral aspects of joint liability, for which group size (2, 4, or 8) matters only weakly (Abbink et al. 2006). Second, not allowing for communication or feedback on partner's choices allows us to focus on the effect of liability. Third, automatic enforcement of joint liability reduces the decision space, which is important to focus on repayment choices and ex-post moral hazard. Moreover, it is a realistic representation of how microfinance institutions put joint liability into practice. For example, our partner organization instructs the loan officer to prolong the weekly group meeting until all repayments are made.

We allow for peer punishment in *JL* treatments but not in *IL* treatments. Our peer punishment is specific to borrowers who are jointly liable for each other, and thus cannot be directly applied to individual-liability contracts.[14] We do not separate joint liability from peer pressure, as joint-liability lending is fundamentally based on social capital (Besley and Coate 1995; Ghatak and Guinnane 1999; Karlan 2007; Stiglitz 1990). Finally, we restrict punishment to one random period. It would have been more realistic to condition punishments on past repayment history, or to allow punishments

[14]Peer punishment between borrowing peers is conceptually distinct from general third-party punishment, such as reputational loss in the community following norm violations. The latter is captured in our separate norm elicitation experiment (Section 5.4).

every period. However, the dimensionality of this would have been prohibitive in our experimental setting. Our single-period punishments can be used to calculate the expected level of punishment for any three-period strategy, see Section 4.3.

Implementation of flexibility Our design focuses on the consumption versus insurance trade-off from giving borrowers the ability to ad-hoc postpone repayments. Three periods and one pass token are the minimum required to do this. The assumption that shocks in period 3 are not insurable scales the expected benefit of repayment by $(1-\theta)$. Since this holds across treatments, it does not affect the game dynamics. We do not attach a price to the use of flexibility to avoid a confound from an increased repayment burden (as in Karlan and Zinman (2009)). Our estimate of the effect of flexibility on ex-post moral hazard is thus a conservative lower bound.

With joint liability, repayment flexibility provides borrowers with an alternative mechanism to insure their repayment against income shocks. For covariate shocks, flexibility provides a superior shock-coping mechanism relative to mutual insurance. For individual shocks, flexibility can help as well. However, using flexibility in these cases may partially crowd out the capacity to provide mutual insurance in real lending groups: In the presence of savings constraints, allowing borrowers to bunch repayment installments together puts added pressure on the current period's income, which decreases their capacity to insure others.[15] Our design is stylized and highlights this trade-off.

2.4 Study Setting and Sample Recruitment

We conducted experimental sessions in 33 borrowing centers of the microfinance institution Ahon Sa Hirap (ASHI), across three provinces of the Philippines: Rizal, Laguna, and Antique. All clients are organized in groups of five borrowers. Each group is part of a borrowing center, consisting of two to eight groups, in which weekly repayment meetings take place. Of the 33 centers, 27 centers (covering 82 percent of our participants) offer joint-liability loans for general business activities. Joint liability is enforced both within the borrowing groups, and between groups on the center level. The remaining six centers – all in the more rural Antique province – offer loans with individual liability for agricultural production. Despite this variation, all clients attend weekly group repayment meetings in their center. Regular joint-liability loans are repaid over 25, 50 or 100 weeks. Individual-liability agricultural clients service only interest payments on a weekly basis, and repay the principal at

[15] Compare also G. Fischer and Ghatak (2016), who show theoretically that small and frequent repayments are more incentive-compatible for present-biased borrowers than allowing them to delay and bunch installments. In our setting, present bias would not affect choices between periods (footnote 11). Rather, present bias is one way to microfound the savings constraints which are built into our design.

harvest time (up to six months after loan disbursal, depending on the crop cycle). Loan sizes range from 2,000 to 100,000 pesos, and average 14,350 pesos (281 EUR) for the most recent loan.[16] The typical annual interest rate is 46 percent on a declining balance.

Importantly for the interpretation of our results, the lender takes various measures to instill a strong culture of good borrowing: In both individual- and joint-liability centers, borrowers and loan officers jointly recite a pledge at every weekly meeting (similar to the Grameen pledge quoted at the start of this paper), in which they promise to faithfully make their repayment installments and take responsibility for one another. In addition to the weekly meetings, social activities are organized at the center level to build solidarity between borrowers.

We identified borrowing centers with at least 20 borrowers and a center meeting hall with seating. From the exhaustive member list of these centers, we randomly selected 20 members to be invited for participation; five members were invited as back-up. Invitation letters were handed out one week in advance during the center meeting. Sessions took place in the center meeting hall on different days than the weekly meetings. Participation was voluntary, and participants were assured that their choices in the experiment would not be revealed to the lender.

In total, 645 participants took part in 33 sessions (one per center) in March and April 2016. Our main analysis sample consists of 577 participants: three participants left after the intake survey, the decisions of 37 participants cannot be analyzed due to enumerator errors in recording answers, and 28 participants did not pass our comprehension test.[17] Table 1 presents background characteristics of our participants, and shows that session type (IL vs. JL) is balanced on observables. Our sample is predominantly female, and on average 47 years old. Around half have completed secondary school. The main sources of household income are own non-farm businesses (47 percent) and farming (26 percent). Forty-five percent of our sample households live below the national poverty line (national average: 21 percent), as measured by the PPI index. Eighty percent are connected to the electricity grid, 23 percent to piped water, and two percent to the landline telephone grid. Three-quarters live in a house with an iron roof (as opposed to a palm roof).

3 Empirical Strategy

A key advantage of our lab-in-the-field experiment is that we can observe repayment choices separately from realized outcomes. Given our interest in ex-post moral hazard, our analysis focuses

[16]This average includes "incentive loans" for non-investment purposes, which serve as a dynamic incentive for borrowers who have repaid their regular loan on-time without delinquencies.

[17]We excluded participants from our main analysis if less than 75 percent of test questions overall or 50 percent of test questions from any one treatment were answered correctly. This exclusion does not affect our results (see Appendix C).

on individual choices in response to contract design features. In particular, we examine choice data regarding loan repayment, the use of flexibility, and peer punishment.

Overall Loan Repayment We identify ex-post moral hazard as the fraction of participants who fails to repay their loan in the absence of income shocks, i.e., despite being fully capable to repay. To do so, we follow individuals' repayment choices along the *no-shock path*: In each period, participants choose to repay or not, conditional on not suffering a shock in the current period, but without knowing whether shocks will arrive in the future. The no-shock path refers to the path of the game tree where shocks are possible ex ante, but do not arrive ex post. This is a useful concept for analysis purposes: The borrower is able to repay in all periods, and any failure to do so must be the result of moral hazard.[18] Our main outcome of interest is a binary variable for full repayment – meaning the individual either repays every period, or uses flexibility and then repays.[19] To make treatments comparable, we apply this variable definition to choices in both individual- and joint-liability conditions, and abstract from group repayment outcomes. We estimate the effect of flexibility on repayment using the linear probability model

$$Repay_{its} = \alpha + \beta_F flexible_t + \lambda_s + \epsilon_{its} \tag{1}$$

where $Repay_{its}$ indicates full repayment of individual i in treatment t in session s, and the binary variable $flexible_t$ switches on for treatments with flexible repayment conditions (*IL-flex* or *JL-flex*). Repayment regressions use within-individual variation in flexibility, and are run separately by session type: We estimate the effect of flexibility on repayment in IL-sessions by comparing choices in treatments $t = \{IL, IL\text{-}flex\}$ and in JL-Sessions by comparing $t = \{JL, JL\text{-}flex\}$. The coefficient β_F thus estimates the effect of flexible repayment for a given liability structure. We include session fixed effects λ_s and cluster errors ϵ_{its} at the level of the individual. We additionally estimate the effect of joint liability on repayment by running

$$Repay_{its} = \alpha + \beta_L joint_t + \lambda_s + \epsilon_{its} \tag{2}$$

for treatments $t = \{IL, JL\}$, using the within-individual variation in liability structure contained in JL-Sessions. The indicator $joint_t$ is equal to one if treatment $t = JL$ and zero otherwise, other variables are defined as above.

[18]The concept of the no-shock path has no bearing on the way choices were incentivized (see Section 2.3).

[19]In individual-liability conditions, this is equivalent to the repayment of three income tokens. In joint-liability conditions, full repayment costs between three and six income tokens given automatic enforcement (see Figure F.5).

Use of Flexibility We further study the effect of the liability structure on the use of flexibility, i.e., the choice to defer payments. The liability structure was randomized between sessions, leading to a between-subject design that compares the *IL-flex* and *JL-flex* treatments. Two distinctions are necessary: Flexibility can be used in case of shocks or absent shocks, and it can be used earlier (period 1, thus foregoing insurance if used absent shocks) or later (period 2). We index the resulting four scenarios by $c =$ {*T1 no shock, T1 shock, T2 no shock, T2 shock*}, and create a binary variable $Flexuse^c$ for whether a participant chooses to use flexibility in a given scenario. We use a linear probability model to estimate

$$Flexuse^c_{its} = \alpha + \beta^c_U joint_t + \epsilon^c_{its} \tag{3}$$

where $Flexuse^c_{its}$ indicates flexibility use in scenario c by individual i in treatment t in session s. The binary variable $joint_t$ now switches on for treatment $t =$ *JL-flex* (the omitted category is *IL-flex)*, and β_U is the effect of the liability structure on the use of flexibility. Finally, since liability was randomized between sessions, we cluster errors ϵ^c_{its} at the session level, resulting in 33 clusters for equation 3.[20]

Punishment Our two joint-liability treatments, *JL* and *JL-flex*, allow for peer punishment. We analyze punishment for repayment and flexibility choices, conditional on shock realizations. Since flexible repayment expands the choice set, we compare the non-repayment choice in *JL* to each non-repayment action with flexibility, i.e., misuse of flexibity, non-repayment, and non-repayment of flexibility. For each choice pair, we run OLS on

$$Punish_{its} = \alpha + \beta_P flexible_t + \lambda_s + \epsilon_{its} \tag{4}$$

where $Punish_{its}$ denotes the level of punishment by individual i in treatment t in session s. For a given choice of the partner, the level of punishment is the number of punishment tokens chosen (0, 1, or 2). We express punishment as a proportion [0,1] of the maximum possible punishment to facilitate later comparisons with our norm elicitation study. The treatment variable $flexible_t$ is an indicator for treatment $t =$ *JL-flex* (the omitted category is *JL*), and β_P is the effect of flexible repayment on punishment for a given choice combination. As in previous specifications using within-individual variation in flexibility, we include session fixed effects λ_s and cluster errors ϵ_{its} at the level of the individual.

[20]We also provide wild cluster bootstrapped confidence intervals that account for the small number of clusters.

4 Results

The next two sections present our key results on the cost and benefits of flexible repayment. Regarding ex-post moral hazard, we find no difference in repayment rates across liability structures, but higher strategic default with flexibility. Regarding social pressure, we document high levels of peer punishment which are reduced when introducing flexibility. We show that both these results are consistent with the presence of social norms, and provide several pieces of supporting evidence in Section 5.

4.1 Overall Loan Repayment

Overall, repayment rates are high: In all treatments, more than 50 percent of participants repay, despite the fact that repayment was designed to be monetarily unprofitable (see Section 2.1). As Figure 2 shows, flexibility has a substantial impact on repayment behavior: *IL-flex* reduces repayment by 16.5 percentage points relative to *IL*. Equivalently, strategic default increases by 46 percent. Numbers are similar with joint liability: Looking at individual choice data, *JL-flex* reduces repayment by 16.4 percentage points relative to *JL*, equivalent to a 58 percent increase in strategic default on the overall loan (when evaluated at the individual level).

We find no significant differences across liability structures, neither with nor without flexibility. In *IL* and *JL*, 66.2 and 66.5 percent of participants fully repay their loan, whereas in *IL-flex* and *JL-flex*, 51.3 and 50.0 percent do so.[21] The group features in the joint liability setting – mutual insurance and peer punishment – do not appear to influence individual repayment choices on average.[22] We summarize these findings in Result 1.

Result 1. Repayment rates are high relative to monetary incentives, and do not differ across liability structure. Flexibility increases strategic default by 16 percentage points (50 percent).

4.2 Use of Flexibility

Each participant has one pass token, which allows her to postpone a repayment either in period 1, in period 2, or not at all. Borrowers can use this flexibility to insure their repayment capacity against an income shock (henceforth 'self-insure using flexibility'), but they can also misuse it to increase early consumption absent shocks. We focus on flexibility use in period 1, when the trade-off between

[21] Results are robust to including individual fixed effects, see Table A.2.

[22] We also find no differences by an incentivized Binswanger-style measure of risk aversion (Tables A.12, A.13 and A.14), suggesting that the added consumption uncertainty from joint liability does not drive our results.

early consumption and insurance is most pronounced.[23] We observe near-universal use of flexibility in case a shock hits, with no difference between the *IL-flex* and *JL-flex* treatments (left panel of Figure 3). This indicates that participants understand the insurance value of flexibility. For participants in *JL-flex*, we additionally infer that self-insurance against income shocks is widely preferred to mutual insurance by their borrowing peer. This is notable insofar as self-insurance through flexibility requires a double repayment in the next period, while mutual insurance does not.

We also observe substantial use of flexibility absent shocks: 55 percent of participants in *IL-flex* and 29 percent of participants in *JL-flex* misuse flexibility when no shocks arrive in period 1 (right panel of Figure 3). The lower rate of misuse in *JL-flex* is consistent with increased costs: Flexibility use eliminates mutual insurance possibilities in the current and the next period.[24]

Result 2. Flexibility is used to insure income shocks. However, there is substantial misuse of flexibility to increase early consumption, especially in individual-liability contracts.

4.3 Peer Punishment

The strategy method provides us with punishment choices for each action the partner can take (see Figure F.6). We first discuss punishment for single-period actions. Because borrowers plausibly choose three-period strategies rather than independent actions, we subsequently calculate the expected level of punishment for key three-period strategies, internalizing the risk of shocks.

Single-Period Punishment To facilitate later comparisons with our norm elicitation, we report punishment levels in shares of the maximum possible punishment (two tokens). Punishment is widely used. Figure 4 shows that non-repayment absent shocks is punished with 61 (60) percent in *JL* (*JL-flex*). In *JL*, we also find high levels of punishment (38 percent) when the partner cannot repay due to a shock. Since shocks are fully observable and make it impossible to repay, this result underscores recent concerns about excessive or anti-social peer pressure in microfinance (we discuss this further in Section 5.1). Surprisingly, even repayment actions receive non-zero punishment levels (14 and 9 percent, respectively), potentially to uphold a general sense of pressure.

Flexibility gives rise to additional actions. When hit by a shock, participants can either self-insure by using their pass token, or rely on their partner to repay. Punishment for these two cases is

[23] We present results for period 2 in Section 5.3.

[24] Given the crowd-out between self-insurance and mutual insurance, there is some evidence that peers attempt to coordinate their use of flexibility: Using (non-incentivized) beliefs about partner's behavior, we find that participants' use of flexibility correlates with their belief that their partner will use flexibility, both in the case of shocks (Spearman's ρ=0.137, p=0.010) and without (Spearman's ρ=0.279, p$<$0.001).

shown in the middle two bars of the right panel in Figure 4. Using flexibility reduces the level of punishment by 20 percentage points, corresponding to a 53 percent decrease relative to punishment for a shock in *JL*.[25] In contrast, relying on one's partner to insure shocks increases the level of punishment by 14 percentage points (37 percent) as compared to punishment for a shock in *JL*.[26] This behavior indicates that self-insurance through flexibility is clearly preferred over relying on the partner to repay.

Absent shocks, flexibility provides a way to reduce the punishment for defaulting on one's loan: The right panel of Figure 4 reveals that misuse of flexibility is punished less (39 percent) than simply defaulting on an installment (60 percent), despite the fact that no repayment occurs in either case. This is not compensated by a significantly higher punishment for defaulting on the subsequent double installment (66 percent). Columns 4 and 8 of Table A.4, Panel A, confirm that neither the level nor the incidence of punishment increase significantly when comparing single-installment to double-installment default.[27]

Expected Punishment for Strategies While single-period punishments are illustrative, what borrowers ultimately face is the expected punishment across their three-period strategy. Recall that punishment was paid out for a random period. The expected punishment for a strategy is the average punishment over all one-period actions in it, where each action is weighted by the probability that the action is applicable (which depends on the arrival of shocks). Figure 5 and Appendix E show the expected punishment for key strategies. All main insights from single-period punishments are confirmed when considering the expected punishment for entire strategies. In particular, the expected punishment for a full repayment strategy is reduced by 30 percent in the presence of flexibility, as participants can self-insure income shocks and thus suffer fewer penalties from having to rely on their partner. On the other hand, we observe that flexibility can reduce the expected punishment for strategic default on the overall loan: Borrowers who want to default can dodge 15 percent of the punishment by first using flexibility to postpone repayments, and then defaulting on all subsequent installments. This strategy is monetarily equivalent to straight default, but socially more sophisticated. Appendix E provides further details.

Result 3. Flexibility reduces the punishment for missing installments due to shocks (excessive punishment) by half, and thus the punishment of borrowers who repay their loan and use flexibility

[25]Regression results for both level and incidence of punishment are shown in Panel B of Table A.4.

[26]For implementation reasons, we were not able to distinguish situations by whether the pass token is still available but not used, or no longer available. Thus, 'Don't use flex (shock)' refers to any situation in *JL-flex* where a shock hits and flexibility is not used.

[27]Ceiling effects cannot explain this phenomenon: Recall that participants assigned either zero, one, or two punishment points for a given action of their partner. Single-installment default is punished with zero (12 percent), or one point (56 percent), which means that a majority is able to increase the punishment for double-installment default if they want to.

responsibly. However, flexibility also reduces the expected punishment for strategic default.

4.4 Discussion of Results

We presented several key repayment results. First, we find that repayment rates are high relative to monetary incentives, and do not differ across liability structures. Second, repayment decreases with flexibility, again independent of the liability structure.

The high share of borrowers choosing to repay in all three periods in our lab-in-the-field experiment is consistent with the near-complete repayment rates that the partner institution reports for its borrowers: In the years 2014 to 2018, the repayment rate was consistently around 96 percent. A qualification is warranted that loans are not considered in default and written off until several years after maturity, and that delinquencies were substantially more common. Using administrative data from the lender, there is a weak positive correlation of real-life behavior as captured by different borrower quality measures and behavior in our experiment (see Table A.10).

Our finding that flexibility lowers repayment is in line with Field et al. (2013), who find higher defaults with a grace period, and Brune et al. (2022), who document larger defaults with repayment skips. It stands in contrast to Barboni and Agarwal (2021) and Battaglia et al. (2021), who find lower defaults with temporary repayment waivers made available to selected clients with a good repayment record. All four studies include endogenous project selection (ex-ante moral hazard), and defaults that may be driven by shocks.[28] We abstract from these and show that *strategic* defaults increase with flexibility. The estimated effect size is large, but in line with other microfinance experiments. For example, Giné et al. (2010) find treatment effects of around 30 percent on loan repayment with variations in dynamic incentives, monitoring, and partner choice.

Our finding that the liability structure does not affect repayment is in line with Giné and Karlan (2014) and Attanasio et al. (2015), who find similar repayment rates in individual- and joint liability contracts in randomized field experiments. Taken together, these results suggest that the liability structure per se (including the threat of punishment from jointly liable peers) does not meaningfully drive repayment behavior.

If punishment from borrowing peers is ineffective, then why do people punish? We observe high levels of punishment for missing installments due to (fully observable) shocks. Flexibility halves this punishment for borrowers who use flexibility to self-insure their loan repayment, but it also reduces punishment for strategic default. It is difficult to rationalize the observed levels

[28]One way to reconcile the mixed evidence on overall default rates in the field is that different flexibility designs affect ex-ante project choice (and thus risk) in different ways.

of punishment with expected payoff maximization. Punishment is costly, and not credible in the sense that punishment decisions are revealed only after repayment choices have been made (see Section 2.1). However, non-credible punishment is frequently observed in the literature (Fehr and Gächter 2000, 2002; Henrich et al. 2010, 2006; Masclet et al. 2003). There is broad consensus that peer punishment depends on intentions for noncooperation (R. Akerlof 2016; Battigalli et al. 2019; Charness and Levine 2007; Rand et al. 2015). Peers' inferences about intentions or types may explain the punishment we observe for default or flexibility misuse, but it does not explain why peers punish for shock-induced non-repayment (see also Czura (2015b)). Alternatively, Aina et al. (2018) highlight that unfulfilled expectations about material outcomes may cause frustration, leading to punishment that is based on outcomes rather than intentions. Outcome-based punishment may explain why peers punish when they have to repay for their partner, but it would predict the same level of punishment for all types of non-repayment, irrespective of shock arrival. Explaining the punishment patterns we observe with existing theories would thus require a mixture of intention-based and outcome-based frustration. An explanation based on anger and frustration is made less likely by our use of the strategy method, which is generally understood to produce a lower bound for emotionally motivated outcomes (Aina et al. 2018; Brandts and Charness 2011).

We propose an alternative explanation in the following section: Our punishment patterns may reflect the existing social norms. Disciplining peers may be part of the norm (see pledge in Grameen Foundation (2010)). Alternatively, punishment may be driven by resentment of actions that violate norms (Kimbrough and Vostroknutov 2020, 2023). In either case, peer punishment will mirror participants' attitudes regarding socially desirable repayment behavior, even when it has no deterrent effect. If flexibility increases the uncertainty about socially desirable behavior, social norms can explain both lower punishments and lower repayment rates.

5 Evidence for Social Norms

In this section, we present several pieces of evidence to support social norms as a consistent explanation of our repayment and punishment results, assuming that participants apply their existing borrowing norms in the lab. In doing so, we follow research showing that lab behavior captures real-life norms (e.g., Huang and Low 2017), especially with framed instructions (e.g., Chang et al. 2019). We start with a brief general discussion of social norms in microfinance in Section 5.1. Section 5.2 proceeds with a simple theoretical framework of loan repayment in the presence of an exogenous social norm. We derive theoretical predictions, and use these in Section 5.3 to re-examine our empirical findings. Section 5.4 reports the results from an incentivized norm elicitation following the methodology of Krupka and Weber (2013).

5.1 Social Norms and Microfinance

We hypothesize that lender-induced social norms may be an important missing puzzle piece in understanding the existing evidence on microfinance repayment. Many microfinance institutions, including our partner organization, put great emphasis on shaping the picture of what constitutes a good borrower.[29] A prominent illustration is that borrowers recite a pledge at the beginning of every meeting to pay all weekly installments, support each other, and help to maintain discipline within the group (Grameen Foundation (2010); also see the weekly joint oath discussed in Breza (2014)). As in the Grameen pledge quoted at the beginning of this paper, this is true even when explicit joint liability has been abandoned in favor of individual liability contracts. Existing evidence linking borrower's repayment choices to social norms is largely qualitative.[30] Further suggestive evidence comes from substantial default and delinquency rates in mobile lending, which lacks the personal interactions that may be required to instill social norms (Fiorin et al. 2023; Kaffenberger et al. 2018).

The existence of social norms may reconcile several puzzles observed in microfinance research. First, empirical studies find no repayment differences between individual liability and joint liability (Attanasio et al. 2015; Giné and Karlan 2014), and speculate that social image concerns are sufficient to maintain the consistently high observed repayment rates.[31] Social image concerns may directly emerge from norms for good borrower behavior. Second, the reputation of microfinance group lending has long been tarnished with reports of excessive pressure and monitoring (Karim 2008; Montgomery 1996; Rahman 1999), culminating in the borrower suicides which led to the 2010 Andhra Pradesh microfinance crisis (studied e.g. in Breza and Kinnan (2018)). Czura (2015b) quantifies peer punishment in a lab-in-the-field experiment with microcredit borrowers in rural India. She confirms that borrowers punish excessively relative to both game-theoretical and fairness-related benchmarks, and speculates that borrowers have internalized the mission indoctrination of the

[29] Such behavior is in line with predictions from a principal-agent model, in which the principal can instill a sense of organizational identity in the agent (G. A. Akerlof and Kranton 2005).

[30] For example, repayment in Morocco is low when microfinance institutions are perceived as illegitimate or loans are perceived as development aid (Morvant-Roux et al. 2014). Osmani (2016) claims that strict rules helped establish a social norm for repayment in Bangladesh.

[31] Importantly, both studies hold weekly group meetings constant across liability type: Giné and Karlan (2014) have meetings in both types, and Attanasio et al. (2015) in neither. De Quidt et al. (2016) and Feigenberg et al. (2013) argue that group meetings create social capital, which may then be used to generate implicit joint liability (side contracting) even in individual liability contracts. However, this continues to predict differential repayment patterns across liability structures. We rule out side contracting in our experiment, and propose instead that group meetings may help lenders to publicly instill social norms.

microlender. Lender-induced social norms can explain excessive punishment if disciplining peers is part of the norm, or if punishment is driven by resentment of norm violations. Finally, a social norm that induces borrowers to make each weekly (or monthly) installment may explain why the introduction of repayment flexibility reduces repayment rates. Having discretion on whether to repay at a given moment or not creates uncertainty in the socially prescribed behavior.[32] This may offer borrowers a way to dodge some of the punishment usually associated with strategic default.

5.2 Theory: Loan Repayment with Social Norms

The following section presents a simple model of loan repayment that is consistent with our empirical findings. To illustrate the basic mechanism, we focus on the case of individual liability, and impose an exogenous social norm on repayment. We show that social norms and dynamic incentives act as substitutes in encouraging high repayment rates. As in the experiment, we then introduce flexibility as the option to postpone individual repayments. We observe that flexibility unambiguously leads to higher repayment rates absent social norms. However, when repayment is sustained by social norms, introducing flexibility can lead to the erosion of these norms, and increase default rates.

As discussed previously, the motivation for repayment flexibility is to allow borrowers to condition repayment timing on shock realizations. However, it creates a trade-off as flexibility can be misused to increase early consumption. The simplest possible model which captures the trade-off between consumption and insurance has three repayment periods. We thus model a repayment game where an agent invests a loan into a risky project, which requires a repayment R in periods $t=1,2,3$. Repaying the loan in full yields a continuation value V in $T=n$, which represents dynamic incentives such as access to future loans. The project generates a risky income of $y_t=2R$ with probability $1-\theta$, and $y_t=0$ with probability θ. There are no savings. We diverge from our experimental design in assuming that there are no shocks in the last period, i.e., that $y_3=2R$ with certainty. This eases the tractability of our model without affecting the game dynamics.[33] To show the simplest possible case, we assume risk-neutral borrowers who discount exponentially over time. Lifetime utility is:

$$U=c_1+\delta c_2+\delta^2 c_3+\delta^n V. \tag{5}$$

We now introduce a social norm for good borrowing behavior, which asks clients to faithfully repay

[32] This uncertainty is not simply due to introducing a new repayment scheme, and may not necessarily resolve as social norms adapt over time. We discuss this further in the conclusion.

[33] Period 3 is used as a catch-up period to repay postponed installments from period 2. In our experiment, we allowed for shocks in period 3. Since these were not insurable and triggered contract default, this assumption impacts the probability of obtaining the continuation value V, but not borrowers' relative incentives across the treatments.

when each installment is due. As a consequence, clients suffer a psychological cost κ each time they fail to make a scheduled repayment, including in the case of income shocks.[34] Assume $0<\kappa<R$ to avoid that repayment becomes trivial. One possible way to microfound such a norm is using models of organizational identity (G. A. Akerlof and Kranton 2005): Borrowers derive utility from their social category – here, being a diligent borrower. However, they suffer a disutility from diverging from the ideal behavior for their category, due to cognitive dissonance, guilt, shame, or reputation loss.[35]

Benchmark: Rigid repayment and social norms As in our empirical analysis, we focus on strategic repayment choices, and thus on the game path where shocks are possible ex ante, but do not materialize ex post. When the borrower makes the first decision in period 1, she already knows there is no shock in period 1. If there is a shock, the loan installment cannot be paid, and the borrower is in default. The assumption $\kappa<R$ ensures that it is not optimal to make repayments after a default.

Absent shocks in period 1, the borrower decides to repay R (and consume $c_1=y_1-R=R$), or to default. Repayment yields

$$U_1^R= \ R \ + \ (1-\theta)(\delta R+\delta^2 R+\delta^n V) \ + \ \theta(-\delta\kappa+\delta^2(2R-\kappa)). \tag{6}$$

Defaulting yields

$$U_1^D= \ 2R-\kappa \ + \ (1-\theta)(\delta(2R-\kappa)+\delta^2(2R-\kappa)) \ + \ \theta(-\delta\kappa+\delta^2(2R-\kappa)). \tag{7}$$

The repayment condition without flexibility is thus

$$\delta^{n-1}V \geq (R-\kappa)[\frac{1}{(1-\theta)\delta}+1+\delta]. \tag{8}$$

For a given level of patience and income uncertainty, the borrower repays for sufficiently high levels of dynamic incentives, or sufficiently strong social norms.

Flexible repayment and social norms We now introduce a pass token, which allows the borrower to postpone a current repayment obligation to the next period. The pass token can be used in periods 1

[34] It is possible to condition κ on whether the borrower fails to repay due to moral hazard (κ_M), or due to income shocks (κ_S). The simplifying assumption $\kappa_S=\kappa_M$ increases the insurance value of flexibility, but does not qualitatively change our results.

[35] Utility from identity is typically modeled as $I_s - e - \ t_s|e^*(s) - e|$, where I_s is the identity utility derived from social category s, e is the cost of a financial effort, and $t_s|e^*(s)-e|$ is the disutility from diverging from the ideal effort or behavior for category s. In the short run, borrowers choose their behavior e but not their social category s (Kranton 2016). Since I_s drops out as a constant, our model is nested for $t_s = \kappa/R$ and $e^*(s) = R$. In a richer model, lenders can invest in strengthening identity I_s and disutility t_s.

or 2, with or without shocks. It is tempting for the borrower to use flexibility in period 1: Immediate consumption increases by R, while the repayment of flexibility is discounted to $\delta(1-\theta)R$. But there is a trade-off: Using the pass token in period 1 means it cannot be used to insure shocks in period 2.

Assume that there is uncertainty regarding the social norm for flexibility. The social norm compels the borrower to make a repayment when asked, but now she is given discretion whether or not to repay at a given point in time. As a result, the social norm is either weakened or uncertain. The psychological cost for not repaying (while invoking flexibility) becomes $\lambda\kappa$, with $0<\lambda<1$ representing alternatively a scale parameter, or a probability that the cost κ will be incurred. Since the social norm imposes a penalty for not repaying when asked, we assume that the penalty for defaulting on the subsequent double repayment is still κ. We present empirical support for this assumption in Section 5.4.

Using flexibility is always dominant in the case of shocks. Furthermore, straight default is now dominated by using flexibility at first, and then defaulting. This is because the penalty for invoking flexibility $\lambda\kappa$ is weaker than that for simple non-repayment, κ. Focusing on choices when no shocks arrive, and starting in period 1, borrowers are left with four strategies:

1. *Flex-S*: Use flexibility only if shocks arrive, repay in periods 1, 2, and 3.
2. *Flex-2*: Use flexibility in period 2 (shocks in period 2 are insured), repay in periods 1 and 3.
3. *Flex-1*: Use flexibility in period 1 (and forfeit shock insurance in period 2),
   repay in periods 2 and 3.
4. *Flex-D*: Use flexibility in period 1, then default.

Characterizing the equilibrium behavior for the full parameter space is complicated, since no strategy is dominated. We summarize key patterns here, and defer to Appendix B for full derivations. We restrict our attention to cases where repayment is sustained by the social norm. In other words, the dynamic incentive V is sufficiently low that borrowers default at $\kappa=0$.

A first insight is that the four strategies can be ordered by their sensitivity to the penalty κ (i.e., their expected utilities can be strictly ranked by their slope in κ, see equations A1–A4): At low levels of κ, default (Flex-D) is the most attractive. However, the appeal of this strategy quickly decreases as κ increases (equation A4 steeply decreases in κ). The appeal of flexibility use in period 1 (Flex-1) is the second-most sensitive to κ, followed by Flex-2, and finally by Flex-S.

For sufficiently low λ – in other words, if flexibility creates sufficient uncertainty in the socially prescribed behavior – the result is a profile where the borrower moves from Flex-D to Flex-1, then Flex-2, and finally Flex-S as κ increases from 0 to R.[36] Flexibility is misused (strategies Flex-1 and

[36] Specifically, this holds for $\lambda \leq \bar{\lambda} \equiv \frac{(1-\delta^2)}{(2+\delta)-\delta^{n-1}\frac{V}{R}}$. For higher values of λ, the borrower moves from Flex-D (for low κ)

Flex-2) for intermediate values of κ (equation A8). For low values of κ, the borrower defaults. For high κ, flexibility is only used for shocks. The repayment condition with flexibility becomes

$$\delta^{n-1}V \geq R[2+\delta]-\kappa[1+\delta]. \tag{9}$$

Inequality 9 can be compared to the repayment condition without flexibility (inequality 8). The repayment condition with flexibility is stronger (holds less often) whenever

$$\kappa \geq (1-(1-\theta)\delta)R. \tag{10}$$

Default rates will be higher under flexible repayment contracts for large κ, large δ, and small θ.

The comparison of repayment conditions across treatments is illustrated in Figure 6. For a given κ, each curve states the minimum value of δ for which the respective strategy is preferred to default. The figure shows that at $\kappa=0$, repayment is more incentive-compatible with flexibility: Being able to delay an installment allows the agent to insure against shocks, and thus increase the probability that V can be obtained. As κ increases, repayment becomes incentive-compatible for lower δ across all conditions, since agents increasingly repay to avoid norm penalties, rather than to obtain V. However, this shift in the individuals' objective means that there are relatively more defaults in the flexible condition: Flexibility erodes the social norm of repayment, by reducing the penalty for the first non-repayment (the use of the pass token) to $\lambda\kappa$. Relatively speaking, default is more costly in the benchmark condition of rigid repayment, where κ is incurred for each missed installment. The model leads to the following testable predictions (see Appendix B.2 for details):

Prediction 1. [Overall repayment] *In the presence of strong social norms, repayment is higher under a rigid repayment contract than under a flexible repayment contract.*

Corollary 1. *Absent social norms, repayment flexibility leads to strictly higher repayment rates.*

Prediction 2. [Default path] *Using flexibility at first and then defaulting strictly dominates defaulting straight away.*

Prediction 3. [Flexibility misuse] *The insurance value of flexibility decreases over time. Thus, misuse of flexibility (use of flexibility absent shocks) will increase over time, conditional on flexibility still being available.*

directly to Flex-S (for higher κ), without misusing flexibility for early consumption. This is inconsistent with our data, which shows that 87 percent of participants in *IL-flex* misuse flexibility absent shocks. We thus focus on the case where $\lambda \leq \bar{\lambda}$ holds.

Prediction 4. [Partial repayments] *In rigid repayment contracts, partial payments are always dominated by full repayment or full default. With repayment flexibility, partial repayments can be optimal: Borrowers may comply with single but not double installments.*

5.3 Loan Repayment Paths

We can now re-examine our findings in light of a repayment model which features installment-based social norms. While the model in Section 5.2 is based on individual liability, the basic mechanisms of repayment under social norms apply independent of liability structure. We thus report whether our results for both IL- and JL-treatments are intuitively consistent with the model predictions. Our main result on repayment is in line with Prediction 1: We observe that repayment is higher under rigid than under flexible repayment conditions. To gain more detailed insights, we create repayment profiles which classify participants with respect to their behavior across all three repayment installments. As for overall repayment, all profiles refer to behavior on the no-shock path, and thus focus on ex-post moral hazard. Without flexibility, participants can *fully repay* all three tokens, they can *fully default* by not repaying any token, or they can *partially repay* by paying only one or two tokens. In the flexibility treatments, we distinguish for each profile whether flexibility has been used. If a participant chooses to repay all three tokens and uses flexibility in either period 1 or 2, she will be classified as *fully repay (with flex use)*. If she chooses to default on all three repayment installments, but uses flexibility to postpone repayment first, she will be classified as *fully default (with flex use)*. If a participant repays one or two tokens and has used flexibility, her profile is *partially repay (with flex use)*.

Figure 7 displays the repayment profiles for each treatment. The green bars indicate full repayment and correspond to the bars in Figure 2. Between *IL* and *JL* the distribution of repayment profiles is nearly identical: 66 percent fully repay all three tokens, 15 percent fully default on all three installments, and 18 to 19 percent partially repay one or two installments. The lack of differences between *IL* and *JL* is not only in line with other empirical studies (Giné and Karlan 2014), but also with the interpretation that borrowers bring their existing norms into the lab, which influences lab repayment choices in *IL* and *JL* alike. Between *IL-flex* and *JL-flex* the distribution of repayment profiles also looks very similar. In *IL-flex* (*JL-flex*), 51.3 (50.0) percent fully repay all three tokens, with 45.4 (40.2) percent using flexibility to do so. Finally, 15.6 (13.4) percent fully default on all three installments, and 33.1 (36.6) percent partially repay one or two installments.

What do we learn from these results in view of our model? In addition to Prediction 1, our data are consistent with Prediction 2: Using flexibility to postpone repayments before defaulting on them appears to dominate defaulting straight away – especially in individual-liability contracts, where there is no crowd-out with mutual insurance. While the overall share of full default is not affected by flexible repayment conditions, default with flexibility use largely replaces straight default in both *IL-flex* (85 percent of defaults), and *JL-flex* (62 percent). Prediction 3 refers to the timing of flexibility use. Consistent with the prediction, flexibility misuse increases over time, conditional on being available: In *IL-flex* (*JL-flex*), flexibility use increases from 55 percent (29 percent) in period 1 to 72 percent (52 percent) in period 2, conditional on not having been used previously (see also Table A.3).[37]

Our results are more puzzling in light of Prediction 4: We find substantial amounts of partial repayments in all repayment conditions. Moreover, the drop in overall repayment in the flexibility treatments is exclusively driven by an increase in partial repayments. The model predicts that partial repayments are dominated in the *IL* benchmark condition with rigid repayment (though not in *JL*, given the possibility of free-riding). With flexible repayment conditions, strong social norms ($R \leq \kappa < 2R$) may make it optimal to repay single but not double installments. However, when we examine partial repayments with flexibility use, over half of participants repay the double installment and default on the remaining single installment (see Table A.7 for key statistics on partial repayments).

Why do participants partially repay their loan, and why does partial repayment increase with flexibility? Three explanations come to mind. First, participants might be confused, and this confusion increases with flexibility. Several arguments speak against this explanation: All main findings are robust to the inclusion of participants who failed to pass the comprehension test. Furthermore, we predict partial repayments with treatment indicators, and control for a treatment-specific comprehension score (Table A.8). We find that high comprehension correlates negatively with partial repayments at baseline, but does not significantly interact with flexibility. We discuss confusion-based robustness tests in more detail in Appendix C.

Second, an installment-based social norm may interact with the induced artificial discounting of consumption tokens (Section 2.1): The cost of repayment is discounted (40, 30, and 20 pesos in periods 1, 2, and 3). Since periods were close together in real time, a psychological cost κ may be constant across periods. For κ in the relevant range, it may therefore be optimal to default in period 1 and repay in period 3. The data reveal the opposite: 70 (65) percent of partial repayers in *IL* (*JL*) pay the first installment, while only 38 (41) percent pay the second (Table A.7). We infer that borrowers are unlikely to have internalized this feature.

Anecdotal evidence from post-session conversations with participants suggests a third explanation: Flexible repayment conditions generated uncertainty in socially appropriate behavior,

[37] Due to the selection problem in conditioning on an endogenous variable, statistics for flexibility use in period 2 should be interpreted with caution. Summing across periods, 88 percent of participants in *IL-flex* and 66 percent of participants in *JL-flex* misuse flexibility when no shocks arrive.

which gave participants room for strategic misinterpretation.[38] Flexibility effectively excused non-repayment of an installment in one period, albeit with strict conditions on repayment and the number of times it can be used. In line with a large literature on motivated beliefs and biased information processing (see Gino et al. 2016 for a review), as well as on moral wriggle room (Dana et al. 2007), borrowers may have convinced themselves that the occasional missing of installments is approved by the lender, even beyond the use of the pass token.[39] Such borrowers may fully repay under rigid repayment conditions, where social norms provide no slack for missing installments. Flexible conditions may provide a welcome excuse to move to partial repayment, while maintaining the image of a good borrower to both oneself and to the lender.

5.4 Norm Elicitation Experiment

The last section discussed the empirical support for social norms in our experimental repayment data. We hypothesize that our participants bring these social norms from their real-life borrowing context to our experiment (see Section 2.4 for a summary of the lender's efforts to instill a culture of good borrowing).[40] To measure prevailing social norms in microcredit directly, we collected additional data in April 2019 on borrowers of the same lender in eight centers in Laguna province.

First, we present suggestive evidence on repayment norms from short surveys administered to 23 clients (see Table A.9). All but one borrower agree or strongly agree with statements that repaying is the moral thing to do, that they have learned this in their initial group training, and that the loan officer highlights the importance of repaying each week. These arguably lender-induced views are further reinforced in many borrowing groups: Two-thirds agree that the undesirability of non-repayment is discussed between group members.

Second, we use state-of-the-art experimental methods to elicit existing norms in an incentivized manner. Consistent with Krupka and Weber (2013), we elicit norms in a separate sample drawn from the same borrower population.[41] In eight sessions, a total of 44 clients evaluate different

[38] The idea that uncertainty in norms may lead people to strategically distort their beliefs was recently introduced by Bicchieri et al. (2020) in the context of a lab experiment on lying.

[39] Our punishment design imposes a constant linear penalty for each missed installment. As a result, it would not capture an alleged acceptability of occasionally missing installments when repayment is flexible.

[40] We also see some indications of this in administrative records. Data from the lender are consistent with heterogeneity in how much borrowers have internalized social norms on repayment, and norms becoming more powerful over time: We find positive associations of full repayment in the experiment with a measure of borrower quality, namely always repaying loan installments on time ($p = 0.10$). Repayment is lower for those who became clients within the last 12 months (see Table A.10), suggesting that norms may take effect over time. Furthermore, new clients are significantly less likely to use flexibility responsibly (i.e., only to insure shocks).

[41] This prevents concerns about self-serving answers as participants do not face the incentives from the experiment. Erkut et al. (2015) show elicited norms are similar between stakeholders and spectators.

vignettes that closely mirror the scenarios in our experiment. Borrowers rate the social appropriateness of each possible repayment choice. They are incentivized not to reveal their own valuations, but to match those of others. Krupka and Weber (2013) show that norms emerge as the focal point in a matching coordination game.

Vignettes describe the loan repayment behavior of a fictitious client, Maria, in a nearby joint-liability loan center (see Appendix H). As in the experiment, vignettes refer to single-period actions under either rigid or flexible repayment conditions, and build in observable income shocks.[42] Participants rated the social appropriateness of each vignette on a six-point Likert-scale from 1 'very socially inappropriate' to 6 'very socially appropriate', using different smileys to illustrate the options. Participants received a bonus payment if their rating of a randomly drawn vignette matched the rating of another randomly drawn participant:[43] A bonus of 50 pesos was paid for an exact match, and 20 pesos for a one-point deviation. All participants received a participation payment of 50 pesos. Analogous to the experiment, the order of the vignettes was kept constant to ease the exposition (first rigid, then flexible repayment conditions), and all participants rated all vignettes.

These data allow us to answer the following questions: (i) Is there a social norm for repayment? (ii) If yes, do norms mirror the punishment patterns we observe in our experiment? (iii) Does the applicability of these norms become more uncertain in the presence of flexibility? Table 2 sets out the results. On (i), we find strong evidence for the existence of social norms for repayment: 73 percent of participants rate repayment as 'very appropriate' and 25 percent as 'mostly appropriate,' suggesting a strong focal point in the coordination game. Perhaps more surprisingly, 84 percent rate non-repayment following a shock as either 'very inappropriate' or 'mostly inappropriate', suggesting that social norms do not excuse non-repayment even when it is unavoidable. On (ii), we find suggestive evidence that the punishment we observe reflects the underlying social norms: The ranking of actions by appropriateness is the same in the norm elicitation as in the punishment choices (see Table 2). Non-repayment due to a shock is rated nearly as inappropriate as strategic default – an even more extreme result than for punishment, potentially due to the fact that punishment was intrinsically costly while appropriateness ratings were incentivized on coordination. A further parallel is that norms clearly favor using flexibility to self-insure against shocks, rather than to rely on peers. This provides useful insights on what may drive punishment behavior in the experiment: Section 4.4 discusses frustration as a possible driver of punishment, as participants were directly affected by their peer's repayment

[42]To approach a more realistic borrowing scenario, we ask about repayment choices in week 3 (and in some cases, week 4) of a 25-week repayment cycle, when the first two weeks were repaid. Empirically, default on the first repayment installments is virtually non-existent. We avoid moving later into the repayment cycle and specifying the full previous repayment history, as this may be misperceived as a signal about the borrower's type or intentions.

[43]This is strategically equivalent to matching on the modal response, but was easier to understand for our participants.

decisions. In contrast, norm study participants are unaffected observers. The fact that we continue to see this pattern is more consistent with norms as a direct driver of experimental punishment.

Does the applicability of norms become uncertain in the presence of flexibility? The most direct evidence for question (iii) is the dispersion of appropriateness ratings: The more participants struggle to coordinate on the same rating, the more uncertainty there is in what constitutes socially desirable behavior. Table 2 shows that the modal rating for strategic default (very inappropriate) is chosen by 57 percent of participants, compared to 39 percent of participants who choose the modal rating (mostly inappropriate) for flexibility misuse, despite the fact that both equate to the non-repayment of an installment. Moreover, all six rating options are chosen by at least 5 percent of participants for flexibility misuse, with 26 percent giving a *positive* rating (for strategic default: 4 percent). The dispersion of ratings becomes even larger for flexibility use in case of shocks: Ratings are distributed nearly symmetrically, with 39 percent rating flexibility use as 'mostly appropriate' and 30 percent rating it as 'mostly inappropriate'. Our results suggest substantial uncertainty in how to apply existing repayment norms to flexible repayment conditions.

An additional way to test question (iii) comes from Prediction 2 as well as our repayment results (Figure 7): Both in theory and empirics, using flexibility first and then defaulting largely dominates defaulting straight away. In Appendix E, we show that the expected punishment for strategic default is lower when repayments are postponed by first using flexibility (Figure 5). We observe a similar pattern in social appropriateness rankings, in a sample disconnected from our experiment: In addition to the vignettes about repayment behavior in a given week, we added selected vignettes about repayment in two consecutive weeks (see Appendix H). Averaging the appropriateness rating of two consecutive non-repayments yields 0.88 (SD 0.17). In contrast, misusing flexibility and then defaulting on the double installment appears to be less inappropriate (average 0.79 (SD 0.20), Wilcoxon signed-rank test p-value<0.001). This result is driven by an increased relative appropriateness of misusing flexibility, while defaulting on a double installment is considered as inappropriate as a second single-installment default (Table 2). This is consistent with norms centered on repaying when due, rather than on the repayment amount. Figure A.2 shows the distribution of the combined ratings and confirms an increased dispersion with flexibility, suggesting higher uncertainty.

If flexibility creates uncertainty in socially appropriate behavior, can our norm elicitation explain the observed frequency of partial repayments? Section 5.3 speculated that participants may have interpreted flexibility as a signal that occasional non-repayment is acceptable, in line with studies on moral wriggle room and motivated reasoning (Gino et al. 2016). Unfortunately, this mechanism would not show up in our norm elicitation: Motivated reasoning needs a motivation. In contrast to our experimental participants, norm elicitation respondents are not directly affected by their interpretation of the repayment conditions, and thus have no incentive to strategically misinterpret them.

6 Conclusion

We study repayment choices under both rigid and flexible repayment conditions in a lab-in-the-field experiment with microcredit borrowers in the Philippines. Although repayment is not payoff-maximizing in our setting, we find high repayment rates across both individual- and joint-liability contracts. The introduction of flexibility increases strategic defaults on the overall loan by 50 percent. Flexibility also reduces peer punishment in joint-liability contracts – both when it is used to insure income shocks, and when it is used to increase early consumption absent shocks.

Our results are consistent with a strong social norm on repayment, which participants bring to our sessions from their real-life borrowing experience. Through meeting and reciting pledges every week, clients internalize what it means to be a good borrower: to pay installments every week, and to discipline peers. We draw parallels to existing work on identity in organizations (G. A. Akerlof and Kranton 2005), and point out that microfinance organizations have strong incentives to inculcate a sense of identity in their borrowers, with associated norms for good behavior. Such norms could help explain not only the high repayment rates and punishment patterns in our experiment, but also two recent puzzles in microfinance research: First, why repayment rates do not differ between individual- and joint-liability contracts. Second, why peer pressure appears to be excessive and sequentially irrational. Furthermore, if social norms refer to weekly installments, the discretion introduced by repayment flexibility means that applying the norm may no longer be straightforward. In turn, uncertainty in socially prescribed behavior may increase ex-post moral hazard. We present supporting evidence for this explanation using a theoretical framework, and from the first incentivized norm elicitation study in microfinance.

Our results also broaden the recent discussion on flexible repayment in microfinance. Existing evidence suggests that flexible repayment can increase profits by facilitating high-risk, high-return investment (Aragón et al. 2020; Barboni and Agarwal 2021; Battaglia et al. 2021; Field et al. 2013). Our results reveal an additional benefit of flexible repayment schemes: they may reduce excessive social pressure in group lending (documented in Czura (2015b), Karim (2008), Rahman (1999), and Montgomery (1996)), by providing borrowers with a way to self-insure against income fluctuations. However, our results also suggest that flexibility may destabilize high-repayment equilibria which are sustained by social norms. We hypothesize that motivated beliefs may act to exacerbate the consequences of norm uncertainty. This is particularly important in settings with observable actions, as small changes in feedback can lead to unraveling of a norm (Bursztyn et al. 2020b; Hill et al. 2012).

Several caveats apply. First, our experiment newly introduces flexible repayment terms. We cannot speak directly to whether and how norms would adjust to flexibility over time, and whether increased strategic defaults would persist. Having said that, the nature of flexibility is to give the borrower discretion in whether to repay or not at a given moment. This may create uncertainty in socially appropriate behavior which does not simply resolve over time: a norm for flexibility would have to distinguish between responsible and irresponsible use. Lenders may be constrained in what kind of norms may be induced: A major benefit of rigid, no-exceptions rules on weekly repayments is that they are *simple*, and can easily be integrated into the pledge that borrowers recite at the start of every meeting.[44] In contrast, inducing a norm on exactly *when* flexibility use is acceptable may be more complicated, especially since repayment capacity is hard to quantify in practice.

Furthermore, we study a particular type of flexibility – discretion in the timing of repayment (also studied in Brune et al. (2022), Battaglia et al. (2021), and Czura (2015a)). Different flexibility designs may have different implications for norm uncertainty. In particular, neither the two-months grace period in Field et al. (2013) nor the (pre-planned) repayment holidays in Barboni and Agarwal (2021) give borrowers any discretion in whether to repay at a given point in time. While these flexibility designs make socially appropriate behavior more straightforward, they do not provide insurance against shocks. This is important in light of Battaglia et al. (2021), who show that increased profits from flexibility appear to be driven by insurance provision rather than by the easing of credit constraints. Our results point to a fundamental trade-off in the design of repayment: Giving borrowers the ability to condition repayments on unobservable (or uncontractable) shocks necessarily requires giving them discretion in whether to repay at a given point in time. Discretion may increase moral hazard, both through present bias (studied theoretically in G. Fischer and Ghatak (2016)), and through uncertainty in social norms.

Future research is needed on the nature of social norms in lending, on how these are formed, and how they respond to changes to contract terms. The idea that repayment equilibria are sustained by social norms may help to explain a range of phenomena, including recent evidence that debt relief programs increase moral hazard even among borrowers who were not at risk of default (Giné and Kanz 2018; Kanz 2016). We believe that models of organizational identity may provide a promising avenue for understanding such norm dynamics.

[44] This is consistent with our result that norms do not seem to allow for shock-induced non-repayment. See e.g. Yoeli et al. (2022), who show theoretically that it is easier to condition norms on categorical distinctions than on continuous variation.

References

Abbink, Klaus, Bernd Irlenbusch, and Elke Renner (2006): "Group size and social ties in microfinance institutions." *Economic Inquiry* 44 (4), pp. 614–628.

Aina, Chiara, Pierpaolo Battigalli, Astrid Gamba, et al. (2018): "Frustration and anger in the ultimatum game: An Experiment." *Working Paper*.

Akerlof, George A and Rachel E Kranton (2000): "Economics and identity." *Quarterly Journal of Economics* 115 (3), pp. 715–753.

— (2002): "Identity and schooling: Some lessons for the economics of education." *Journal of Economic Literature* 40 (4), pp. 1167–1201.

— (2005): "Identity and the Economics of Organizations." *Journal of Economic Perspectives* 19 (1), pp. 9–32.

Akerlof, Robert (2016): "Anger and enforcement." *Journal of Economic Behavior & Organization* 126, pp. 110–124.

Aleem, Irfan (1990): "Imperfect Information, Screening, and the Costs of Informal Lending: A Study of a Rural Credit Market in Pakistan." *World Bank Economic Review* 4 (3), pp. 329–349.

Aragón, Fernando, Alexander Karaivanov, and Karuna Krishnaswamy (2020): "Credit lines in micro-credit: Short-term evidence from a randomized controlled trial in India." *Journal of Development Economics* 146, p. 102497.

Armendáriz, Beatriz (1999): "On the design of a credit agreement with peer monitoring." *Journal of Development Economics* 60 (1), pp. 79–104.

Armendáriz, Beatriz and Jonathan Morduch (2010): *The economics of microfinance*. Cambridge, Massachusetts, USA: MIT press.

Athreya, Kartik, Juan M. Sánchez, Xuan S. Tam, and Eric R. Young (May 2018): "Bankruptcy and delinquency in a model of unsecured debt." *International Economic Review* 59 (2), pp. 593–623.

Attanasio, Orazio, Britta Augsburg, Ralph De Haas, Emla Fitzsimons, and Heike Harmgart (2015): "The impacts of microfinance: Evidence from joint-liability lending in Mongolia." *American Economic Journal: Applied Economics* 7 (1), pp. 90–122.

Baland, Jean-Marie, Lata Gangadharan, Pushkar Maitra, and Rohini Somanathan (2017): "Repayment and exclusion in a microfinance experiment." *Journal of Economic Behavior and Organization* 137, pp. 176–190.

Baland, Jean-Marie, Catherine Guirkinger, and Charlotte Mali (2011): "Pretending to be poor: Borrowing to escape forced solidarity in Cameroon." *Economic Development and Cultural Change* 60 (1), pp. 1–16.

Barboni, Georgia and Parul Agarwal (2021): "How do flexible microfinance contracts improve repayment rates and business outcomes? Experimental Evidence from India." *Working Paper*.

Bari, Faisal, Kashif Malik, Muhammad Meki, and Simon R. Quinn (2021): "Asset-based microfinance for microenterprises: Evidence from Pakistan." *Working Paper*.

Battaglia, Marianna, Selim Gulesci, and Andreas Madestam (2021): "Repayment flexibility and risk taking : Experimental evidence from credit contracts." *Working Paper*.

Battigalli, Pierpaolo, Martin Dufwenberg, and Alec Smith (2019): "Frustration, aggression, and anger in leader-follower games." *Games and Economic Behavior* 117, pp. 15–39.

Bauer, Michal, Julie Chytilova, and Jonathan Morduch (2012): "Behavioral foundations of Microcredit: Experimental and survey evidence from rural India." *American Economic Review* 102(2), pp. 1118–1139.

Besley, Timothy and Stephen Coate (1995): "Group lending, repayment incentives and social collateral." *Journal of Development Economics* 46 (1), pp. 1–18.

Bicchieri, Cristina (2016): *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press.

Bicchieri, Cristina, Eugen Dimant, and Silvia Sonderegger (2020): "It's not a lie if you believe the norm does not apply: Conditional norm-following with strategic beliefs." *Working Paper*.

Brandts, Jordi and Gary Charness (2011): "The strategy versus the direct-response method: A first survey of experimental comparisons." *Experimental Economics* 14 (3), pp. 375–398.

Breza, Emily (2014): "Peer effects and loan repayment: Evidence from the Krishna Default Crisis." *Working Paper*.

Breza, Emily and Cynthia Kinnan (2018): "Measuring the equilibrium impacts of credit: Evidence from the Indian Microfinance Crisis." *Working Paper*.

Brune, Lasse, Xavier Giné, and Dean Karlan (2022): "Give Me a Pass: Flexible Credit for Entrepreneurs in Colombia." *NBER Working Paper* 30634.

Bu, Di and Yin Liao (2022): "Shaming Microloan Delinquents: Evidence from a Field Experiment in China." *Management Science* 68 (5), pp. 3768–3790.

Bursztyn, Leonardo, Georgy Egorov, and Stefano Fiorin (2020a): "From extreme to mainstream: The erosion of social norms." *American Economic Review* 110 (11), pp. 3522–48.

Bursztyn, Leonardo, Alessandra L. González, and David Yanagizawa-Drott (Oct. 2020b): "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia." *American Economic Review* 110 (10), pp. 2997–3029.

Cason, Timothy N., Lata Gangadharan, and Pushkar Maitra (2012): "Moral hazard and peer monitoring in a laboratory microfinance experiment." *Journal of Economic Behavior and Organization* 82 (1), pp. 192–209.

Cassar, Alessandra, Luke Crowley, and Bruce Wydick (2007): "The effect of social capital on group loan repayment: Evidence from field experiments." *Economic Journal* 117 (517), F85–F106.

Chang, Daphne, Roy Chen, and Erin Krupka (2019): "Rhetoric matters: A social norms explanation for the anomaly of framing." *Games and Economic Behavior* 116, pp. 158–178.

Charness, Gary and David I. Levine (2007): "Intention and stochastic outcomes: An experimental study." *Economic Journal* 117 (552), pp. 1051–1072.

Coutts, Alexander (2022): "Identifying communication spillovers in lab in the field experiments." *Journal of Development Economics* 157, p. 102845.

Czura, Kristina (2015a): "Do flexible repayment schedules improve the impact of microcredit? Evidence from a randomized evaluation in rural India." *Working Paper*.

— (2015b): "Pay, peek, punish? Repayment, information acquisition and punishment in a microcredit lab-in-the-field experiment." *Journal of Development Economics* 117, pp. 119–133.

Dana, Jason, Roberto A Weber, and Jason Xi Kuang (2007): "Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness." *Economic Theory* 33 (1), pp. 67–80.

De Quidt, Jonathan, Thiemo Fetzer, and Maitreesh Ghatak (2016): "Group lending without joint liability." *Journal of Development Economics* 121, pp. 217–236.

— (2018): "Commercialization and the decline of joint liability microcredit." *Journal of Development Economics* 134, pp. 209–225.

Dhami, Sanjit, Junaid Arshad, and Ali al-Nowaihi (2022): "Psychological and social motivations in microfinance contracts: Theory and evidence." *Journal of Development Economics* 158, p. 102912.

Erkut, Hande, Daniele Nosenzo, and Martin Sefton (2015): "Identifying social norms using coordination games: Spectators vs. stakeholders." *Economics Letters* 130, pp. 28–31.

Fehr, Ernst and Simon Gächter (2000): "Cooperation and punishment in public goods experiments." *American Economic Review* 90 (4), pp. 980–994.

— (2002): "Altruistic punishment in humans." *Nature* 415, pp. 137–140.

Feigenberg, Benjamin, Erica Field, and Rohini Pande (2013): "The economic returns to social interaction: Experimental evidence from microfinance." *Review of Economic Studies* 80 (4), pp. 1459–1483.

Field, Erica and Rohini Pande (2008): "Repayment frequency and default in microfinance: Evidence from India." *Journal of the European Economic Association* 6 (2-3), pp. 501–509.

Field, Erica, Rohini Pande, John Papp, and Natalia Rigol (2013): "Does the classic microfinance model discourage entrepreneurship among the poor? Experimental evidence from India." *American Economic Review* 103 (6), pp. 2196–2226.

Fiorin, Stefano, Joseph Hall, and Martin Kanz (2023): "How Do Borrowers Respond to a Debt Moratorium? Experimental Evidence from Consumer Loans in India." *CEPR Working Paper*.

Fischer, Greg and Maitreesh Ghatak (2016): "Repayment frequency in lending contracts." *Working Paper*.

Fischer, Paul and Steven Huddart (2008): "Optimal contracting with endogenous social norms." *American Economic Review* 98 (4), pp. 1459–75.

Gertler, Paul J., Brett Green, and Catherine D. Wolfram (2021): "Digital Collateral." *Working Paper*.

Ghatak, Maitreesh and Timothy W. Guinnane (1999): "The economics of lending with joint liability: Theory and practice." *Journal of Development Economics* 60 (1), pp. 195–228.

Giné, Xavier, Jessica Goldberg, and Dean Yang (2012): "Credit Market Consequences of Improved Personal Identification: Field Experiment Evidence from Malawi." *American Economic Review* 102 (2009), pp. 2923–2954.

Giné, Xavier, Pamela Jakiela, Dean Karlan, and Jonathan Morduch (2010): "Microfinance games." *American Economic Journal: Applied Economics* 2 (July), pp. 60–95.

Giné, Xavier and Martin Kanz (2018): "The economic effects of a borrower bailout: Evidence from an emerging market." *Review of Financial Studies* 31 (5), pp. 1752–1783.

Giné, Xavier and Dean Karlan (2014): "Group versus individual liability: Short and long term evidence from Philippine microcredit lending groups." *Journal of Development Economics* 107, pp. 65–83.

Gino, Francesca, Michael I Norton, and Roberto A Weber (2016): "Motivated Bayesians: Feeling moral while acting egoistically." *Journal of Economic Perspectives* 30 (3), pp. 189–212.

Grameen Foundation (2010): *Guidelines for establishing and operating Grameen-style microcredit programs*. Tech. rep.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (Aug. 2013): "The Determinants of Attitudes toward Strategic Default on Mortgages." *Journal of Finance* 68 (4), pp. 1473–1515.

Henrich, Joseph, Jean Ensminger, Richard McElreath, Abigail Barr, and Clark Barrett (2010): "Markets, religion, community size, and the evolution of fairness and punishment." *Science* 327 (5972), pp. 1480–1484.

Henrich, Joseph, Richard McElreath, Abigail Barr, Jean Ensminger, and Clark Barrett (2006): "Costly punishment across human societies." *Science* 312 (5781), pp. 1767–1770.

Hill, Ruth Vargas, Eduardo Maruyama, and Angelino Viceisza (2012): "Breaking the norm: An empirical investigation into the unraveling of good behavior." *Journal of Development Economics* 99, pp. 150–162.

Huang, Jennie and Corinne Low (2017): "Trumping norms: Lab evidence on aggressive communication before and after the 2016 US presidential election." *American Economic Review* 107 (5),

pp. 120–124.
Huck, Steffen, Dorothea Kübler, and Jörgen Weibull (2012): "Social norms and economic incentives in firms." *Journal of Economic Behavior and Organization* 83 (2), pp. 173–185.
Iyer, Rajkamal, Asim Ijaz Khwaja, Erzo F. P. Luttmer, and Kelly Shue (2016): "Screening Peers Softly: Inferring the Quality of Small Borrowers." *Management Science* 62 (6), pp. 1554–1577.
Jackson, Matthew O. and Leeat Yariv (2014): "Present bias and collective dynamic choice in the lab." *American Economic Review* 104 (12), pp. 4184–4204.
Jakiela, Pamela and Owen Ozier (2016): "Does Africa need a rotten kin theorem? Experimental evidence from village economies." *Review of Economic Studies* 83 (1), pp. 231–268.
Kaffenberger, Michelle, Edoardo Totolo, and Matthew Soursourian (2018): "A digital credit revolution." *Working Paper*.
Kanz, Martin (2016): "What does debt relief do for development? Evidence from India's bailout for rural households." *American Economic Journal: Applied Economics* 8 (4), pp. 66–99.
Karim, Lamia (2008): "Demystifying micro-credit: The Grameen Bank, NGOs, and neoliberalism in Bangladesh." *Cultural Dynamics* 20 (1), pp. 5–29.
Karlan, Dean (2007): "Social connections and group banking." *Economic Journal* 117 (517), pp. 52–84.
Karlan, Dean and Sendhil Mullainathan (2007): "Is microfinance too rigid?" *Financial Access Initiative Concept Note*.
Karlan, Dean and Jonathan Zinman (2009): "Observing unobservables: Identifying information asymmetries with a consumer credit field experiment." *Econometrica* 77 (6), pp. 1993–2008.
Kimbrough, Erik and Alexander Vostroknutov (2020): "A theory of injunctive norms." *Working Paper*.
— (2023): "Resentment and punishment." *Working Paper*.
Kranton, Rachel E (2016): "Identity economics 2016: Where do social distinctions and norms come from?" *American Economic Review* 106 (5), pp. 405–409.
Krupka, Erin L. and Roberto A. Weber (2013): "Identifying social norms using coordination games: Why does dictator game sharing vary?" *Journal of the European Economic Association* 11 (June), pp. 495–524.
Labie, Marc, Carolina Laureti, and Ariane Szafarz (2017): "Discipline and flexibility: A behavioural perspective on microfinance product design." *Oxford Development Studies* 45 (3), pp. 321–337.
Masclet, David, Charles Noussair, Steven Tucker, and Marie-Claire Villeval (2003): "Monetary and nonmonetary punishment in the voluntary contributions mechanism." *American Economic Review* 93 (1), pp. 366–380.
Meyer, Richard L. (2002): "The demand for flexible microfinance products: Lessons from Bangladesh." *Journal of International Development* 14 (3), pp. 351–368.
Montgomery, Richard (1996): "Disciplining or protecting the poor? Avoiding the social costs of peer pressure in micro-credit schemes." *Journal of International Development* 8 (2), pp. 289–305.
Morvant-Roux, Solène, Isabelle Guérin, Marc Roesch, and Jean-Yves Moisseron (2014): "Adding value to randomization with qualitative analysis: The case of microcredit in rural Morocco." *World Development* 56, pp. 302–312.
Osmani, S.R. (2016): "Models of microcredit delivery and social norm." *Working Paper*.
Rahman, Aminur (1999): "Microcredit initiatives for equitable and sustainable development: Who pays?" *World Development* 27 (1), pp. 67–82.
Rand, David G., Drew Fudenberg, and Anna Dreber (2015): "It's the thought that counts: The role of intentions in noisy repeated games." *Journal of Economic Behavior & Organization* 116, pp. 481–499.
Rosenberg, Richard (1999): "Measuring microcredit delinquency: Ratios can be harmful to your

health." *CGAP Occasional Paper* (3).

Roth, Alvin E. (1995): "Bargaining experiments." *Handbook of Experimental Economics*. Ed. by John Kagel and Alvin E. Roth. Princeton University Press, pp. 253–348.

Selten, Reinhard (1967): "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments." *Beiträge zur Experimentellen Wirtschaftsforschung*. Ed. by Heinz Sauermann. Tübingen, Germany: J.C.B. Mohr.

Stiglitz, Joseph E. (1990): "Peer Monitoring and Credit Markets." *World Bank Economic Review* 4 (3), pp. 351–366.

Wydick, Bruce (1999): "Can social cohesion be harnessed to repair market failures? Evidence from group lending in Guatemala." *Economic Journal* 109 (457), pp. 463–475.

Yoeli, Erez, Bethany Burum, N Aygun Dalkiran, Martin Nowak, and Moshe Hoffman (2022): "The emergence of categorical norms." *Working Paper*.

Tables and Figures

Figure 1: Experimental Design (IL and IL-flex)

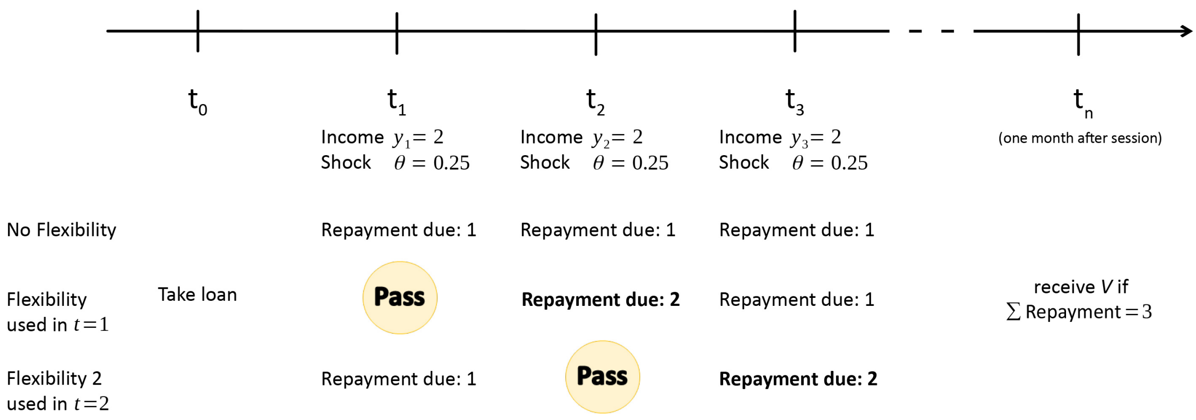
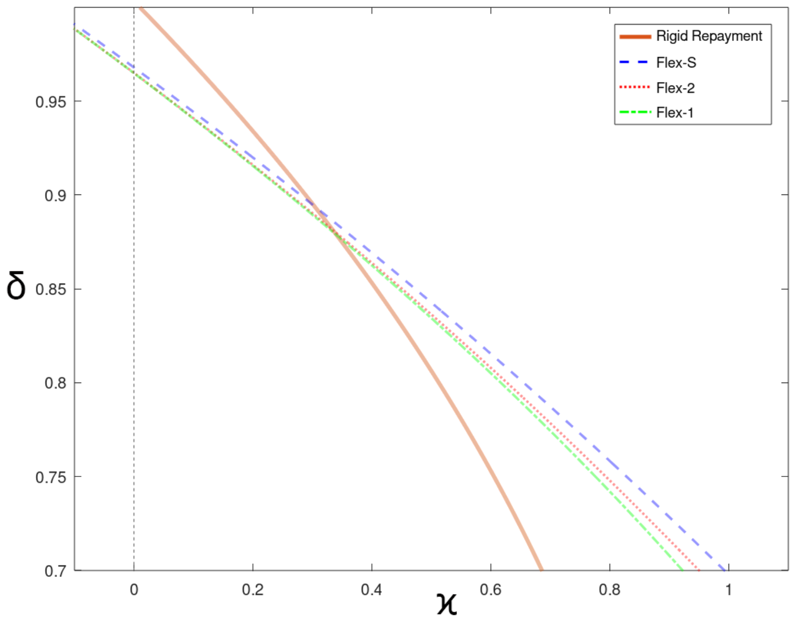
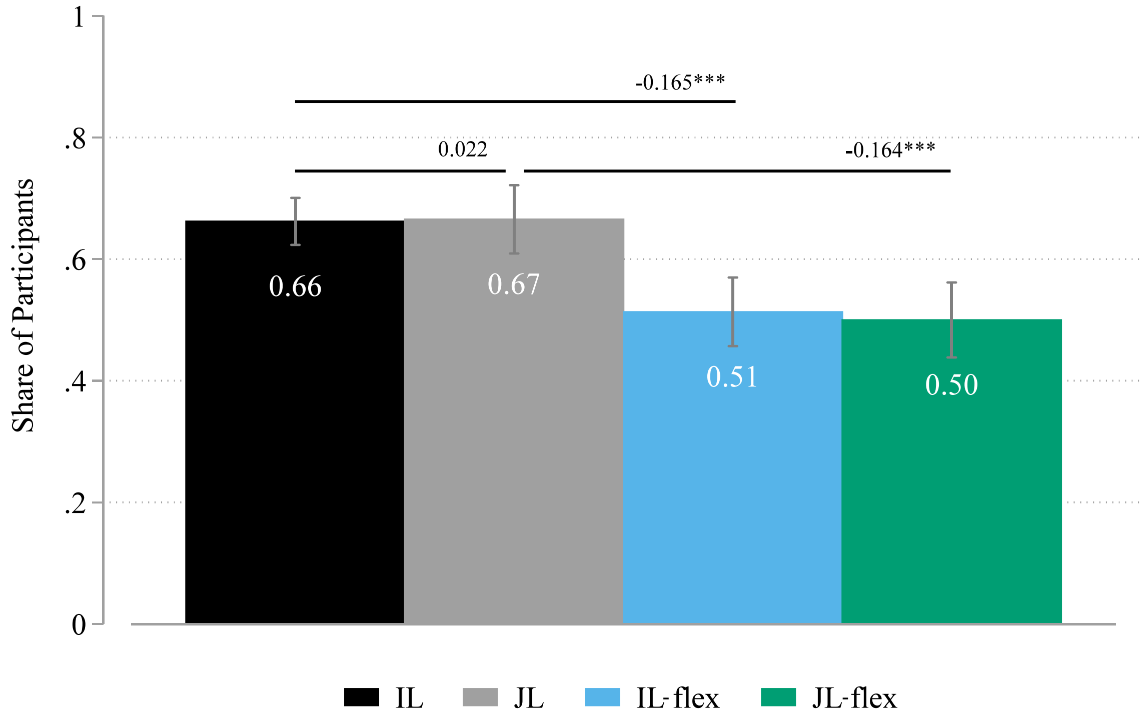


Figure 6: Repayment Conditions with Rigid and Flexible Repayment



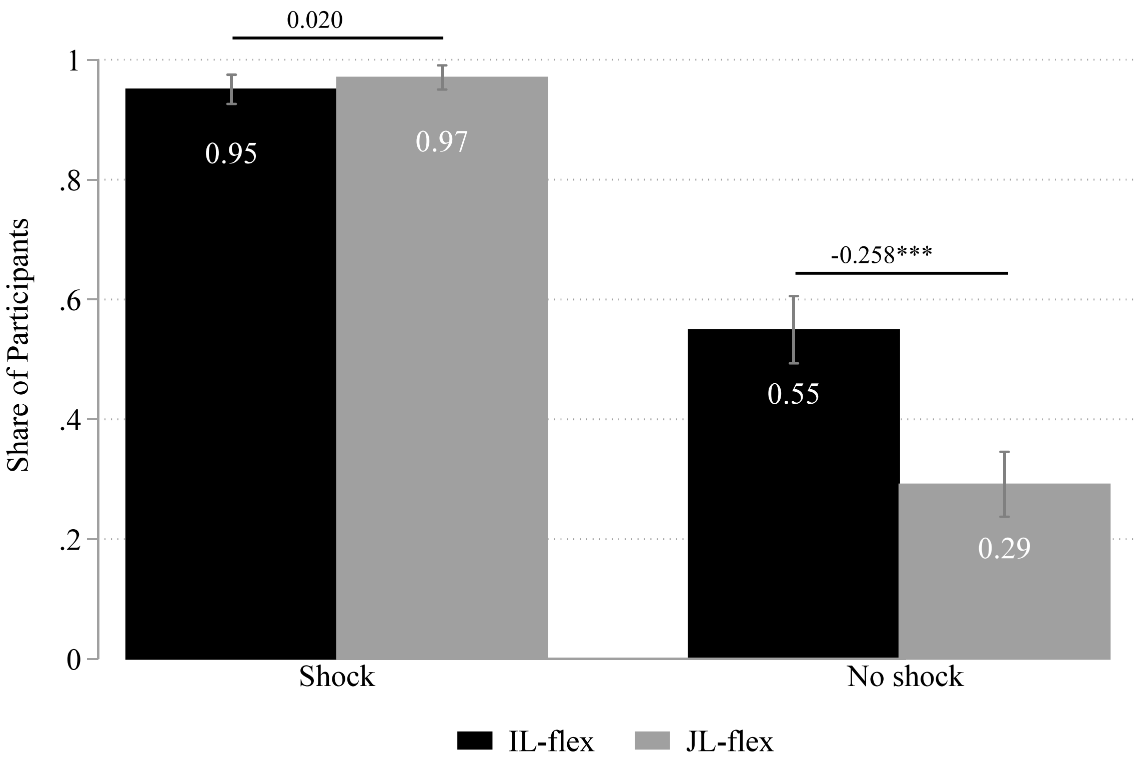
Notes: For a given κ, each curve states the minimum value of δ for which the respective strategy is preferred to default (equation 7 in rigid repayment and equation A4 in flexible repayment). For instance, the Flex-2 curve compares equations A2 and A4. The overall repayment condition under flexibility is given by the lowest curve of Flex-1, Flex-2, and Flex-S. This figure shows simulations using $V = 3.3$, $R = 1$, $n = 4$, $\theta = 0.25$, and $\lambda = 0.2$. Note the experiment induced $V/R = 3.3$, $\theta = 0.25$ and $\delta \approx 0.8$ (see Section 2.1).

Figure 2: Individual Full Repayment



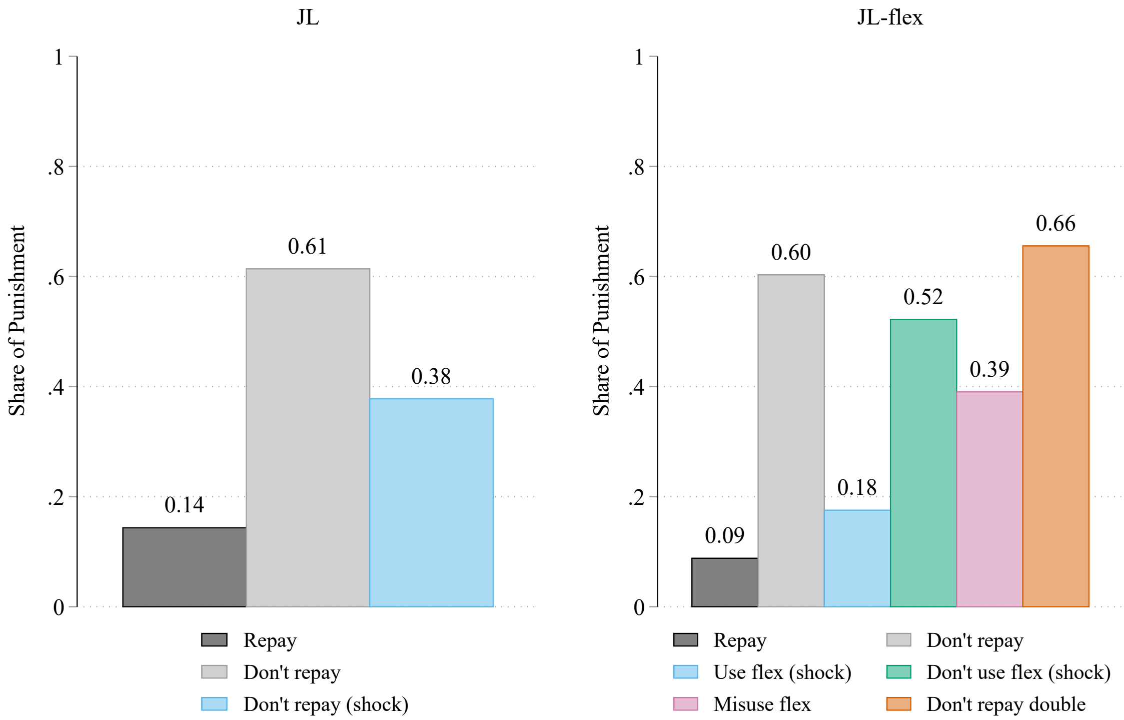
Notes: Binary indicator for full repayment. Coefficients from OLS regressions with session fixed effects and standard errors clustered at the individual level (shown in Appendix Table A.1). ***p<0.01, **p<0.05, *p<0.10.

Figure 3: Use of Flexibility



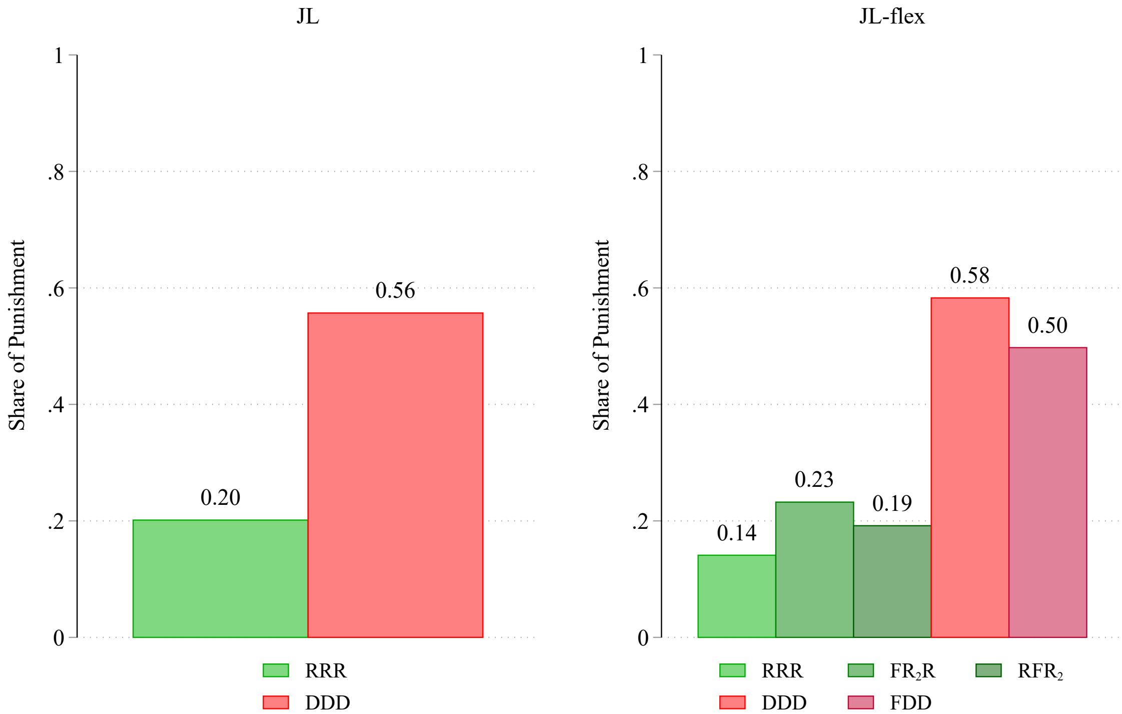
Notes: Share of participants who use flexibility in period 1. Coefficients from four OLS regressions comparing the use of flexibility in the respective scenario, with *IL-flex* as the reference category and standard errors clustered at session level (shown in Table A.3). *** p<0.01, ** p<0.05, * p<0.1.

Figure 4: Level of Punishment (single-period actions)



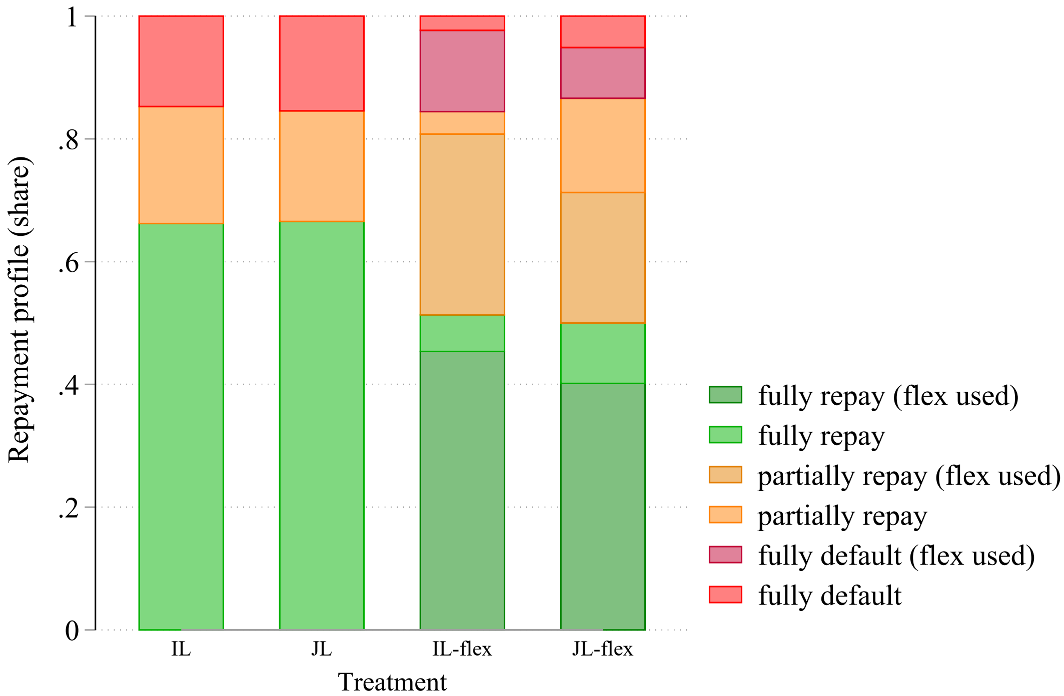
Notes: Level of punishment: share of punishment tokens relative to the maximum possible punishment (two tokens). Incidence of punishment shown in Figure A.1. Punishment choices are conditional on partner's action and shock arrival as indicated.

Figure 5: Expected Punishment



Notes: The expected punishment for a strategy is the average punishment over all one-period actions, where each action is weighted by the probability that it is played. Strategy names refer to actions on the no-shock path: *R* is repayment, *D* is default, *F* is flexibility use, and R_2 is a double repayment following flexibility use. In *JL-Flex*, all full-repayment strategies assume that flexibility is used in case of shocks (consistent with the results in Section 4.2). Strategy *DDD* assumes that flexibility is not used in any state of the world.

Figure 7: Repayment Profiles (no-shock path)



Notes: Fraction of participants who fully repay, partially repay and fully default in each treatment. For the flexibility treatments, the graph additionally indicates for each of the three scenarios whether flexibility has been used.

Table 1: Borrower Characteristics (administrative data)

Variable	Means Total (1)	IL-Session (2)	JL-Session (3)	Difference IL vs. JL (4)
Female	0.931	0.954	0.907	-0.046
	(0.254)	(0.210)	(0.290)	(0.483)
Age	46.546	46.606	46.483	-0.122
	(11.745)	(12.180)	(11.299)	(0.944)
Probability of living below NPL	45.218	47.213	43.142	-4.071
	(31.591)	(32.666)	(30.353)	(0.500)
Electricity	0.802	0.783	0.821	0.038
	(0.399)	(0.413)	(0.384)	(0.659)
Tap Water	0.230	0.181	0.278	0.097
	(0.421)	(0.386)	(0.449)	(0.321)
Landline Phone	0.022	0.028	0.016	-0.012
	(0.147)	(0.166)	(0.125)	(0.512)
Education: Secondary graduate	0.506	0.473	0.541	0.067
	(0.500)	(0.500)	(0.499)	(0.362)
Loan Amount in PHP 1000	14.350	13.765	14.963	1.198
	(11.087)	(10.156)	(11.974)	(0.438)
Main income: Enterprise	0.466	0.463	0.469	0.007
	(0.499)	(0.500)	(0.500)	(0.955)
Main income: Farming	0.261	0.242	0.278	0.035
	(0.439)	(0.429)	(0.449)	(0.806)
Iron Roof	0.754	0.715	0.794	0.079
	(0.431)	(0.452)	(0.405)	(0.363)
IL-loan center	0.179	0.174	0.184	0.010
	(0.383)	(0.380)	(0.388)	(0.942)
Observations	577	305	272	577

Notes: The table presents means and standard deviations in parentheses for administrative variables. Column (2) shows data for participants in IL-Sessions (who play IL and IL-Flex). Column (3) shows data for participants in JL-Sessions (who play IL as a benchmark, then JL and JL-Flex). See Section 2.2 for details on procedures. Column (4) reports differences and p-values in parentheses from regressions with standard errors clustered at the session level. NPL refers to the national poverty line. All variables except age, probability of living below NPL, and loan amount are binary. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Norm Results

Panel A	**Punishment vs. Norms**								
	Level of Punishment in Experiment (0-1)	Norm Elicitation: Inappropriateness (0-1)							
				Percent of Respondents					
Action	Mean	Mean	SD	+++	++	+	-	- -	- - -
No flexibility									
Repay	0.14	0.06	0.12	**73**	25	0	2	0	0
Don't repay (shock)	0.38	0.85	0.15	0	0	2	14	**43**	41
Don't repay	0.61	0.86	0.19	0	2	2	14	25	**57**
Flexibility									
Repay	0.09	0.23	0.27	**48**	14	20	11	7	0
Use flex (shock)	0.18	0.47	0.3	5	**39**	11	11	30	5
Misuse flex	0.39	0.69	0.28	5	7	14	14	**39**	23
Don't repay (shock)	0.52	0.77	0.28	7	2	2	18	30	**41**
Don't repay	0.6	0.9	0.16	0	0	5	5	30	**61**
Action in subsequent period									
No flex: Don't repay, second time		0.89	0.21	0	7	0	2	23	**68**
Flex: Don't repay double	0.66	0.9	0.18	0	2	2	5	25	**66**

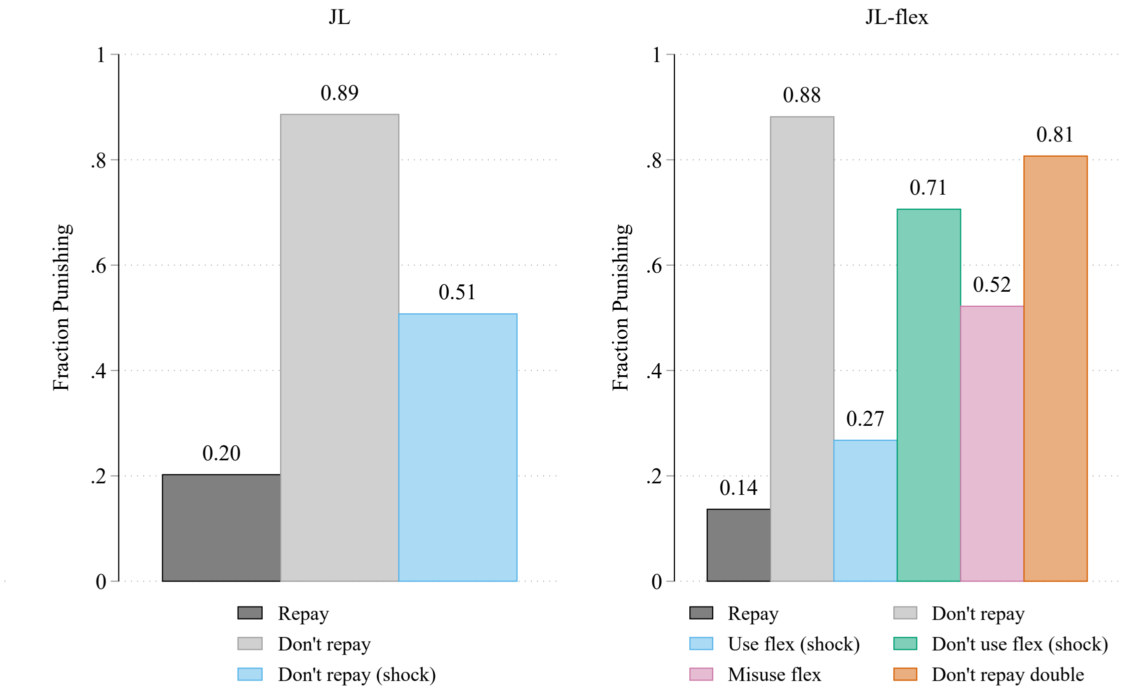
Panel B	**Norms: Main Comparisons**		
Action	Mean	SD	Wilcoxon signed-rank test
Use of flex in case of shock			$p < 0.001$
No flex: Don't repay (shock) vs.	0.85	0.15	
Flex: Use flex (shock)	0.47	0.3	
Don't repay in two consecutive periods (per-period averages)			$p < 0.001$
No flex: Don't repay & Don't repay, second time vs.	0.88	0.17	
Flex: Misuse flex & Don't repay double	0.79	0.20	

Notes: The table is ordered according to the severity of actions (measured in terms of punishment or inappropriateness). Punishment refers to choices in the experiment in the *JL* and *JL-flex* treatments. Norm vignettes refer to the repayment choice in week 3 in a 25-week loan cycle, except for 'Action in subsequent period,' which refers to week 4. Inappropriateness of a given action is measured on a six-point Likert-scale (+++ indicating very high social appropriateness, — very low social inappropriateness), rescaled for comparability to 0-1 with higher numbers indicating higher inappropriateness. For implementation reasons, the inappropriateness rating of 'Don't repay double' conditions on previous misuse of flexibility, while the corresponding punishment does not condition on why flexibility was used.

Online Appendix

A Additional Tables and Figures

Figure A.1: Incidence of punishment



Notes: Incidence of punishment: fraction of participants who decide to allocate at least one punishment point. Punishment choices are conditional on partner's action and shock arrival as indicated.

Table A.1: Full Repayment

	Main Sample			Incl. those who failed comprehension test		
	IL vs. IL-flex	IL vs. JL	JL vs. JL-flex	IL vs. IL-flex	IL vs. JL	JL vs. JL-flex
	(1)	(2)	(3)	(4)	(5)	(6)
IL-flex	-0.165***			-0.182***		
	(0.0355)			(0.0350)		
JL		0.0221			0.0209	
		(0.0342)			(0.0332)	
JL-flex			-0.164***			-0.175***
			(0.0398)			(0.0388)
Mean of DV	0.679	0.643	0.665	0.679	0.641	0.661
Observations	607	544	526	632	574	555
No. of participants	305	272	272	318	287	287
R^2	0.076	0.085	0.081	0.081	0.081	0.084
Session FE	yes	yes	yes	yes	yes	yes

Notes: Dependent variable is a binary indicator for full repayment. Columns 1–3 show regressions for the main sample. Columns 4–6 include participants who failed the comprehension test, as defined in footnote 17. OLS regressions with session fixed effects and standard errors clustered on individual level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.2: Full Repayment: Individual FE

	Main Sample			Incl. those who failed comprehension test		
	IL vs. IL-flex	IL vs. JL	JL vs. JL-flex	IL vs. IL-flex	IL vs. JL	JL vs. JL-flex
	(1)	(2)	(3)	(4)	(5)	(6)
ILflex	-0.166***			-0.185***		
	(0.0497)			(0.0490)		
JL		0.0221			0.0209	
		(0.0477)			(0.0463)	
JLflex			-0.154***			-0.164***
			(0.0571)			(0.0556)
Mean of DV	0.679	0.643	0.665	0.679	0.643	0.665
Observations	607	544	526	632	574	555
No. of participants	305	272	272	318	287	287
R^2	0.618	0.659	0.604	0.617	0.663	0.606
Individual FE	yes	yes	yes	yes	yes	yes

Notes: Dependent variable is a binary indicator for full repayment. Columns 1–3 show regressions for the main sample. Columns 4–6 include participants who failed the comprehension test, as defined in footnote 17. OLS regressions with individual fixed effects and standard errors clustered on individual level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.3: Use of Flexibility

	Main Sample				Incl. those who failed comprehension test			
	Shock		No Shock		Shock		No Shock	
	T1	T2	T1	T2	T1	T2	T1	T2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
JL-flex	0.0198	-0.0433*	-0.258***	-0.207***	0.0253	-0.0414*	-0.251**	-0.208***
	(0.0185)	(0.0255)	(0.0939)	(0.0698)	(0.0179)	(0.0244)	(0.0942)	(0.0686)
	[-0.0220, 0.0531]	[-0.0848, 0.0101]	[-0.4365, -0.0904]	[-0.3371, -0.0806]	[-0.0051, 0.0588]	[-0.0849, 0.0054]	[-0.4310, -0.0720]	[-0.3379, -0.0696]
Mean of DV in IL-flex	0.951	0.970	0.549	0.723	0.943	0.971	0.552	0.718
Observations	575	327	575	329	603	340	603	342
No. of participants	575	327	575	329	603	340	603	342
R^2	0.003	0.009	0.068	0.043	0.004	0.008	0.064	0.044

Notes: Dependent variable is a binary indicator for using flexibility in a given scenario. Using flexibility in T2 is conditional on still having it, i.e. on not having used it in T1. Columns 1–4 show regressions for the main sample. Columns 5–8 include participants who failed the comprehension test, as defined in footnote 17. OLS regressions with standard errors clustered on session level in parentheses and wild cluster bootstrapped 95% confidence intervals in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.4: Punishment: Main Sample

Panel A: Punishment in case of no shock

	Level of Punishment				Incidence of Punishment			
Choice in JL	Repay	Don't Repay	Don't Repay	Don't Repay	Repay	Don't Repay	Don't Repay	Don't Repay
Choice in JL-flex	Repay	Don't Repay	Misuse Flex	Don't Repay Double	Repay	Don't Repay	Misuse Flex	Don't Repay Double
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
JL-flex	-0.0450**	-0.00623	-0.220***	0.0422	-0.0470	-0.00653	-0.367***	-0.0856***
	(0.0211)	(0.0269)	(0.0341)	(0.0296)	(0.0296)	(0.0246)	(0.0383)	(0.0285)
Mean of DV in JL	0.14	0.61	0.61	0.61	0.20	0.89	0.89	0.89
Observations	499	499	499	499	499	499	499	499
No. of Participants	272	272	272	272	272	272	272	272
R^2	0.052	0.104	0.169	0.151	0.058	0.148	0.247	0.148
Session FE	yes	yes	yes	yes	yes	yes	yes	yes

Panel B: Punishment in case of shock

	Level of Punishment			Incidence of Punishment		
Choice in JL	Don't Repay	Don't Repay	Don't Repay	Don't Repay	Don't Repay	Don't Repay
Choice in JL-flex	Use Flex	Don't Use Flex	Don't Repay Double	Use Flex	Don't Use Flex	Don't Repay Double
	(1)	(2)	(3)	(5)	(6)	(7)
JL-flex	-0.203***	0.144***	-0.0536*	-0.231***	0.208***	-0.0818**
	(0.0322)	(0.0330)	(0.0295)	(0.0409)	(0.0409)	(0.0375)
Mean of DV in JL	0.38	0.38	0.38	0.51	0.51	0.51
Observations	498	498	498	498	498	498
No. of Participants	270	270	270	270	270	270
R^2	0.128	0.143	0.117	0.129	0.127	0.133
Session FE	yes	yes	yes	yes	yes	yes

Notes: Dependent variable in Columns 1-4 is the level of punishment as measured by (0-1 rescaled) punishment points and in Columns 5-8 incidence of punishment, i.e. percent of participants who punish. The omitted category is *JL*. Choices with and without flexibility are matched on i) shock realization and ii) the amount repaid. Number of responses in *JL*: 272 (270) and *JL-flex*: 227 (228) for no shock (shock) scenario. OLS regressions with session fixed effects and standard errors clustered on individual level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.5: Punishment: Robustness

Panel A: Level of punishment in case of no shock

	Main Analysis Sample				Incl. those who failed comprehension test			
Choice in JL	Repay	Don't Repay	Don't Repay	Don't Repay	Repay	Don't Repay	Don't Repay	Don't Repay
Choice in JL-flex	Repay	Don't Repay	Misuse Flex	Don't Repay Double	Repay	Don't Repay	Misuse Flex	Don't Repay Double
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
JL-flex	-0.0450**	-0.00623	-0.220***	0.0422	-0.0432**	0.000413	-0.220***	0.0534*
	(0.0211)	(0.0269)	(0.0341)	(0.0296)	(0.0206)	(0.0264)	(0.0336)	(0.0292)
Mean of DV in JL	0.14	0.61	0.61	0.61	0.14	0.62	0.62	0.62
Observations	499	499	499	499	523	523	523	523
No. of Participants	272	272	272	272	287	287	287	287
R^2	0.052	0.104	0.169	0.151	0.058	0.102	0.168	0.144
Session FE	yes	yes	yes	yes	yes	yes	yes	yes

Panel B: Level of punishment in case of shock

	Main Analysis Sample				Incl. those who failed test questions			
Choice in JL	Don't Repay	Don't Repay	Don't Repay		Don't Repay	Don't Repay	Don't Repay	
Choice in JL-flex	Use Flex	Don't Use Flex	Don't Repay Double		Use Flex	Don't Use Flex	Don't Repay Double	
	(1)	(2)	(3)		(5)	(6)	(7)	
JL-flex	-0.203***	0.144***	-0.0536*		-0.210***	0.134***	-0.0579**	
	(0.0322)	(0.0330)	(0.0295)		(0.0314)	(0.0325)	(0.0294)	
Mean of DV in JL	0.38	0.38	0.38		0.39	0.39	0.39	
Observations	498	498	498		522	522	522	
No. of Participants	270	270	270		285	285	285	
R^2	0.128	0.143	0.117		0.138	0.134	0.112	
Session FE	yes	yes	yes		yes	yes	yes	

Notes: Dependent variable is the level of punishment as measured by (0-1 rescaled) punishment points. Columns 1–4 show regressions for the main sample. Columns 5–8 include participants who failed the comprehension test, as defined in footnote 17. The omitted category is *JL*. Choices with and without flexibility are matched on i) shock realization and ii) the amount repaid. OLS regressions with session fixed effects and standard errors clustered on individual level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.6: Punishment: Individual Fixed Effects

Panel A: Punishment in case of no shock

	Level of Punishment				Incidence of Punishment			
Choice in JL	Repay	Don't Repay	Don't Repay	Don't Repay	Repay	Don't Repay	Don't Repay	Don't Repay
Choice in JL-flex	Repay	Don't Repay	Misuse Flex	Don't Repay Double	Repay	Don't Repay	Misuse Flex	Don't Repay Double
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
JLflex	-0.0419	-0.00881	-0.222***	0.0396	-0.0441	-0.00881	-0.370***	-0.0881**
	(0.0306)	(0.0393)	(0.0499)	(0.0433)	(0.0432)	(0.0359)	(0.0559)	(0.0416)
Mean of DV in JL	0.14	0.61	0.61	0.61	0.20	0.89	0.89	0.89
Observations	499	499	499	499	499	499	499	499
No. of Participants	272	272	272	272	272	272	272	272
R^2	0.716	0.640	0.601	0.647	0.694	0.703	0.636	0.684
Individual FE	yes	yes	yes	yes	yes	yes	yes	yes

Panel B: Punishment in case of shock

	Level of Punishment				Incidence of Punishment			
Choice in JL	Don't Repay	Don't Repay	Don't Repay		Don't Repay	Don't Repay	Don't Repay	
Choice in JL-flex	Use Flex	Don't Use Flex	Don't Repay Double		Use Flex	Don't Use Flex	Don't Repay Double	
	(1)	(2)	(3)		(5)	(6)	(7)	
JLflex	-0.202***	0.145***	-0.0526		-0.228***	0.211***	-0.0789	
	(0.0470)	(0.0481)	(0.0431)		(0.0596)	(0.0596)	(0.0546)	
Mean of DV in JL	0.38	0.38	0.38		0.51	0.51	0.51	
Observations	498	498	498		498	498	498	
No. of Participants	270	270	270		270	270	270	
R^2	0.645	0.676	0.744		0.647	0.650	0.715	
Individual FE	yes	yes	yes		yes	yes	yes	

Notes: Dependent variable in Columns 1-4 is the level of punishment as measured by (0-1 rescaled) punishment points and in Columns 5-8 incidence of punishment, i.e. percent of participants who punish. The omitted category is *JL*. Choices with and without flexibility are matched on i) shock realization and ii) the amount repaid. Number of responses in *JL*: 272 (270) and *JL-flex*: 227 (228) for no shock (shock) scenario. OLS regressions with individual fixed effects and standard errors clustered on individual level in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.7: Partial Repayments

	N	Partial Repayers	Share Repaying
IL	577	110	
T1 repaid			.70
T2 repaid			.38
T3 repaid			.44
JL	272	49	
T1 repaid			.65
T2 repaid			.41
T3 repaid			.59
IL-flex	302	100	
T1 repaid			.31
T1 flex used			.45
T2 repaid			.38
T2 flex used			.44
T3 repaid			.37
Use flex, repay single but not double			.39
Use flex, repay double but not single			.50
No flex used			.11
JL-flex	254	93	
T1 repaid			.42
T1 flex used			.17
T2 repaid			.33
T2 flex used			.41
T3 repaid			.46
Use flex, repay single but not double			.23
Use flex, repay double but not single			.35
No flex used			.42

Notes: Share of participants (of those who only partially repay) who make a given installment.

Table A.8: Partial Repayment and Confusion

	(1)	(2)	(3)
High Comprehension	-0.0807***	-0.0153	0.00142
	(0.0237)	(0.0267)	(0.0312)
Flexibility		0.153***	0.173***
		(0.0342)	(0.0404)
High Comprehension x Flex			-0.0359
			(0.0496)
Mean of DV	0.250	0.250	0.250
Observations	1,405	1,405	1,405
No. of Participants	577	577	577
R^2	0.008	0.033	0.033

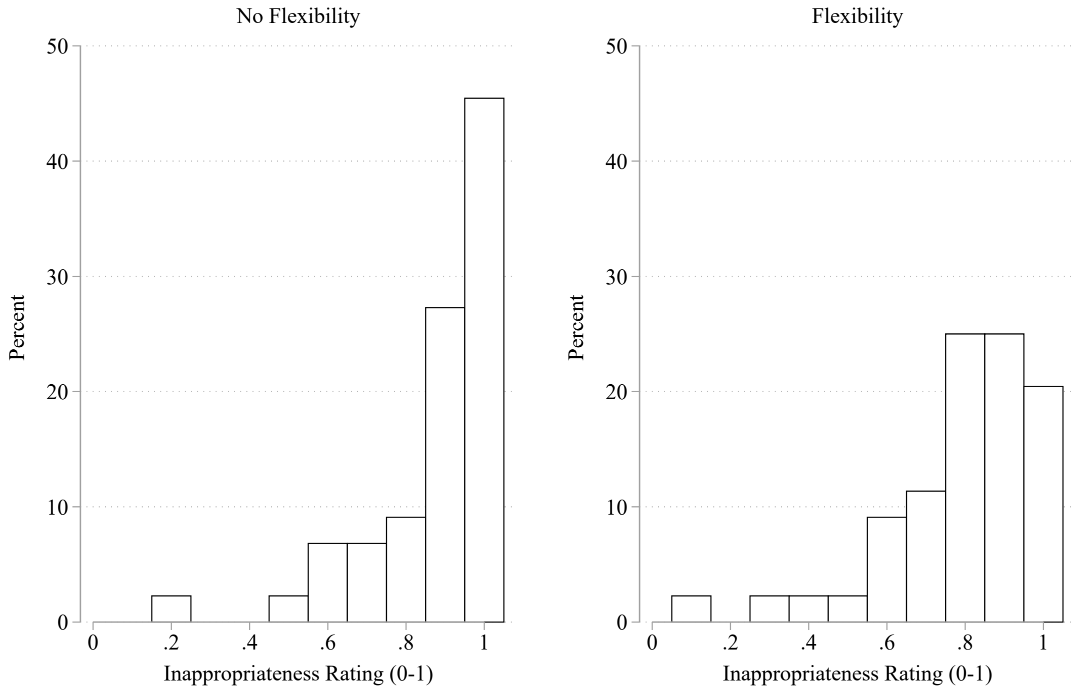
Notes: Dependent variable is a binary indicator for partial repayment in a given treatment. High comprehension is an indicator for above-median performance on test questions for a given treatment. Flexibility is a binary indicator for *IL-flex* and *JL-flex* treatments. OLS regressions with standard errors clustered on session level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.9: Qualitative Interviews on Norms

	Share Agree	Share Agree strongly	Mean Likert scale (1 to 4)	Std. Dev. Likert scale (1 to 4)
Repaying is the right/moral thing to do	0.35	0.61	3.57	0.59
We learned in our group training that repayment is essential.	0.35	0.61	3.57	0.59
Our loan officer emphasizes that we should repay each week.	0.43	0.57	3.57	0.51
The group has discussed the undesirability of non-repayment extensively with each other.	0.57	0.13	2.78	0.74

Notes: Agreement of 23 borrowers is measured on a four-point Likert-scale (1-4): 1) disagree strongly, 2) disagree, 3) agree, 4) agree strongly.

Figure A.2: Histogram of the Injunctive Norm for Defaulting Twice



Notes: No flexibility: averaged inappropriateness ratings of actions i) 'Don't repay' and ii) 'Don't repay, second time.' Flexibility: averaged inappropriateness ratings of actions i) 'Misuse flex' and ii) 'Don't repay double.'

Table A.10: Correlations of Lab Behavior and Real-Life Measures

	Full repayment		Flex only for shock		Punish only voluntary default	
	(1)	(2)	(3)	(4)	(5)	(6)
No delinquencies	0.0753		0.0231		0.0396	
	(0.0461)		(0.0568)		(0.0979)	
New member		-0.0611		-0.140***		-0.147
		(0.0594)		(0.0491)		(0.114)
Mean of DV in omitted	0.527	0.602	0.213	0.141	0.357	0.397
Observations	1077	1095	1768	1796	212	215
No. of Participants	442	449	442	449	212	215
R^2	0.00281	0.00141	0.000355	0.0106	0.000754	0.00628

Notes: Dependent variable in Columns 1–2 is an indicator for fully repaying the loan in the microcredit game, in Columns 3–4 it is an indicator for using flexibility only in case of experiencing a shock, and in Columns 5–6 it is an indicator for punishing only voluntary default in *JL*. *No delinquencies* is an indicator for always repaying loan installments on time, and *new member* is an indicator for participants who are clients of our partner organization for less than a year. OLS regressions with standard errors clustered at the participant level. *** p<0.01, ** p<0.05, * p<0.1.

Table A.11: Balance by Session Type and Order Effects

	Repaid in IL (full) (1)	Repaid in IL (by period) (2)	Repay after shock default (3)	Repay after shock default (4)
JL-Session	-0.0353	0.00435		
	(0.0559)	(0.0428)		
Period 2		-0.0525***		
		(0.0177)		
Period 3		-0.0393**		
		(0.0192)		
JL-Session x Period 2		-0.0174		
		(0.0313)		
JL-Session x Period 3		-0.0232		
		(0.0292)		
JL			-0.00735	-0.00735
			(0.0453)	(0.0588)
IL-flex			-0.0168	-0.0138
			(0.0291)	(0.0367)
JL-flex			-0.0163	-0.0140
			(0.0422)	(0.0544)
Mean of DV in IL	0.679	0.793	0.234	0.234
Observations	577	1,731	1,379	1,379
No. of participants	577	577	577	577
R^2	0.001	0.004	0.130	0.622
Session FE			yes	no
Individual FE			no	yes

Notes: Dependent variable as indicated. OLS regressions with standard errors clustered on session level in parentheses. The sample in Column 2 includes one observation per participant and period in *IL*. The sample in Columns 3 and 4 includes one observation per participant and treatment played. *** p<0.01, ** p<0.05, * p<0.1.

Table A.12: Full Repayment and Risk Aversion

	IL (1)	JL (2)	ILflex (3)	JLflex (4)	IL vs ILflex (5)	IL vs JL (6)	JL vs JLflex (7)
Risk Averse	0.0142	-0.0224	0.0457	-0.0176	0.0168	-0.0418	-0.0676
	(0.0453)	(0.0733)	(0.0625)	(0.0732)	(0.0554)	(0.0682)	(0.0693)
IL-flex					-0.184***		
					(0.0528)		
IL-flex x Risk Averse					0.0381		
					(0.0712)		
JL						-0	
						(0.0451)	
JL x Risk Averse						0.0583	
						(0.0689)	
JL-flex							-0.202***
							(0.0484)
JL-flex x Risk Averse							0.0964
							(0.0840)
Mean of DV (in omitted)	0.662	0.665	0.513	0.500	0.679	0.643	0.665
Observations	577	272	302	254	607	544	526
No. of participants	577	272	302	254	305	272	272
R^2	0.113	0.083	0.060	0.108	0.077	0.086	0.083
Session FE	yes	yes	yes	yes	yes	yes	yes

Notes: Dependent variable is a binary indicator for full repayment. Columns 1–4 show regressions separately by treatment, Columns 5–7 replicate Columns 1–3 of Table A.1 including an indicator for above median risk aversion and an interaction of this variable with the respective treatment indicator. OLS regressions with session fixed effects and standard errors clustered on individual level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.13: Use of Flexibility and Risk Aversion

	Shock		No Shock		Shock		No Shock	
	IL-flex	JL-flex	IL-flex	JL-flex				
	T1	T1	T1	T1	T1	T2	T1	T2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Risk Averse	-0.0178	0.0320	0.0112	0.141**	-0.0178	0.000439	0.0112	0.0332
	(0.0278)	(0.0196)	(0.0820)	(0.0634)	(0.0274)	(0.0217)	(0.0809)	(0.0584)
JL-flex					-0.00140	-0.0405	-0.306***	-0.198**
					(0.0205)	(0.0330)	(0.109)	(0.0932)
JL-flex x Risk Averse					0.0498	-0.00825	0.129	-0.00981
					(0.0335)	(0.0428)	(0.102)	(0.135)
Mean of DV (in IL-flex)	0.951	0.970	0.549	0.292	0.951	0.970	0.549	0.723
Observations	304	271	304	271	575	327	575	329
No. of participants	304	271	304	271	575	327	575	329
R^2	0.002	0.008	0.000	0.023	0.007	0.009	0.077	0.044

Notes: Dependent variable is a binary indicator for using flexibility in a given scenario. Using flexibility in T2 is conditional on still having it, i.e. on not having used it in T1. Columns 1-4 present results for using flexibility in T1 in different shock scenarios and treatments. Columns 5–8 replicate Columns 1–4 of Table A.3 including an indicator for above median risk aversion and an interaction of this variable with the indicator for JL-flex. OLS regressions with standard errors clustered on session level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.14: Punishment (Incidence) and Risk Aversion

	Means			Difference
	Sample	Less Risk Averse	More Risk Averse	(3)-(2)
	(1)	(2)	(3)	(4)
JL				
punish repay	0.202	0.195	0.214	0.018
	(0.402)	(0.398)	(0.412)	(0.653)
punish non-repayment (absent shock)	0.886	0.876	0.903	0.027
	(0.318)	(0.331)	(0.298)	(0.529)
punish non-repayment (shock)	0.507	0.482	0.549	0.067
	(0.501)	(0.501)	(0.500)	(0.540)
JL-flex				
punish repay	0.137	0.119	0.163	0.045
	(0.344)	(0.324)	(0.371)	(0.278)
punish non-repayment (absent shock)	0.882	0.890	0.870	-0.020
	(0.324)	(0.314)	(0.339)	(0.777)
punish flex use (no shock)	0.522	0.544	0.489	-0.055
	(0.501)	(0.500)	(0.503)	(0.505)
punish double repayment	0.154	0.118	0.207	0.089**
	(0.361)	(0.323)	(0.407)	(0.034)
punish non-repayment after flex use (no shock)	0.807	0.757	0.880	0.123
	(0.396)	(0.430)	(0.326)	(0.175)
punish non-repayment after flex use (shock)	0.417	0.419	0.413	-0.006
	(0.494)	(0.495)	(0.495)	(0.936)
punish flex use (shock)	0.268	0.265	0.272	0.007
	(0.444)	(0.443)	(0.447)	(0.909)
punish not using flex (shock)	0.706	0.735	0.663	-0.072
	(0.457)	(0.443)	(0.475)	(0.358)
Observations	577	318	259	577

Notes: Column 1 shows sample means. Columns 2 and 3 show means for below and above median risk aversion, respectively. Column 4 shows the difference between Columns 3 and 2. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.15: Full Repayment and Present Bias

	(1) IL	(2) JL	(3) ILflex	(4) JLflex	(5) IL vs ILflex	(6) IL vs JL	(7) JL vs JLflex
Present Bias	-0.0432	-0.0918	0.0562	-0.0939	-0.0216	-0.102	-0.128
	(0.0534)	(0.0863)	(0.0720)	(0.0878)	(0.0674)	(0.0842)	(0.0848)
IL-flex					-0.182***		
					(0.0397)		
IL-flex x Present Bias					0.0822		
					(0.0889)		
JL						0.0174	
						(0.0365)	
JL x Present Bias						0.0302	
						(0.103)	
JL-flex							-0.176***
							(0.0428)
JL-flex x Present Bias							0.0719
							(0.116)
Mean of DV (in omitted)	0.673	0.683	0.498	0.505	0.680	0.665	0.665
Observations	577	272	302	254	607	544	526
R^2	0.114	0.087	0.060	0.112	0.077	0.089	0.086
Session FE	yes	yes	yes	yes	yes	yes	yes

Notes: Dependent variable is a binary indicator for full repayment. Columns 1–4 show regressions separately by treatment, Columns 5–7 replicate Columns 1–3 of Table A.1 including an indicator for present bias (being less patient today than in one month) and an interaction of this variable with the respective treatment indicator. OLS regressions with session fixed effects and standard errors clustered on individual level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

B Theory Appendix

B.1 Equilibrium Behavior with Flexibility

Using flexibility is dominant in the case of shocks. Furthermore, straight default is dominated by using flexibility at first, and then defaulting. This is because the penalty for invoking flexibility $\lambda\kappa$ is weaker than that for simple non-repayment, κ. Focusing on the game path without shocks, borrowers are left with four non-dominated strategies:

1. Flex-S: Use flexibility only if shocks arrive, repay in periods 1, 2, and 3.

$$U_1^{Flex-S} = R + (1-\theta)(\delta R + \delta^2 R + \delta^n V) + \theta(-\delta\lambda\kappa + 0 + \delta^n V) \quad \text{(A1)}$$

2. Flex-2: Use flexibility in period 2 (shocks in period 2 are insured), repay in periods 1 and 3.

$$U_1^{Flex2} = R + (1-\theta)(\delta(2R - \lambda\kappa) + 0 + \delta^n V) + \theta(-\delta\lambda\kappa + 0 + \delta^n V). \quad \text{(A2)}$$

3. Flex-1: Use flexibility in period 1 (and forfeit shock insurance in period 2), repay in periods 2 and 3.

$$U_1^{Flex1} = 2R - \lambda\kappa + (1-\theta)(0 + \delta^2 R + \delta^n V) + \theta(-\delta\kappa + \delta^2(2R - \kappa)). \quad \text{(A3)}$$

4. Flex-D: Use flexibility in period 1, then default.

$$U_1^{Flex-D} = 2R - \lambda\kappa + (1-\theta)(\delta(2R - \kappa) + \delta^2(2R - \kappa)) + \theta(-\delta\kappa + \delta^2(2R - \kappa)) \quad \text{(A4)}$$

We restrict our attention to cases where repayment is sustained by a social norm. In particular, we assume that the continuation value V is sufficiently low that borrowers default at $\kappa = 0$ in both rigid and flexible repayment conditions (as shown below):

$$\delta^{n-1} V < R(2 + \delta) \quad \text{(A5)}$$

To identify the weakest condition that will guarantee repayment, it is necessary to identify the profile of strategies that will be played as κ increases from zero to R. It is straightforward to show that, at $\kappa=0$, $U_1^{Flex-D} > U_1^{Flex-1} > U_1^{Flex-2} > U_1^{Flex-S}$. While default is most attractive at $\kappa=0$, the expected utility from default decreases quickly as κ increases: $\frac{\partial U^{Flex-D}}{\partial \kappa} = -\lambda - (\delta+\delta^2)$. The strategies have the reverse order in their slope in κ, with Flex-D being the most sensitive to κ, and Flex-S being the least sensitive:

$$\frac{\partial U^{Flex-D}}{\partial \kappa} < \frac{\partial U^{Flex-1}}{\partial \kappa} < \frac{\partial U^{Flex-2}}{\partial \kappa} < \frac{\partial U^{Flex-S}}{\partial \kappa} < 0. \tag{A6}$$

This can be seen in Figure B.3. As κ increases, the borrower moves from Flex-D to Flex-1. Comparing equalities A3 and A4, Flex-1 is preferred to Flex-D if

$$\kappa \geq \frac{R[2+\delta] - \delta^{n-1}V}{[1+\delta]}. \tag{A7}$$

As κ increases further, Flex-2 and finally Flex-S become attractive. Flexibility is misused to increase early consumption as long as Flex-2 is preferred to Flex-S, i.e., if A2 is larger than A1: $(1-\delta)R > \lambda\kappa$ Combining the two inequalities yields that flexibility is misused (Flex-1 and Flex-2) for intermediate values of κ:

$$\frac{R[2+\delta] - \delta^{n-1}V}{[1+\delta]} \leq \kappa < \frac{(1-\delta)}{\lambda} R. \tag{A8}$$

This yields a positive interval for κ if

$$\lambda \leq \frac{(1-\delta^2)}{[2+\delta] - \delta^{n-1}\frac{V}{R}} \tag{A9}$$

Condition A9 is important because it guarantees that the order of strategies is Flex-D, Flex-1, Flex-2, and finally Flex-S as κ increases. In turn, this implies that $U_1^{Flex-1} \geq U_1^{Flex-D}$ (inequality A7) is the weakest condition that guarantees repayment under flexibility. For higher values of λ, the borrower moves from Flex-D (for low κ) directly to Flex-S (for higher κ), without misusing flexibility. As a result, the weakest repayment condition becomes $U_1^{Flex-S} \geq U_1^{Flex-D}$ (curve 'Flex-S' in Figure B.3). Because 87 percent of participants in our sample misuse flexibility absent shocks, we focus on the case where condition A9 holds.

Rearranging inequality A7 for an easier comparison with the rigid benchmark case, the repayment condition under flexibility is

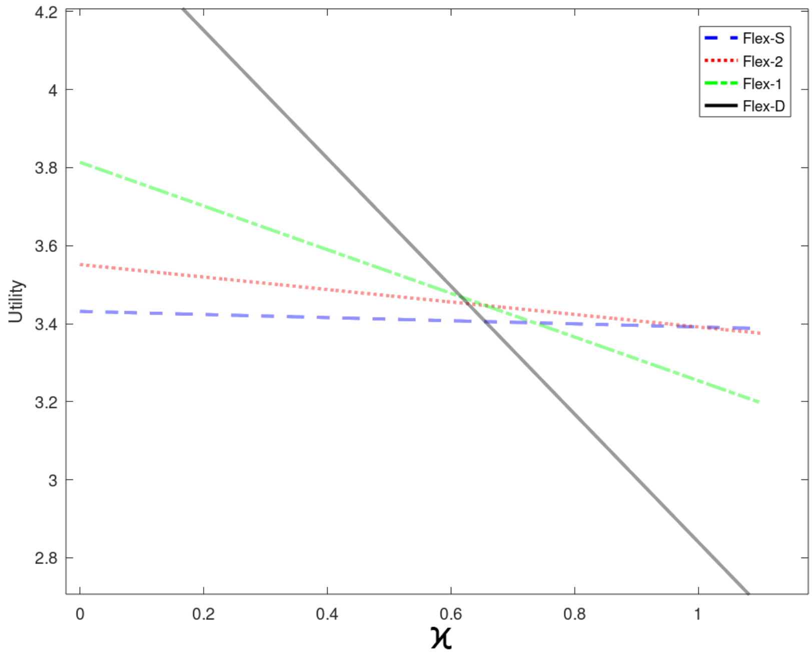
$$\delta^{n-1}V \geq R[2+\delta] - \kappa[1+\delta]. \tag{A10}$$

Comparing inequality A10 with the repayment condition from the benchmark case (inequality 8), flexibility imposes stronger conditions (causes more defaults) than rigid repayment if

$$\kappa \geq (1-(1-\theta)\delta)R. \tag{A11}$$

Default rates will be higher under flexible repayment contracts for large κ, large δ, and small θ.

Figure B.3: Repayment Behavior with Flexible Repayment



Notes: Curves represent the expected utility from a given strategy under flexible repayment (equations A1 to A4). This figure shows simulations using $V=3.3$, $R=1$, $n=4$, $\theta=0.25$, and $\lambda=0.2$. Note the experiment induced $V/R=3.3$, $\theta=0.25$, and $\delta\approx0.8$ (see Section 2.1).

B.2 Derivation of Predictions

Prediction 1. [Overall repayment] *In the presence of strong social norms, repayment is higher under a rigid repayment contract than under a flexible repayment contract.*

Corollary 1. *Absent social norms, repayment flexibility leads to strictly higher repayment rates.*

The prediction and its corollary follow from inequality A11. Prediction 1 follows for sufficiently large κ, while its corollary follows for $\kappa=0$.

Prediction 2. [Default path] *Using flexibility at first and then defaulting strictly dominates defaulting straight away.*

The expected utility from defaulting can be inferred from equation A4 with λ set to 1. Prediction 2 follows from $\lambda < 1$.

> **Prediction 3.** [Flexibility misuse] *The insurance value of flexibility decreases over time. Thus, misuse of flexibility (use of flexibility absent shocks) will increase over time, conditional on flexibility still being available.*

Using flexibility in absence of a shock yields early consumption $R - \lambda\kappa$ in exchange for a delayed payment δR, and is thus attractive as long as $\lambda\kappa < (1-\delta)R$ (see condition A8). Comparing equations A2 and A3, flexibility use in period 1 is unconditionally preferred to flexibility use in period 2 if

$$\theta\delta^n V < R\cdot[1+\delta(\theta\delta - 2 + \delta + 2\theta)] - \kappa[\lambda + \delta(\theta\delta + \theta - \lambda)] \tag{A12}$$

Prediction 3 makes a simpler, conditional statement: In the moment of using it, flexibility has the same benefit in periods 1 and 2, but the cost of using it is lower in period 2. Thus, conditional on still being available, flexibility use should be higher in period 2 than in period 1.

> **Prediction 4**. [Partial repayments] *In rigid repayment contracts, partial payments are always dominated by full repayment or full default. With repayment flexibility, partial repayments can be optimal: Borrowers may comply with single but not double installments.*

In rigid repayment contracts, adding individual repayments to the default path (equation 7) incurs a cost of $R - \kappa$, and does not obtain the continuation value V. For $\kappa < R$, making no repayments dominates making partial repayments. For $\kappa \geq R$, making all repayments dominates partial repayments.

In flexible repayment contracts, the borrower faces a double installment $2R$ after using flexibility to defer repayment. By assumption, the social norm imposes a penalty for not repaying when asked, and thus the penalty for defaulting on a double repayment is still κ. Starting from a full repayment strategy with flexibility use (e.g. equation A3), defaulting on the double repayment yields $2R - \kappa - \delta^{n-1}V$. In cases where $R \leq \kappa < 2R$, and for low V or δ, borrowers may thus find it profitable to comply with single installments but default on double installments.

C Potential Confounds

Balance across session types We randomly allocated centers to either IL- or JL-Sessions. To avoid confounds in our analysis, it is necessary to rule out systematic differences between sessions. As shown in Table 1, observable characteristics are balanced across session types. In both types of sessions, we administer a benchmark *IL* treatment, such that choices in this treatment serve as an additional randomization check. Column 1 in Table A.11 shows that the share of participants who fully repays their IL loan is the same in both session types. This also holds true if we examine the three periods in *IL* separately (Column 2).

Order effects Our treatments are run in a constant order (see Section 2.2), giving rise to potential concerns about order effects between our treatments. Several arguments alleviate such concerns. First, since we elicited choices with the strategy method and only realized shock outcomes, matching (where applicable) and corresponding earnings after all decisions had been made, participants did not receive intermediate feedback which may have permitted learning.

Second, the mere repetition of experimental decisions may have led participants to realize that repayment is not monetarily profitable, and thus to default more. This mechanism is inconsistent with the fact that the observed increase in overall loan defaults in flexible treatments is entirely driven by an increase in partial repayments (Figure 7): Partial repayments are monetarily dominated by full repayment, full default, or both. Our results are thus not explained by participants learning how to maximize payoffs over time. This conclusion is further supported by a technical detail: To account for the uncertainty regarding future shock arrival when making repayment choices, we also elicit a repayment choice conditional on being in default following a shock. Around 22 percent repay after a shock-induced default, with no variation between treatments (Table A.11). If learning took place, we should see a decrease in this fraction over time, and thus in later treatments.

Other effects relating to the order of treatments are possible; for instance participants' concentration may decrease over time. Several pieces of evidence speak against order effects in a more general sense: Repayment behavior is identical in the first two treatments in JL-Sessions (*IL* and *JL*). Two interpretations are possible: Either, there are no order effects and no effects of joint liability on repayment (consistent with the existing evidence, e.g., Giné and Karlan (2014)). Or these two effects exactly cancel each other out. Since there is no theoretical reason to expect the latter, we deem the former to be more plausible. Furthermore, overall repayment drops when we introduce flexibility, but this introduction happens in the second treatment in IL-Sessions, and in the third treatment in JL-Sessions. Despite both the difference in order and the difference in liability structure, the drop in repayment has a similar magnitude in both session types. Our evidence thus suggests that participants do not simply change their behavior as a function of time and treatment order, but that the treatments themselves cause behavior to change.

Confusion Our main analysis excludes participants who failed to meet our comprehension threshold, based on the test questions described in Section 2.4 and footnote 17. To make sure this is not driving our results, we repeat our analyses including all participants. All observed effects of flexibility and joint liability on overall loan repayment (Table A.1), as well as the effect of joint liability on flexibility use (Table A.3) are robust to the inclusion. The same holds true for punishment choices: In Table A.5, we replicate results regarding the level of punishment from Table A.4 with very similar coefficients. The fact that our analyses are robust to the inclusion of confused participants suggests that lack of comprehension is unlikely to be an important driver of our results.

We further explore in how far confusion can explain partial repayments. We first study the relationship between partial repayments and the treatment-specific comprehension score. In Column 1 of Table A.8, we find that above-median comprehension predicts an 8 percentage point (27 percent) decrease in the incidence of partial repayments. Confusion thus helps explain the presence of partial repayments at baseline. This effect is small relative to the 15 percentage point increase in partial repayments in flexibility treatments (see Column 2 and Figure 7). We test whether this increase is driven by reduced comprehension in flexibility treatments, but find no significant interaction between the two. Confusion thus helps explain the presence of partial repayments, but not their increase in the flexibility treatments.

Communication Spillovers The sessions were conducted in microfinance centers, with only one session run in each center. The microfinance centers were geographically spread across three provinces of the Philippines: Rizal, Laguna, and Antique. Where two microfinance centers were close to each other (within approx. 5km), their sessions were conducted on the same day (morning and afternoon). This means communication spillovers would have had a maximum of two hours to move from one center to a neighboring center. Note that Coutts (2022) finds that spillovers cease to be significant for distances greater than 2.5 km, even when information has several days or weeks to travel.

D Expected Payouts

Strategies

In line with our analysis of overall repayment, we focus on strategies which lead to either full repayment or no repayment on the no-shock path.[45] In the *IL* and *JL* conditions, this leads to two

[45] With three periods, eight states of the world depending on shock realizations, and between one and three possible actions per period and state, it is impractical to document expected payouts for all possible strategies.

strategies: Repay all installments (denoted *RRR*), or default on all installments (*DDD*). In the *IL-flex* and *JL-flex* conditions, participants who wish to repay have three main strategies at their disposal (note that *F* refers to flexibility use, and R_2 refers to a subsequent double repayment): They can repay every period (*RRR*), they can misuse flexibility in period 1 (FR_2R, thus foregoing shock insurance in period 2), or they can misuse flexibility in period 2 (RFR_2). In all three full-repayment strategies, we assume that flexibility is used *in case of shocks* (consistent with our experimental findings in Section 4.2). On the opposite extreme, participants may choose to default on the overall loan by not making any repayments (*DDD*). To provide a benchmark and allow comparisons to the *IL* and *JL* conditions, this strategy assumes that flexibility is not used in any state of the world. Finally, we consider a strategy that is monetarily equivalent to straight default, but socially more sophisticated: Participants hide their plan to default by using flexibility in the first period (regardless of shock arrival), and default on all installments starting in period 2 (*FDD*).

Calculations

Expected payouts are calculated from the values of income tokens used for consumption, as well as from the continuation value V, which was a 100 pesos cash payment after one month. Payouts account for the shock probability $\theta = 0.25$ in each period, and the action prescribed by each strategy in each state of the world. For instance, the expected payout for strategy *RRR* in condition *IL-Flex* is

$$E(\pi_{RRR}|IL\text{-}Flex) = (1-\theta)R^{t1} + (1-\theta)^2 R^{t2} + ((1-\theta)^2 + 2\theta^2(1-\theta))R^{t3} + (1-\theta)^2(1+\theta)\delta_M \cdot V$$

where R^{ti} is the value of an income token in period i, and δ_M reflects any discounting of the cash payment after one month. To build intuition, note that a participant playing *RRR* consumes one income token in period 1 as long as no shock occurs in period 1 (probability $(1-\theta)$). The participant consumes one income token in period 2 as long as there are no shocks in periods 1 and 2 (probability $(1-\theta)^2$). If a shock occurs in period 1, the participant insures herself using flexibility, but then needs to make a double repayment in period 2, and thus consumes zero in this period. If there is a shock in period 2 (but not in period 1), the participant uses flexibility to insure herself in period 2, but again consumes zero in this period. In period 3, the participant consumes one income token if there is no shock in periods 2 and 3 (probability $(1-\theta)^2$) – period 1 shocks are irrelevant since these are insured. Note that a shock in period 2 (but not in period 1) would lead to a double repayment in period 3, and hence to zero consumption. Finally, if there are shocks in *both* periods 1 and 2, the loan is in default. In the case of contract default (group default in joint liability treatments), we assume that participants will not continue to make repayments. Hence, assuming there is no shock in period 3 (probability $\theta^2(1-\theta)$), the participant consumes *two* income tokens in this period. The probability of receiving the continuation value V is $(1-\theta)^2$ that no shocks arrive in periods 2 or 3, plus $(1-\theta)^2\theta$ that there is a shock in period 2, but not in periods 1 and 3.

Expected payouts for other strategies are obtained analogously. As outlined in Section 2.1, we induced discounting across periods in the lab using income tokens with declining value: $R^{t1}=40$, $R^{t2}=30$, $R^{t3}=20$. Table D.16 presents payouts using different values of δ_M (any payouts cited in the main text use $\delta_M=1$). For simplicity, we show payouts for the joint liability treatments for the cases that the partner plays either *RRR* or *DDD*.

Interpretation

In all four treatments, *RRR* is monetarily less beneficial than *DDD*. In addition, the expected payoff from a given strategy is similar across treatments: The expected payoff from *DDD* is the same in *IL*, *IL-flex*, and in the two *JL* treatments in the case the partner also plays *DDD*. In case the partner plays *RRR*, *DDD* is more beneficial in the *JL* treatments due to the possibility of free-riding on the partner. A key difference is that *RRR* yields a higher payoff in *IL-flex* than in *IL* due to the possibility of self-insuring shocks. In contrast, *RRR* yields a *lower* payoff in *JL-flex* than in *JL*: This is driven by our assumption that borrowers use self-insurance rather than mutual insurance to insure shocks. Even though self-insurance is more costly given the need for a double repayment, it is almost uniformly preferred by our participants (Section 4.2), consistent with a desire to avoid norm-related penalties. Finally, expected payoffs for *FDD* are the same as for *DDD*, both in *IL-flex* and *JL-flex*, independent of the partner's play.

Overall, the payoff structures in Table D.16 do not explain the treatment differences that we find. For example, more participants repay in *IL* than in *IL-flex*, despite the expected payoffs from default being the same and the expected payoffs from repaying being *higher* in *IL-flex* due to the insurance value of flexibility. We conclude that our findings cannot be explained by payoff maximization.

E Expected Punishments

Calculation

The expected punishment for a strategy (see Appendix D) is the average punishment over all one-period actions, where each action is weighted by the probability that the action is applicable. The calculations also take into account that punishment was paid out for one random period. In *JL*, we

Table D.16: Expected Payouts

Strategy	One-month discount factor for V		
	1	0.8	0.6
PANEL A: INDIVIDUAL LIABILITY TREATMENTS			
IL			
RRR	122	113	60
DDD	135	135	135
IL-Flex			
RRR	130	116	102
FR_2R	135	124	113
RFR_2	139	125	111
DDD	135	135	135
FDD	135	135	135
PANEL B: JOINT LIABILITY TREATMENTS (PARTNER PLAYS RRR)			
JL			
RRR	137	121	104
DDD	177	169	160
JL-Flex			
RRR	129	115	101
FR_2R	137	125	113
RFR_2	124	113	102
DDD	177	169	160
FDD	135	135	135
PANEL C: JOINT LIABILITY TREATMENTS (PARTNER PLAYS DDD)			
JL			
RRR	67	58	50
DDD	135	135	135
JL-Flex			
RRR	67	58	50
FR_2R	135	135	135
RFR_2	75	75	75
DDD	135	135	135
FDD	135	135	135

In treatments with flexibility, all repayment strategies assume that flexibility is used in case of shocks (motivated by our experimental result that self-insurance is preferred to mutual insurance). Calculations for all strategies assume that the agent stops repaying after certain (group) default.

calculate expected punishment $E[P]$ of strategy S containing actions a_{it} in case of no shock in period t and a_{jt} in case of a shock as follows:

$$E[P(S)] = \frac{1}{3}\sum_{t=1}^{3}(1-\theta)P(a_{it}) + \theta P(a_{jt}) \tag{A13}$$

where $P(a)$ denotes the sample mean punishment of action a in the respective treatment, t the respective period, and θ the shock probability ($\theta = 0.25$ in our experiment). Given our joint-liability setting with mutual insurance, we assume that an individual resumes her initial strategy after a shock (i.e., someone playing RRR resumes repayment after shocks). Note that in *JL*, $a_{it}=R$ for all t in the strategy RRR and $a_{it}=D$ for all t in the strategy DDD. In both cases, $a_{jt}=$*don't repay (shock)* as the individual is not able to repay, so the calculation simplifies to

$$E[P(S)]=(1-\theta)P(a_{it})+\theta P(a_{jt}) \tag{A14}$$

In *JL-flex*, flexibility gives rise to additional strategies, both in case of a shock and without. For all strategies that involve repayment and/or the use of flexibility *(RRR, FR_2R, RFR_2,* and *FDD)*, we assume that flexibility would be used in case of a shock, so $a_{jt}=$ *use flex (shock)* if flexibility is available. This implies that we need to consider additional actions triggered by the use of flexibility (i.e., repay $2R$ or default on $2R$). For example, $E[P(FR_2R)]$:

$$\begin{aligned} E[P(FR_2R)]=[&\underbrace{(1-\theta)P(\textit{misuse flex})+\theta P(\textit{use flex (shock)})}_{\text{Period 1}} \\ &+\underbrace{(1-\theta)P(\textit{repay double})+\theta P(\textit{don't repay double (shock)})}_{\text{Period 2}} \\ &+\underbrace{(1-\theta)P(\textit{repay})+\theta P(\textit{don't repay (shock)})}_{\text{Period 3}}]/3 \end{aligned} \tag{A15}$$

For *DDD* in *JL-flex*, we assume $a_{jt}=$*don't use flex (shock)*, so the calculation is analog to *DDD* in *JL*.

Results

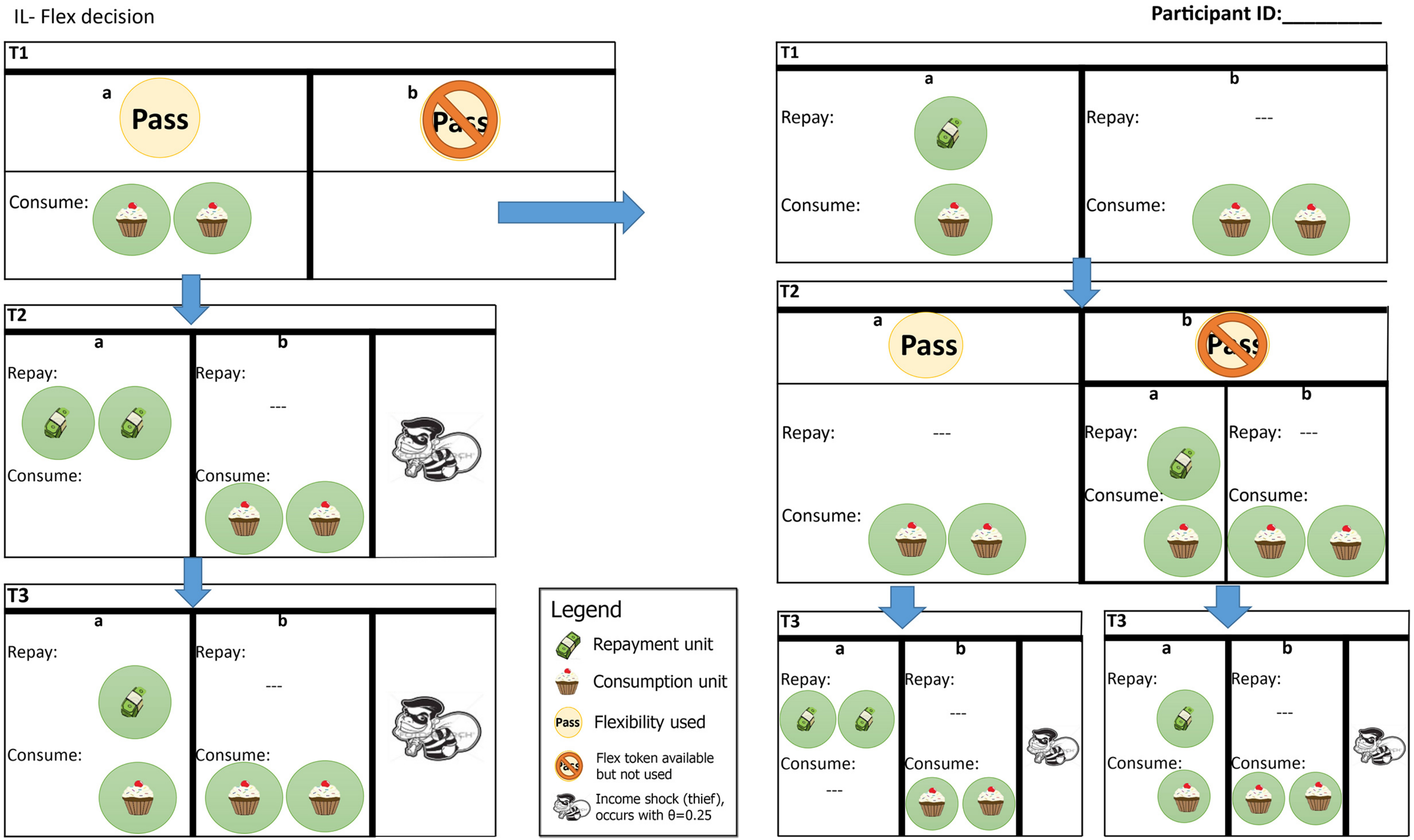
Figure 5 shows the expected punishment for key strategies. Consistent with the intuition from the single-period punishment discussion, we find that flexibility reduces excessive punishment for those who choose to repay: Participants can self-insure income shocks, and thus suffer fewer penalties from having to rely on their partner. The expected punishment for *RRR* is reduced by 30 percent, though this difference is not statistically significant ($p=0.107$, Wilcoxon signed-rank test comparing *RRR* in *JL* vs. *JL-flex*). Misusing flexibility comes at a cost, which increases in its foregone insurance value: Misuse in period 1 is punished 21 percent more than misuse in period 2 ($p<0.001$, Wilcoxon signed-rank test comparing FR_2R vs. RFR_2). Finally, we observe that flexibility can reduce the expected punishment for strategic default on the overall loan: Borrowers who plan to default can dodge 15 percent of the punishment by first using flexibility to postpone repayments ($p<0.001$, Wilcoxon signed-rank test comparing *DDD* vs. *FDD* within *JL-flex*).

F Details on Elicitation of Choices

Flexibility As with all other choices, we elicited decisions on the use of flexibility using the strategy method. First, subjects were asked whether they wanted to use their pass token in period 1 when there is no shock. If they decided to use it, they were subsequently asked about their repayment decisions in period 2 and 3. If they chose not to use it, they had the choice to use their pass token in period 2: first, when there is no shock and second, in case of a shock in period 2. Last, all participants decided whether to use their pass token in period 1 in case of a shock. This procedure lays out nearly the complete strategy on flexibility use.

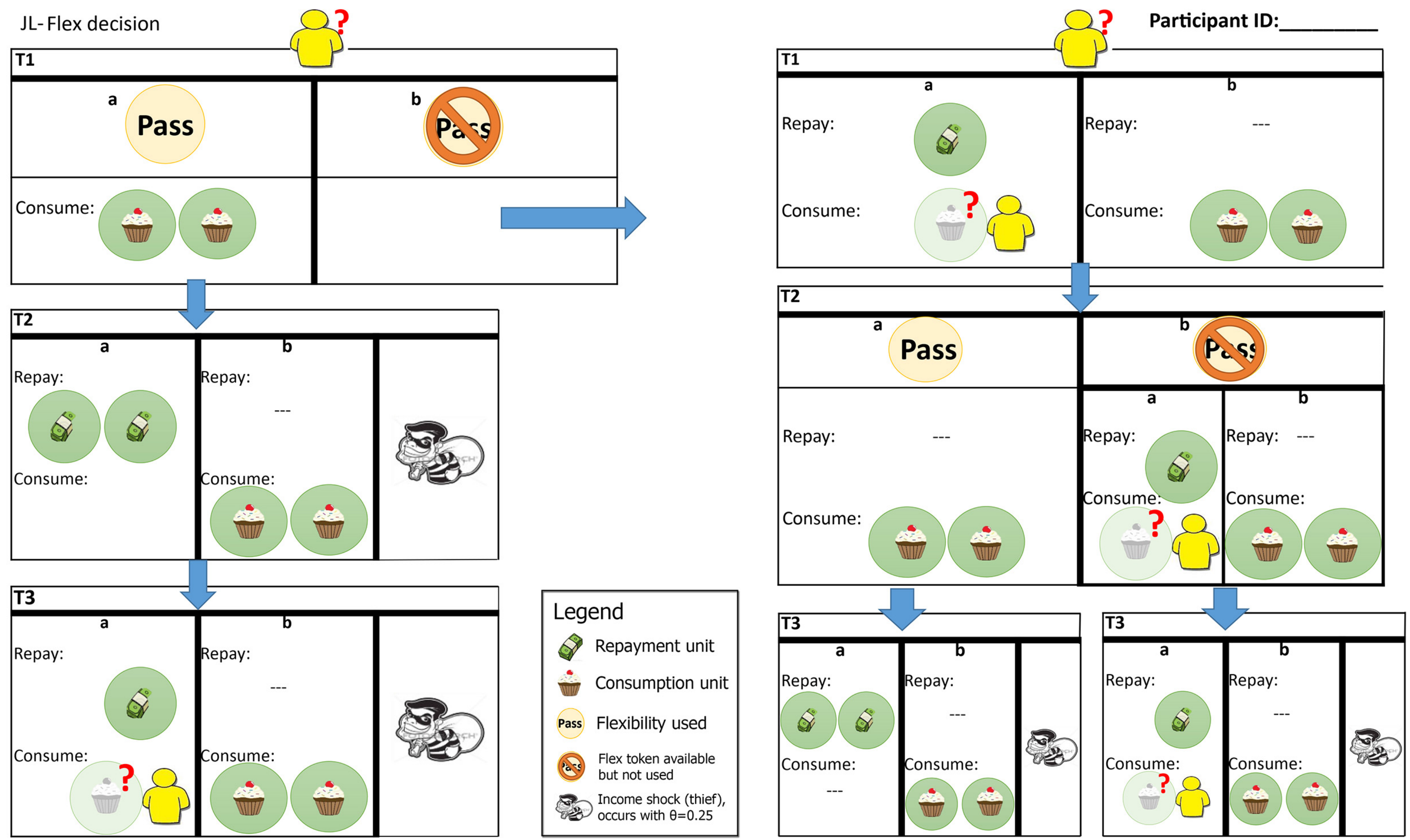
Punishment In total, we elicited punishment decisions in *JL-flex* for eight different choices of the partner: repayment, non-repayment, flexibility use in case of a shock, misuse of flexibility, double repayment (repaying flexibility), non-repayment of flexibility, not using flexibility in case of a shock and non-repayment of flexibility due to a shock.

Figure F.4: Decisions in IL-flex (no-shock path)



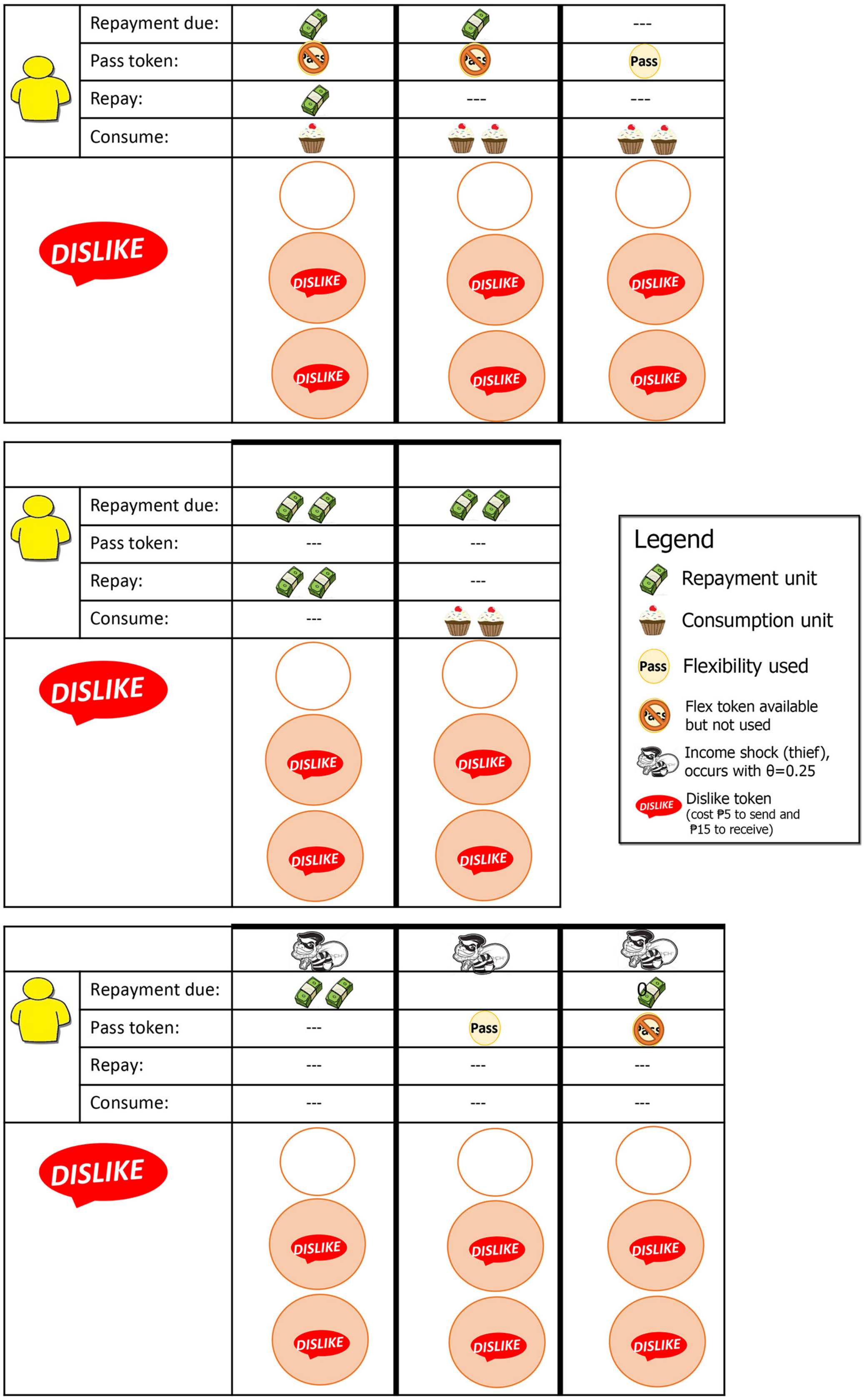
Notes: This figure shows one of the decision sheets from the *IL-flex* treatment, following the path where shocks are possible ex-ante but do not arrive ex-post. Choices following shocks are elicited on additional decision sheets. The borrower starts in the top left panel, and in the case that no shocks arrive in T1. If she decides to use her pass token in T1, she consumes 2 in T1 and moves to T2 in the middle left panel. In T2, a shock arrives with $\theta = 0.25$. If no shock arrives in T2, she decides between repaying the pass token with $2R$, and consuming 2. In T3, a shock arrives with $\theta = 0.25$. If no shock arrives in T3, she decides between repaying 1 (and consuming 1), and consuming 2. Note that repayment after default is possible but dominated. Alternatively, the borrower may choose not to use her pass token in T1, in which case she moves to the right decision panels as shown. Decisions are repeated for other shock realizations following the strategy method (see Section 2.2 for procedures).

Figure F.5: Decisions in JL-flex (no-shock path)



Notes: This figure shows one of the decision sheets from the *JL-flex* treatment, following the path where shocks are possible ex-ante but do not arrive ex-post. Decisions are analog to those described in Figure F.4, except that repayment may now cost either one or two tokens, depending on the (unknown) choice of an anonymous partner. See Section 2.2 for procedures.

Figure F.6: Punishment Decisions in JL-flex



Notes: This figure shows the decision sheet for punishments in the *JL-flex* treatment. Each column shows a possible behavior of the partner and whether a shock has arrived, and asks the participant to assign zero, one, or two dislike tokens. Each token costs five pesos to send and 15 pesos to receive. For instance, the first column corresponds to the punishment for 'Repays' in the *JL-flex* panel of Figure 4. The second column corresponds to 'Does not repay', the third column to 'Misuses flex', and so on. For the randomly selected payoff-relevant treatment, punishment is paid out for a randomly selected period within this treatment (Section 2.2).

G Experimental Setup

Income tokens earned in the microcredit game could be traded for items from the consumption table containing a variety of products such as sweets, food staples, household items and beauty products, offered at typical market prices. Participants were encouraged to familiarize themselves with the items before the start of the session with the help of a consumption catalog that displayed all items and their value, and all items were visible throughout the session.

Figure G.7: Consumption Table

Notes: Consumption table displaying items for in-kind experimental payouts: income tokens earned in the microcredit game could be traded for these items.

H Norm Elicitation: Vignettes

We will now conduct a study. In the following we will describe various decision-making situations and we will portray the behavior of Maria. Think of Maria who is a fictitious member in a nearby ASHI center, similar to yours. Maria took a loan to run a small internet shop with four old computers. At her shop, people can pay to use a computer and go online.

Scenario 1: no flexibility, week 3 Maria has a 25-week loan. It's the third week of her loan cycle [SHOW SCHEDULE]. In weeks 1 and 2, she has repaid. Today is the next center meeting and her third repayment is due.

1. Maria repays her loan installment. As far as you know, her business is going ok and she does not have any unexpected financial difficulties at the moment.
2. Maria does not repay her loan installment. You know that there were many power cuts this week, and she could not earn much from her internet shop because her computers don't work without power (she does not have a generator).
3. Maria does not repay her loan installment. As far as you know, her business is going ok and she does not have any unexpected financial difficulties at the moment.

Scenario 2: no flexibility, week 4 Now let's look at the next week. Remember that Maria has repaid in weeks 1 and 2. Let's look at the following situation: last week, as far as you know Maria's business was going ok and she did not suffer from any emergency. However, she did not repay her loan installment. Today is the center meeting and Maria's fourth repayment installment is due.

1. Maria repays her loan installment. As far as you know, her business is going ok and she does not have any unexpected financial difficulties at the moment.
2. Maria does not repay her loan installment. You know that there were many power cuts this week, and she could not earn much from her internet shop because her computers don't work without power.
3. Maria does not repay her loan installment. As far as you know, her business is going ok and she does not have any unexpected financial difficulties at the moment.
4. Suppose Maria has not repaid either in weeks 3 or in week 4. As far as you know, her business is going ok and she does not have any unexpected financial difficulties at the moment. How socially appropriate do you believe Maria's behavior to be, when you *jointly* consider her behavior in weeks 3 and 4?

Scenario 3: flexibility, week 3 Now imagine that Maria has a new loan product, the 'pass token' [SHOW SCHEDULE]. The pass token schedule means that she has to repay a weekly installment every week, but you get 2 pass tokens with which you can pass the repayment amount due in this week to the next week. The following week, you have to pay a double repayment installment. For example, in some weeks you may have problems repaying your installment because someone in your family gets sick, your children have school activities, or you had a bad business week, etc. The pass token allows you to pass your week's repayment installment to the next period without defaulting to ASHI. ASHI does not check the reasons for you using the pass token so you can use it whenever you want. The only restriction is that you only have 2 pass tokens in total, and you cannot use the two tokens in a row. You have to repay the first before the second token can be used.
Again, Maria has a 25-week loan and this time, she has two pass tokens. It's the third week of her loan cycle [SHOW SCHEDULE]. In weeks 1 and 2, she has repaid. Today is the next center meeting and her third repayment installment is due.

1. Maria repays her loan installment. As far as you know, her business is going ok and she does not have any unexpected financial difficulties at the moment.
2. Maria uses one of her pass tokens, which postpones this week's installment to next week. You know that there were many power cuts this week, and she could not earn much from her internet shop because her computers don't work without power.
3. Maria uses one of her pass tokens, which postpones this week's installment to next week. As far as you know, her business is going ok and she does not have any unexpected financial difficulties at the moment.
4. Maria does not repay her loan installment. As far as you know her business is going ok and she does not have any unexpected financial difficulties at the moment. She does not use one of her pass tokens.
5. Maria does not repay her loan installment. You know that there were many power cuts this week, and she could not earn much from her internet shop because her computers don't work without power. She does not use one of her pass tokens.

Scenario 4: flexibility, week 4 Now we are in the next week. Remember that Maria has repaid in weeks 1 and 2. Imagine that last week, her business was going ok, she had no unexpected financial difficulties and that she used her pass token. Today is the fourth center meeting and her third and fourth installments are due, that is two installments in total.

1. Maria repays the two installments. As far as you know, her business is going ok and she does

not have any unexpected financial difficulties at the moment.

2. Maria does not repay the two installments. You know that there were many power cuts *this week*, and she could not earn much from her internet shop because her computers don't work without power.
3. Maria does not repay the two installments. As far as you know, her business is going ok and she does not have any unexpected financial difficulties at the moment.
4. Suppose Maria has used her pass in week 3, and does not repay the two installments in week 4. As far as you know, her business is going ok and she does not have any unexpected financial difficulties at the moment. How socially appropriate do you believe Maria's behavior to be, when you *jointly* consider her behavior in weeks 3 and 4?