Peer Effects and Loan Repayment: Evidence from the Krishna Default Crisis*

Emily Breza†

Abstract

I estimate loan repayment peer effects by analyzing a natural experiment during which 100% of borrowers temporarily defaulted on their microloans. Because the defaults occurred simultaneously, the timing of the shock induced variation in both individual- and peer-level repayment incentives. Using variation in the peer group's incentives to instrument for peer repayment, I find that if a borrower's peers all repay, she is 10-15pp more likely to repay. I estimate the benefit of peer effects to the lender using a dynamic discrete choice model and find that overall, the peer effects improve repayment rates and profits relative to a counterfactual.

1 Introduction

Group lending has always been one of the hallmarks of microfinance. In order to overcome weak contracting institutions, limited borrower wealth and collateral, and the inability for microentrepreneurs to transfer control rights to creditors, microfinance contracts have traditionally relied on dynamic incentives and social capital to provide repayment incentives.[1] The Grameen Bank website claims "there is more to the bank than just the balance sheet; it ties lending to a process of social engineering."[2] The peer lending context has been exported

*I thank Esther Duflo, Abhijit Banerjee, Stephen Ryan, Robert Townsend and Sendhil Mullainathan for encouragement and detailed discussion. Also Prathap Kasina for outstanding research assistance and the George and Obie Shultz Fund and the NSF GRFP for financial support. I owe deep gratitude to Padmaja Reddy and Spandana. I also thank seminar and conference participants at Columbia, Duke, Harvard, LBS Corporate Finance Symposium, NEUDC, Northwestern, Notre Dame, Stanford, Tufts, Washington University St. Louis, USC, and Yale for their insightful comments.

†Columbia Business School, Division of Finance and Economics. Email: ebreza@columbia.edu

[1]See Section 3.1 for a discussion of the limitations of dynamic incentives as an enforcement device.

[2]see http://www.grameen-info.org

and replicated across the globe and has remained a key investment for microfinance institutions (MFIs). While microlenders have typically enjoyed very high repayment rates, we know surprisingly little about how the social capital developed during the lending process actually affects a borrower's repayment decision. Furthermore, as large-scale coordinated defaults have begun to creep into the microfinance sector, it is more important than ever to understand if microfinance's "social engineering" stabilizes or exacerbates a crisis.

I examine microfinance peer effects following a large-scale default episode that took place in the Krishna District of Andhra Pradesh, India on March 9, 2006. In order to promote his own financial inclusion agenda, the District Collector[3] announced that his constituents should stop repaying their microloans. Within two days of the announcement, all borrowers had ceased making installment payments. Soon after the defaults, the local microlenders began to reestablish collections in the affected villages, suspended the joint liability feature of the loans, and again offered new loans to repayers. Some individuals resumed payment within a few months of the crisis, and as of November 2009, 40-50% of individuals had fully repaid their liabilities. I investigate whether peer effects helped or hindered collections.

While a mass default might seem like a special case in which to look for peer effects, quantifying negative risks during crises is key to understanding the value and long-run viability of MFIs. Markets for securitized microloans continue to grow, and any positive or negative repayment peer effect should affect both the pricing of such securities and the borrowing costs faced by MFIs themselves. Because these borrowing costs are passed through to customers through interest rates, peer effects could play an important role in affecting the cost of capital for poor individuals across the globe. The role of peer effects is also relevant for shaping government policies for financial inclusion.[4]

The link between social effects and financial behavior is related to the broader question of asset correlations during financial crises. Peer effects appear in other crisis settings such as bank runs (Iyer and Puri (2012), Kelly and Gráda (2000), and Bond and Rai (2009)) and the US housing market. Guiso et al. (2013) find that social norms in a community affect an individual's decision to default strategically on their mortgages. Microfinance is a prime candidate for repayment peer effects because so much emphasis on social capital is built into the loan mechanics. Further, as in housing, strategic default is highly public, and individuals may derive private benefits from their microfinance borrowing groups.

Historically, many group lending schemes have been characterized by group-level joint

[3] The District Collector is a federal bureaucrat who is the head of the district's administration.

[4] In India for example, some new government-sponsored financial access models, such as Business Correspondent banking, do not include peer components.

liability. In these contract structures, there is a direct channel for peer decisions to affect repayment rates. For example, if one member defaults on her loan, then the remaining members must bear the cost of that default if they intend to receive new loans in the future. The non-defaulting borrowers could use a local enforcement technology to coax the defaulter into repaying. Alternately, this extra cost may result in other repayers choosing to default and walk away from the lending relationship. In both scenarios, the actions of the peer group have direct consequences for an individual's own repayment decisions. Several theoretical models examine mechanisms through which joint liability might operate. Candidate pathways include moral hazard in project selection (Stiglitz (1990)), moral hazard in project effort (Banerjee et al. (1994)), adverse selection of borrowers (Ghatak and Guinnane (1999) and Ghatak (1999)), and village sanctions and limited contract enforcement (Besley and Coate (1995)). These models make different predictions for borrower repayment, but all predict that peer behaviors should affect individual decisions. Ahlin and Townsend (2007) and Karlan (2007) investigate empirically the workings of joint liability in microfinance using data from Thailand and Peru, respectively. The Besley and Coate (1995) model of strategic default is the most relevant to my empirical setting and highlights the potential for both virtuous and perverse social repayment equilibria. Giné et al. (2011) find evidence for perverse joint liability effects during a default episode in Karnataka, India.

While the joint liability literature gives a rich theoretical framework for thinking about peer effects in lending, contract structure alone may not fully explain the social effects embedded in microfinance. Following the Grameen II[5] model, many MFIs have abandoned joint liability but have maintained group meetings. A small but growing set of empirical work links social capital and microfinance in the absence of joint liability. Giné and Karlan (2006, 2009) randomize between individual and joint liability loan contracts. While varying the contract structure, they maintain the group format of the repayment meetings and loan disbursements. Over 1- and 3-year horizons, moral hazard does not appear to increase when clients are assigned to the individual liability treatment, default rates do not increase, and the individual liability policy actually attracts more new clients.

In my setting without joint liability, the remaining question is, to what effect does the peer format of the individual liability loans contribute to repayment incentives? Feigenberg et al. (2013) examine social effects in the absence of joint liability. The authors find that individuals randomly assigned to weekly versus monthly repayment schedules form stronger ties with their fellow group members, visit fellow group members more frequently, and exhibit

[5] See "Grameen II," available at www.grameen.com for a description.

more trust.[6] The weekly groups also achieve better repayment. The experimental design does not allow the authors to confirm a direct link between social capital and repayment, but the findings highlight the possibility of such effects. In this paper, I analyze the causal relationship of peer repayment on an individual's own repayment decision in the absence of joint liability. Furthermore, this paper is one of the first pieces of evidence on social effects in response to microfinance defaults.

It is both difficult to find exogenous variation in default and hard to estimate peer effects due to omitted correlated covariates, unobserved correlated shocks, and the reflection problem described by Manski (1993). To circumvent these problems, I exploit the random timing of the default shock and propose a novel identification strategy for estimating social effects in borrowing relationships. Microlenders use the promise of new loans to encourage repayment, so borrowers who were closest to receiving a new loan at the time of the defaults have the biggest potential benefits and lowest costs from repayment once collections restart, and indeed I show that these individuals are the most likely to repay. In the standard microfinance contract, borrowers make 50 weekly installment payments, so a borrower's location in this 50-week credit cycle affects her own repayment incentives. Because borrowers have staggered loan disbursements within the peer group,[7] these cyclical repayment incentives also induce variation in the peer group's overall repayment incentives. Peer groups with a majority of borrowers in weeks 45-50 of their loan cycles have much stronger incentives than peer groups with a majority of borrowers in weeks 0-5. Thus, the 50-week loan cycle provides separate sources of variation to identify both "own" and peer incentives. I use variation in the peer group's overall repayment burden to instrument for the fraction of peers who repay, providing consistent estimation of the effect of peer repayment on individual repayment.

I first employ an instrumental variables technique using the average week in cycle to instrument for average peer repayment. The data set comes from Spandana, one of the largest MFIs in India at the time. I control for continuous functions of the total length of time each individual and each aggregate peer group has spent borrowing from Spandana, to eliminate effects from differences in total time spent borrowing from the MFI. Because the discontinuous repayment incentives identify the peer effect, I also apply a regression discontinuity approach to the problem. I compare repayment rates for individuals whose peers have very strong incentives with individuals whose peers have very weak incentives. Since peer groups on either side of the discontinuity look very similar in all regards aside from

[6] Research such as Townsend (1994) has documented the importance of social networks for risk sharing in developing countries.

[7] I define the relevant notion of peer group in section 2.1.

dynamic repayment incentives, the variation is as good as random. The data set contains information on the full universe of loans, including village of residence and borrowing center membership, and can shed light on which level of peer interaction (i.e., social distance) has the largest effect on repayment. Microfinance is unique because the contract structure defines the relevant peer groups.

I find that borrowers are sensitive to their own dynamic repayment incentives, and that each completed week in the loan cycle before the defaults corresponds to a 1pp increase in eventual repayment. I do find that peer behavior influences repayment. Borrowers are 10-15pp more likely to repay their loans if their entire center[8] repays. (The borrowing center is a smaller unit than the village, and all center members attend the same weekly meeting.) The estimates at the village-level peer group are substantially smaller and insignificant and provide suggestive evidence that the peer effect is a largely local phenomenon, fostered by the regular center meetings. Non-linear estimates of the repayment effect show that there are large increases in the probability of repayment if just one peer repays or similarly if all peers repay. Furthermore, I find evidence that the peer effect is asymmetric and is largely a positive force, pulling individuals with weak incentives out of default.

To address more precisely the relationship between peer effects and asset values, I estimate a structural model of repayment, using weekly time variation. I treat the repayment decision as a dynamic discrete game played with fellow center members. Using the methodology of Aguirregabiria and Mira (2007) to deal with the interdependency of players' actions, I estimate the parameters of a simple model of repayment and predict repayment paths under the regime with peer effects and under a counterfactual without peer effects. A structural model is required for two reasons. First, while the reduced form analysis provides estimates for the effects of peer repayment in relative terms, it does not fix the absolute repayment levels under regimes with and without peer effects. Second, pricing the value of the peer effect requires modeling the time structure of each individual's repayment path. Lenders pay an opportunity cost of capital for every additional week borrowers remain in default. Understanding this time structure requires placing more restrictions on the problem.

I find that firm revenues increase by 10% when peer effects are switched on, helping to partially mitigate the costs of default. The greatest increase in potential lender revenues comes from decreasing the time to repayment of those individuals with the highest numbers of payments outstanding. For these individuals with weak "own" repayment incentives, I estimate that peer effects may increase the value of their cash flows by as much as 40%.

[8] Borrowers are assigned to groups of 6-10 individuals. Every 2-5 groups are then combined to form a center. All members of each center meet together each week to make their loan payments.

While peer effects potentially have both positive and negative effects on repayment, the virtuous peer effect is stronger at luring these types of individuals to repay.

The Krishna default crisis has already been repeated on a much larger scale. In October 2010, the government of the state of Andhra Pradesh issued an emergency ordinance severely restricting the operations of all MFIs in the state. The alleged motivations were almost identical to those of the Krishna District Collector: fears of each of over-indebtedness, usurious interest rates, abusive collections practices, and alleged borrower suicides. The results of this study will be helpful in guiding collection efforts in Andhra Pradesh and in informing lenders how to better incorporate the peer forces in microfinance to increase repayment rates during a crisis and to ultimately decrease borrower interest costs.

The body of the paper proceeds as follows. Section 2 describes the natural experiment and data. Section 3 provides a graphical analysis and outlines the intuition behind the identification strategy. Section 4 describes the empirical model, while Section 5 details the results. I estimate a structural model of the loan repayment decision in Section 6. Section 7 discusses potential mechanisms, and Section 8 concludes.

2 Empirical Setting

2.1 Spandana Group Loans

Before describing the defaults in detail, it is necessary to understand the loan product offered by the MFI sharing its repayment data. Spandana Sphoorty Financial Limited was one of the largest MFIs in India and was one of the primary microlenders operating in the Krishna District at the time of the crisis. The standard loan product operates on a 50-week cycle. After loans are disbursed, individuals make 50 equally-sized, weekly installment payments. Upon successful repayment, individuals are given a new, larger 50-week loan. In normal times, defaulters are sanctioned with the denial of future credit.

These loans, typically only offered to women, usually have joint liability provisions. Before the first loan disbursement, each borrower is assigned to a joint liability borrowing group of 5-10 women. Borrowers in a group tend to all be on the same repayment schedule. Every 2-5 borrowing groups within the same village are then combined to form a center. Within a center, groups may have staggered loan disbursements and thus might be at different places in the loan cycle at any time. All borrowers belonging to a center meet each week to make their installment payments to the credit officer. These meetings begin with a joint oath, which affirms the virtues of making on-time payments and helping fellow borrowers. The credit officer then takes public attendance and collects publicly the payments from each member. In normal times, there is also joint liability at the center level between groups. However, this is rarely, if ever, enforced. Following the crisis, all joint liability was suspended.

For the remainder of the paper, I define the peer group as either the borrowing center or the village.[9] Because I use administrative data, I have near complete records of center and village membership. Unlike some peer effects applications, the social format of the lending product leads to a clear definition of the relevant peer groups.

2.2 The Krishna Crisis

The setting is a natural experiment from the Krishna District of Andhra Pradesh, India. On March 9, 2006, the District Collector Navin Mittal closed over 50 branches of two large MFIs, Share and Spandana. This move resulted in the cessation of all weekly repayments across the district, a potential loss of close to Rs 200cr (~$44mm). The Collector alleged that MFIs were setting interest rates too high, using unethical means to make collections, and stealing clients from state banks and SHGs (self help groups). Furthermore, several farmer suicides were blamed on microfinance. The local media[10] began a campaign to highlight the evils of the microfinance industry. There is some mention in local newspapers that Mittal scheduled his announcement around International Women's Day, which occurs on March 8. Sa Dhan, an association representing community development financial institutions, put pressure on the government to rescind the District Collector's statement. A retraction was made in mid-2006 and the worst of the crisis finally came to an end in early 2007.

The District Collector's announcement spurred a diverse set of reactions. Some borrowers started again repaying their loans within a few months of the announcement and picked up where they left off in their loan cycles. In other areas, angry villagers threatened Spandana and Share field staff and forced longer branch closures. As of November 2009, approximately 45% of the outstanding portfolio on March 9, 2006 had been repaid. Understanding these repayment patterns and the role of peer effects is the primary concern of this study.

Due to the political climate, Spandana was forced to take a measured response. While the defaulted loans had been issued under joint liability, Spandana immediately abandoned publicly its enforcement. The MFI also made the decision to reward repayers with future credit regardless of the time spent in default and was able to satisfy all demand for new credit. In effect, the crisis dismantled the strict discipline generally required by an MFI and

[9] I do not consider the borrowing group due to a lack of variation in my instrument within groups.

[10] Many members of the local media also had financial stakes in other types of financial companies. The collector's statement only affected microlenders and benefited competitors.

gave borrowers the option to extend the maturity of their loans at no additional interest cost. After one year of only marginally successful collection efforts, Spandana also started offering refinancing plans, where small new loans were disbursed to encourage borrowers to begin making regular loan payments and to eventually repay all outstanding debt.

The media's treatment of the crisis highlights the controversy surrounding microfinance and the perceived importance of peer interactions. Between March and June 2006, there were frequent negative articles about microfinance in the Vijayawada edition of *The Hindu*. A common view seems to be that "the microfinance companies sanction loans to [groups] liberally without insisting on security but charge exorbitant interest and collect the installments using peer pressure of the group." It is clear that the media thought that the peer enforcement channel mattered in the loan collections. One strongly-worded article notes, "micro-finance institutions have hit upon a new and unscrupulous method of recovering outstanding loans – pitting members of self-help groups against another" The Hindu (2006). Whether the peer effect led to increased long-term default or repayment is an open question.

Spandana's Krishna defaults represent an ideal natural experiment for studying microloan repayment. First, the defaults were instigated by a federal bureaucrat, not through a grassroots movement, and the defaults did not spread across district borders, indicating that true political upheaval did not drive the crisis. Because loan repayment rates remained close to 100% in neighboring districts, it is safe to assume that in the absence of Mittal's announcement, Krishna repayment rates would also have remained near 100%. Moreover, according to the Microfinance Information Exchange, Spandana had less than 0.01% of its portfolio overdue more than 90 days in both 2004 and 2005. In 2008, after the crisis had subsided and crisis loans had been written-off, the reported portfolio at risk was again very low at 0.02%. Second, Spandana was one of the largest, most efficient MFIs in the world and was able to withstand the liquidity shock from the suspension of loan payments on almost 200,000 loans and fulfill its promise of future credit to all repayers. The large ICICI bank owned many of the defaulted loans, further insulating Spandana from liquidity effects. Spandana was able to retain and pay its field staff even when all collections had ceased. Finally, as a result of the crisis, other MFIs decided not to expand their client base into Krishna district. Spandana and Share also agreed not to add new borrowers in the district after the crisis.

2.3 Data

In November, 2009. I visited the district head quarters and 18 branch offices, and Spandana shared their digitized records with me.[11] The data used in the analysis represent a close to complete set of loans outstanding during the crisis and report on loans serviced by all 23 branches operating in the district in 2009.[12] All of Spandana's borrowers during the crisis were women. The data set includes information on group name, center number, and village or slum name as well as details about each loan including size, date of disbursement, cycle, and repayment information. The raw data set contains information on 194,312 loans.

Borrowers are given the chance to take small, interim loans after making many on-time payments. The interim loans also require fixed installments and add to the client's total liability, but do not affect the borrower's loan cycle. Because the main loans are substantially larger than the interim loans, I drop all interim loans and all loans smaller than Rs. 3,000. The data set does not have unique borrower identifiers, so I use fuzzy matching on borrower name, group name, center number, and village name to identify multiple loans.

For the empirical analysis, I also drop all villages with fewer than 50 borrowers. This is for two reasons: First, in the peer effects regressions, villages with only one group or center would be automatically dropped since there is no extra-group variation available. Second, it is possible that these villages have miscoded place names. I further drop any villages without documented cycle numbers, because they are essential for controlling for functions of an individual's total weeks in her lending relationship with Spandana.

Table 1 gives an overview of the final data set used throughout the paper. There are approximately 115,000 unique borrowers included from 574 villages with an average loan size of Rs 7,640 (~$170). This represents 5,340 borrowing centers, with approximately 9 centers per village. The administrative records also include week-specific payment and delinquency information for a subset of borrowers that allows me to determine when a borrower resumed paying her loan and when she completed making payments on the delinquent loan. Of the 115,000 loans in the analysis sample, 57% were still delinquent in November 2009.

Spandana also records the stated purpose for the loans. The three most common purposes constitute half of the loans in the sample. The most common business use is livestock (26.33%), followed by textiles (saree sales, embroidery, tailoring) and small retail shops.

[11] A research team from the Centre for Microfinance (CMF) collected information from those branches I did not visit personally.

[12] Some branches closed after the defaults, but all defaulted loans were transferred to one of the 23 branches.

3 Graphical Analysis

3.1 Repayment Incentives Across the Loan Cycle

The key task in this analysis is to find plausibly exogenous variation in the repayment behavior of each borrower's peer group. Note that each borrower's own incentive to repay changes over the 50-week cycle. Because loan installments are all the same size and are made weekly, the cost of paying off the remainder of the loan is decreasing as the loan cycle progresses (i.e. as borrowers approach the loan's maturity date). Additionally, MFIs almost universally use dynamic incentives to encourage repayment. Lenders motivate borrowers with the promise of new, often larger amounts of credit. This is true for Spandana even after the defaults, as borrowers who eventually repaid were all offered new loans. As the weeks in the loan cycle progress, the borrower is closer to receiving the next loan disbursement. Hence with discounting, the costs of repaying the remaining loan burden are decreasing and the benefits of paying off the loan in full are increasing. Therefore, repayment incentives are strongest in week 49 and weakest in week 0 of each loan cycle.

Throughout the paper, I rely on the idea that dynamic incentives provide differential repayment motivations depending on where agents fall in their loan cycles at the time of the default. Namely, agents are increasingly motivated to repay when the next loan disbursement is expected in a shorter number of weeks. It is important to note, however, that basic dynamic incentives are not sufficient to provide repayment incentives, even without exogenous default forces. Bulow and Rogoff (1989) make this point with respect to sovereign debt. My identification assumptions do not require dynamic incentives to be the only driver of loan repayment, but do require that agents prefer to repay if they are closer to receiving a new loan. There could be a range of additional motivations that lead to pristine loan repayment records in the absence of large-scale default episodes.[13]

Figure 1 shows stylized (dashed curves) and observed (scatter plot) repayment behavior as a function of the week in the loan cycle. The dashed curves in Figure 1 plot a back-of-the-envelope, hypothetical discounted present value calculation of a borrower's future cash flows from the MFI as a function of the weeks since taking the first loan. There are three simple ingredients to this calculation: outflows (installment payments), inflows (new loans), and a

[13] For example, microlenders could provide other services that borrowers value such as low-cost inputs or technical assistance. Defaulting on loans would result in the denial of both credit and these other non-credit sources of utility. Alternatively, it has been shown that microfinance can serve as a commitment device that also provides agents with utility above and beyond the value of the loans (see Basu (2011) and Fischer and Ghatak (2010)).

discount rate (assumed to be the annualized interest rate paid on microloans).[14] The simple NPV calculation suggests increasing value to the borrower as the week in loan approaches 50, with a sharp discontinuity upon disbursement of a new loan.

The relationship between the week in the loan cycle when the crisis occurred and actual loan repayment following the crisis is displayed in the scatter plot in Figure 1. Each point represents the average full repayment[15] rate as of November 2009[16] across all borrowers in each week with the MFI. The points between 0 and 50 weeks correspond to borrowers in their first loan cycles, while the points between weeks 50-100 correspond to second loans. Note that the actual repayment pattern is quite striking and follows the overall shapes of the stylized NPV curves, with sharp discontinuities at multiples of 50 weeks. Borrowers in the beginning of their loan cycles tend to repay with less than 20% likelihood, while borrowers at the end repay with upwards of 60% likelihood. Note that this pattern looks similar across all three loan cycles in the data. Because the Krishna defaults all occurred within days of each other and because the timing of the initial loan disbursements was staggered, Mittal's announcement induced variation in the repayment incentives of borrowers across the district. Week in the loan cycle is a good candidate for quasi-exogenous variation in repayment.

This modulo 50 weeks argument can be extended to the peers of a borrower. Holding an individual's loan characteristics fixed, we can ask what happens if her peers fall on one side of the discontinuity or the other. The experiment that the exercise most closely mimics is selectively writing off loans and measuring the impact on peer repayment.

Some assumptions are required for identification. First, conditional on observables and holding the individual constant, peer groups that fall on one side or the other of each 50-week threshold must not be systematically different. First, I know when each borrower and overall peer group started taking loans from the MFI, so I can control for smooth functions of this timing variable. A similar argument might be made for loan size. Again, I include controls for functions of loan size and use the remaining variation in timing for identification. A second concern would be that the district collector timed his statement to coincide with the loan cycles of key constituents. A required assumption is that the timing of the announcement was not related to any of the cyclical timings of specific types of borrowers. This is unlikely to be violated because Spandana was only one of several MFIs in the area, which all had

[14] To generate this picture, I assume that the loan size grows faster than the interest rate and that borrowers take loans for 8 loan cycles. However, the saw-toothed pattern is not specific to these assumptions.

[15] For the bulk of the analysis, repayment is an indicator for the individual repaying the entire loan.

[16] November, 2009, which is three years after the resolution of the collector's statement, is an appropriate time at which to separate long term repayers from long term defaulters. The Spandana staff predicted that it might be possible to collect at maximum 10% of the remaining debt outstanding in subsequent months.

different expansion paths across the district. Also, Spandana borrowers were in a range of loan cycles at the time of the defaults, making it hard to privilege any specific group.

It is important that borrowing groups in weeks 45-50 of their cycle 1 loans not be different from borrowing groups in weeks 0-5 of their cycle 2 loans. Figure 2 plots the distribution of the number of groups by the borrowing group's week with Spandana on the date of the defaults. The scatter plot shows the raw data, while the solid lines are local linear regressions, run separately for borrowers in cycles 1, 2 and 3 following Imbens and Kalyanaraman (2012). The largest concentration of loans occurs between 40-50 weeks before the defaults, but there is no significant discontinuity at week 50.[17] It is likely that the large peak in lending activity is due to Spandana's continuous loan expansion pattern. There is no detectable difference in the number of loans at the week 100 discontinuity either. In the Online Appendix, I present similar RD pictures for the most common loan purposes. Due to sparsity in the data, I restrict the analysis to only the threshold at week 50. There is no significant difference in any loan purpose across this discontinuity.

But why did borrowers react so strongly to the collector's statement? Mittal undermined the repayment discipline characteristic to microfinance and allowed clients to both choose when to make payments and to borrow indefinitely at an effective 0% interest rate. I analyze the extent to which persistent peer effects contributed to the 45%-50% repayment rates. If the peer decisions did play a role, was the peer effect on net virtuous or vicious for repayment? The peer channel may encourage either speedy repayment or extended default.

3.2 Preliminary Peer Effect Evidence

Figure 3 provides evidence that there might be repayment peer effects, at least at the center peer group level. The graph plots the probability mass functions (PMFs) of the fraction of repayers across all villages in Panel A and all borrowing centers in Panel B. The resulting village distribution is close to single peaked, with the largest fraction of villages experiencing repayment rates around 50% - 60% and some small excess mass at 0. However, the center-level distribution has a distinctly different shape. There are large spikes in mass at 0 and 1, representing full default and repayment.[18] Few villages, however, have close to universal default or universal repayment. These density plots give preliminary evidence of repayment peer effects at the center level. The striking difference in the repayment profile over villages versus centers also suggests that social mechanisms might be stronger at closer social dis-

[17] The size of the discontinuity using Imbens and Kalyanaraman (2012) is -4.7. The standard error is 10.

[18] The distribution of incentives across the centers is much more smooth and evenly spread out across the 50 weeks in the cycle. See Figure 8 in the Online Appendix.

tances. However, other sources of correlation within peer groups might also be responsible for these patterns, so it is necessary to use the quasi-random variation induced by the timing of the shock to say something more definitive.

Finally, Figure 4 presents even stronger evidence of a repayment peer effect. The figure plots a local linear approximation of the individual's repayment likelihood as a function of the modal borrowing center week in cycle. The relationship is non-parametrically estimated within each loan cycle (i.e. 0-50 weeks, 50-100 weeks, and 100+ weeks) using the optimal RD bandwidth selection of Imbens and Kalyanaraman (2012). Functions of the individual's and center's week in the loan cycle, week with the MFI, and loan size are partialled out of the repayment variable. The picture indicates that the likelihood of repayment jumps downward by approximately 10pp as the center's modal number of weeks crosses from 49 to 51 weeks in the loan cycle and from 99 to 101 weeks. This jump is also statistically distinguishable from zero at standard levels. In other words, borrowers are 10pp more likely to repay if their peers switch from very bad to very good incentives. There is another similar, but much noisier jump at the 100 weeks discontinuity. The discontinuities in the figure suggest quite substantial peer effects in loan repayment, which I further investigate in Sections 4 and 5.

4 Empirical Strategy

Before a further discussion of the results, it is important to understand the statistical inference problem at hand. The causal equation of interest is

$$repay_i = \beta_0 + \beta_1 repay_{p(i)} + \beta_2 weeks_i + \beta_3 X_{i,p(i)} + \varepsilon_{i,p(i)} \tag{4.1}$$

where i indexes the individual and $p(i)$ indexes the peer group net of the individual. The variable, $repay_i$ is a measure of individual i's loan repayment, $repay_{p(i)}$ is a measure of repayment by i's peers, and $X_{i,p(i)}$ is a set of additional individual and peer-level controls.[19]

The biggest problem confronting most peer effects identification strategies is that the peer effect cannot be separated from correlated shocks or other omitted group-level characteristics using an OLS framework. The coefficient of an OLS regression would also contain the "reflection" of any peer effect through the peer group (see Manski (1993)).

Peer effects questions arise frequently in the labor and development literatures, and researchers have developed several approaches to consistently identify the setting-specific equivalents of equation 4.1. One set of approaches involves separating the peer group from

[19] For most of the analysis, $repay_i$ is an indicator for full repayment as of November 2009.

the common shock group.[20] Others have tried an IV approach to solve the identification problem.[21] My identification strategy uses the week in the loan cycle at the time of default conditional on other observables as an instrument for individual and peer repayment incentives. An important assumption is that the timing of Mittal's announcement was as good as exogenous. This assumption seems plausible for the reasons previously discussed.

Table 2 compares the average week in the loan cycle at the time of default with village characteristics from the Census of India. Ideally, village characteristics would not be correlated with the distribution of weeks remaining in the loan cycle. The Indian Census has information on population, education facilities, medical facilities, access to taps or wells, communication facilities, banking facilities and roads. The table only includes information for rural villages and may be incomplete due to difficulties in matching village names between the Spandana data and the Census. However, for the sample of matched villages, the average weeks variable does not seem to be related to many of the covariates available in the census that capture demographic information, land area, access to finance, access to education, and access to health care in a systematic way.

The first column examines the relationship between various village covariates and the average village week in the loan cycle. Each regression coefficient comes from a separate univariate regression. Most of the variables are not correlated with weeks in the cycle. As the identification is really coming from individuals early in the cycle versus late in the cycle, I also show a different specification in columns 2-4. These columns display results of regression of village characteristics on the fraction of the village between weeks 0 and 5 and between weeks 45 and 50 in the loan cycle. Column 4 shows whether the coefficients are significantly different. While distance to town is overall decreasing in the average weeks in the loan cycle variable, there are no significant differences between the villages with more individuals in the extreme early versus extreme late part of the cycle. The only variables that are significantly different are access to medical facilities and access to a railway. While villages with more early cycle borrowers are more likely to have medical facilities, they are less likely to have rail access. Further per capita medical facilities are not significantly different. To check that these characteristics are not driving my conclusions, I find that including village fixed effects

[20] See Conley and Udry (2010)'s treatment of social learning among farmers in Ghana and Munshi and Myaux (2006)'s analysis of social views towards contraception in Bangladesh.

[21] For example, Duflo and Saez (2002) instrument peer behavior with average group characteristics when analyzing social effects on 401K contributions. In studying peer effects in education, Acemoglu and Angrist (2000) use separate instruments for own education and for the education of an individual's peers. In order to estimate education externalities across age cohorts in Indonesia, Duflo (2004) finds an instrument for peer education that is orthogonal to the individual's own educational attainment.

in the center-level peer effects regressions does not change substantively the results.

In the estimation procedure, the implicit first stage for each individual is

$$repay_j = \gamma_0 + \gamma_1 weeks_j + \gamma_2 X_{j,p(j)} + e_{j,p(j)} \tag{4.2}$$

where $repay_j$ is a measure of repayment for individual j. The regressor, $weeks_j$, is the number of weeks elapsed in the 50-week loan cycle at the time of default for individual j. The vector $X_{j,p(j)}$ contains individual and peer group controls including polynomials of loan amount and polynomials of the weeks elapsed since taking the first loan from the MFI. As loan size and overall tenure with the MFI may be endogenous, identification comes from the within-loan cycle timing variable, conditional on loan size and on functions of overall MFI tenure. Recall that loan cycles only last 50 weeks then restart. Those individuals who are early in their second loan cycles should have similar characteristics (i.e. continuous, not discrete differences) to individuals late in the first loan cycle.

My goal is to estimate the causal relationship in Equation 4.1. In the main specification, I assume the peer effect to be linear in means,[22] with $repay_{p(i)} = \sum_{j \in p(i), j \neq i} \frac{repay_j}{N_{p(i)}}$. $N_{p(i)}$ denotes the size of person i's peer group, $p(i)$, excluding person i. Here, the endogenous peer repayment term can be instrumented using the average weeks in the loan cycle conditional on average village loan size and starting dates with Spandana.[23] Let $weeks_{p(i)} = \sum_{j \in p(i), j \neq i} \frac{weeks_j}{N_{p(i)}}$ be the average weeks with the MFI of the peer group. The key requirement for identification is that $weeks_i \perp weeks_{p(i)} | X_{i,p(i)}$. Because of this requirement, all peer group characteristics are calculated excluding the individual's own borrowing group. So, $p(i)$ is always either defined as village ex group or center ex group. The orthogonality requirement is plausible conditional on functions of the group's tenure with the MFI and loan sizes.[24]

The first stage aggregated to the peer group level is

$$repay_{p(i)} = \delta_0 + \delta_1 weeks_{p(i)} + \delta_2 X_{i,p(i)} + \eta_{p(i)} \tag{4.3}$$

where the vector $X_{i,p(i)}$ includes the appropriate individual- and peer-level controls. Because I assume that the functional form of the peer effect is linear, I use a linear first stage in the average weeks in cycle of the peer group.

[22] In Section 5.3, I analyze non-linear peer effects. In Section 5.4, I analyze alternate functional forms.

[23] See Online Appendix A for a discussion.

[24] This requirement can be relaxed by using a regression discontinuity strategy on average peer week.

As an alternate specification, I also use the following aggregate first stage:

$$repay_{p(i)} = \gamma_0 + \sum_{j \in p(i), j \neq i} \left(\gamma_1 \frac{1\left(0 \leq weeks_j < 5\right)}{N_{p(i)}} + \gamma_2 \frac{1\left(45 \leq weeks_j < 50\right)}{N_{p(i)}} \right) + \gamma_3 X_{i,p(i)} + \psi_{p(i)} \tag{4.4}$$

In this specification, the instruments are the fraction of the peer group in weeks 0-5 and the fraction in weeks 45-50 of the loan cycle. I also use dummy variables indicating whether at least a threshold fraction of the peer group falls into one of these categories. To the vector of controls $X_{i,p(i)}$, I add the highest and lowest values of both number of weeks with Spandana and loan size within the relevant peer group.

The figures in Section 3 suggest an RD interpretation of the week in loan cycle variation. One option, used by Angrist and Lavy (1999), is to run the same IV specification, but with the sample restricted to only the data points close to the discontinuities. Imbens and Lemieux (2008) discuss the equivalence between local linear regression on either side of the discontinuity and Two Stage Least Squares and use a dummy variable for data points to right of the threshold as the instrument. This procedure also involves restricting the sample to a small window around the discontinuity. The new first stage is

$$R_{p(i)} = \delta_0 + \delta_1 W_{p(i),T} + \delta_2 X_{i,p(i)} + \eta_{p(i)} \tag{4.5}$$

where $R_{p(i)} = repay_{p(i)}$, $W_{p(i),T} = 1\left(\sum_{j \in p(i)} \frac{(45 \leq weeks_j < 50)}{N_{p(i)}} > T\right)$, and T is the threshold. The regressions are restricted to peer groups for which either

$$1\left(\sum_{j \in p(i)} \frac{(45 \leq weeks_j < 50)}{N_{p(i)}} > T\right) = 1 \text{or } 1\left(\sum_{j \in p(i)} \frac{(0 \leq weeks_j < 5)}{N_{p(i)}} > T\right) = 1$$

These regressions are also performed for both definitions of the peer group and only use information close to the discontinuity. As in the other specifications, the vector $X_{i,p(i)}$ contains smooth functions of the peer group and individual-level running variables (weeks with the MFI) in addition to loan size controls. Finally, I perform these regressions eliminating centers in weeks $0-5$ with a majority of borrowers in the first loan cycle. First cycle borrowers with weak repayment incentives do not have a natural comparison group.

5 Results

5.1 OLS Estimates and Determinants of Repayment

Table 3 details the OLS estimates of equation 4.1 and shows very high associations between own and peer repayment. Throughout the analysis, I focus on the village peer effect and the center peer effect excluding the group.[25] Column 1 shows a relationship of 81pp between village repayment and individual repayment controlling for own and peer loan size as well as functions of own and peer weeks with the MFI and branch fixed effects. This means that when the entire village switches from full default to full repayment, the individual borrower tends to make the same switch with 81% likelihood. Column 2 separates the village peer effect, 30pp, from the center peer effect, 56pp. Note that approximately two-thirds of the peer association comes from the smaller borrowing center. Column 3 shows the center peer effect alone, at 64pp. The high center level repayment correlations are consistent with the shape of Figure 3. Again, caution is necessary in interpreting the OLS results, as estimates tend to be greatly overstated in the case of unobserved correlated covariates or shocks.

Table 4 captures the correlations between individual repayment and loan size, length of experience borrowing from the MFI and the number of payments made in the loan cycle before the defaults occurred.[26] Column 1 shows the coefficients from a simple OLS regression of repayment on week in the loan cycle, loan size, and total number of weeks spent borrowing from Spandana. This basic specification indicates that individuals with larger loan sizes are slightly less likely to repay and that individuals with longer borrowing histories are also less likely to repay. Columns 2-4 include peer level covariates as well as branch fixed effects. The loan size effect loses significance in the more robust specifications. This absence of a large effect may indicate that Spandana is successful in calibrating loan sizes to a borrower's repayment capacity. However, there does appear to be a more robust relationship between repayment and length of a borrower's relationship with Spandana. The column 4 estimates suggest that a borrower in loan cycle 2 is more than 4-5 percentage points (~10%) less likely to repay than a borrower in cycle 1. This evidence might suggest that the dynamic incentives lose their power as borrowers complete more cycles.

In all specifications, one additional week in the loan cycle at the time of the defaults corresponds to a 1 percentage point greater likelihood of repayment. In other words, indi-

[25] All peer variables exclude the borrowing group. Most borrowers within a group receive their loans on the same day, so there is no variation in weeks in loan cycle at the group level. In contrast, groups within a given center may have staggered loan disbursements.

[26] These variables will be used as covariate controls in the subsequent regression specifications.

viduals in week 50 are 50 percentage points more likely to repay their loans than borrowers in the beginning of their loan cycles. The standard errors of all of the estimates are extremely small. Note that the 1pp coefficient is not very sensitive to the inclusion of branch fixed effects or village level peer group controls. Borrowers who had made more payments before the defaults occurred are indeed more likely to repay. This relationship between loan repayment and number of payments made at the time of the defaults is the building block for the instrumental variables approach used to estimate the peer effects.

5.2 Reduced Form and Instrumental Variables Estimates

The first stage regressions used in the IV estimates are aggregated versions (in means) of the specifications appearing in Table 4. The resulting regression coefficients inevitably look quite similar.[27] Panel A of Table 5 includes the first stage coefficients corresponding to equation 4.3 in columns 1-4 for both the village and center definition of the peer group. In all specifications, average peer group repayment increases by approximately 1pp for every unit increase in the average weeks peer group variable. In specifications where I estimate both the village ex center and the center ex group peer effects, I use two instruments, village ex center and center ex group average weeks in cycle. Columns 2 and 3 show both the village and center level first stage regressions for the combined village and center specifications. Note that there is no effect (or an extremely small effect) of village ex center average weeks on center repayment levels in column 4. This regression can be interpreted as a reduced form regression for the effect of village repayment on center repayment. The absence of a relationship is preliminary evidence that the village peer effect is not very strong.

Columns 5-7 of Panel A of Table 5 display results for the reduced form regressions of individual repayment on the peer group weeks in loan cycle variables. Column 5 presents the reduced form using the village-ex-group peers. The weeks variable is not significant in column 1. Column 2 includes village-ex-center and center-ex-group weeks variables. The coefficient on the center level peer group instrument is 0.00105 (and significant at the 1% level), which is much larger than the (insignificant) coefficient on the village peer group variable. However, because the estimate on the village variable is so imprecise, I cannot conclude that the village and center peer effects are different. Column 7 shows results for the reduced form using only the center-level variables. The magnitude of the coefficient suggests that if the whole center moved from week 0 to week 49 in the loan cycle, individual

[27] 5th order polynomials of individual and peer group weeks with the MFI are included in all subsequent first stage, reduced form and instrumental variables regressions. The individual specifications in Table 4 only include linear and squared weeks with the MFI terms.

repayment would increase by 6 percentage points. In general, the reduced form regressions give evidence that there is a repayment peer effect that is stronger at the more local level.

I also estimate the peer effects model using the fraction of peers with extremely good or extremely bad incentives as instruments. Panel B of Table 5 displays the first stage and reduced form results using these extreme weeks instruments. The results are qualitatively very similar to the average weeks instruments described above. If the peer group shifts from from 0% of members in weeks 0-5 (45-50) to 100% of members in weeks 0-5 (45-50), the repayment likelihood decreases (increases) by 20-30pp. The sum of the coefficients on the 0-5 and 45-50 variables are indistinguishable from zero, implying a symmetric effect of weeks in the loan cycle on repayment incentives. Again, notice that the village peer variables are not significant in the center repayment regressions.

The reduced form estimates using the extreme weeks instruments also provide support for the existence of center-level peer effects. Having a peer group move from 0% to 100% of members in weeks 45-50 increases an individual's likelihood of repaying by approximately 4 percentage points. However, there is no detectable effect of having a large fraction of the peer group in weeks 0-5. For the specification in column 6, a Wald test indicates that the sum of the center coefficients is greater than zero. Therefore, the reduced form estimates suggest that the peer effect is asymmetric and driven by repayers pulling individuals out of default rather than defaulters luring individuals to remain in default.

The results of the 2SLS estimation procedure on the full sample are detailed in Table 6. Parameter estimates of equation 4.1 are shown using the average weeks instrument in columns 1-3 and the fraction extreme weeks instruments in columns 4-6. As in the reduced form, only the center peer effects are significantly different from zero. The magnitude of the peer effect is approximately 10pp using the average weeks instruments and 14pp using the extreme weeks instruments. This translates into a 1.0-1.4pp increase in the probability of repaying the loan for every additional 10pp of repayment in the peer group. The peer effect results are also robust to restricting the data set to those individuals with centers with especially high or especially low incentives. Table 7 presents the average weeks IV specifications restricting the sample to centers close to the 50-week discontinuity.

Finally, Table 8 presents the reduced form and IV regression specifications corresponding to the fuzzy RD interpretation as suggested by Imbens and Lemieux (2008). Because the village effects do not appear to be very strong, I limit the analysis to the center peer effect. The regressions are restricted to sub-samples of centers with either very high or very low incentives. Column 1 restricts the sample to centers where at least 75% of borrowers were at weeks 0-5 or weeks 45-50 of the loan cycle when the defaults occurred. Columns 2 and 3 use

80% and 85% thresholds, respectively. Furthermore, I exclude centers in their first loan cycle with extremely weak incentives at the time of the defaults (weeks 0-5 of their first loans). All of the specifications of Table 8 indicate that the peer group moving from low incentives to high incentives (i.e. weeks 0-5 vs. weeks 45-50) increases the repayment probability by 5.5-7.0pp. This translates into IV peer effects of 11-14pp.

The peer effects estimates are robust to a number of alternative specifications. First, the center-level results look very similar under the inclusion of village fixed effects. Second, the estimates are robust to the inclusion varying polynomial degrees of average peer group and own characteristics. All specifications in the tables include fifth order polynomials, but the results are qualitatively the same under second or third order polynomials. The results also do not depend on the inclusion of minimum and maximum peer weeks variables, which partially control for the variance of the peer group's weeks distribution. Lastly, while the specifications in Table 8 pool peer borrowing centers into high and low incentive groups, the results are robust to more flexible regression discontinuity specifications. The reduced form estimates implied by Figure 4 (using local linear regression with Imbens and Kalyanaraman (2012) optimal bandwidth choice) closely correspond to the IV and RD reduced form estimates in the tables.

5.3 Non-Linear Peer Effects

I also investigate whether a non-linear peer effect may be partially driving the stark shape of repayment observed in Figure 3. While linear peer effects do lead to hump shaped adoption/repayment patterns, non-linear peer effects might help to explain why there is so much bunching towards the poles of full repayment and full default. I am interested in estimating

$$repay_{i,p(i)} = \beta_0 + g\left(repay_{p(i)}\right) + \beta_2 weeks_i + \beta_3 X_{i,p(i)} + \varepsilon_{i,p(i)}$$

where the function $g(.)$ may not be linear. Using the nonparametric IV control function approach of Newey et al. (1999), I plot the non-linear center peer effect in Figure 5. The estimates use series regression with fifth order polynomials[28], with bootstrapped error bars. The shape of the peer repayment relationship is not linear. It is characterized by steep regions around 0 and 1 and a much shallower slope between 0.2 and 0.8. It is statistically only possible to conclude that $g(1) > g(0.2) > g(0)$. However, this steepness near full peer default suggests that coordinating on default may not be a stable equilibrium. The curve

[28]The controls and running variable functions are the same as those included in the linear regressions.

also appears quite symmetric with an inflection point near 50% peer repayment, suggesting that the greatest marginal effects come from the first and the last peer deciding to repay.

5.4 Partial Repayment Peer Effects and Alternate Functional Forms

Finally, Online Appendix Table 10 explores other specifications with qualitatively similar results. Columns 1-3 show the IV coefficients using partial repayment[29] as the outcome of interest. The peer effect estimates are quite a bit larger than in the case of full repayment. If an entire village switches from making 0 payments to making 1 payment, then an individual borrower is 35.6 percentage points more likely to make one payment. Column 2 shows that a lot of this effect (22.2pp) comes through the center level, but the coefficient on village ex center repayment is quite large, at 17.4pp, although insignificant. Unlike in the full repayment specifications, the village and center effects are of the same magnitude. The peer effect is two to three times larger for individuals making one payment vs. making the full set of payments. Column 4 of the same table shows the peer effect as a function of the number of peers who fully repay their loans. The effect of each additional peer is approximately the same as one additional week in the loan cycle.

6 Structural Estimation

How much better or worse would MFI revenues have been if the peer mechanism were disabled? Answering this question requires measuring the time spent in default, during which borrowers paid no interest on outstanding balances. While the reduced form identification strategy allows for estimates of the size of the overall repayment peer effect, structural assumptions are required to characterize the repayment time path. Furthermore, the reduced form analysis does not pin down the absolute level of the peer effect. I introduce a dynamic discrete choice model to investigate the effect of peer effects on MFI profits. In essence, this exercise compares the fragility of microlenders, which include social capital in the loan mechanics and repayment meetings, to more traditional, individual lenders.

6.1 A Simple Model of Loan Repayment

I model the default crisis as a breaking of the social contract between the borrowers and the MFI. The previously rigid repayment schedules became completely flexible following the

[29] I define partial repayment as making at least 1 payment to the MFI after the collector's statement.

District Collector's statement, and borrowers were effectively given free options to delay indefinitely making further payments. Borrowers face the following value function:

$$
\begin{aligned}
& V\left(w_i, x_{p(i)}, \varepsilon_i(1)\right) \qquad (6.1)\\
= & \max_{a_i \in \{0,1\}} \left\{\pi_i\left(a_i, w_i, x_{p(i)}\right) + \mathbf{1}\left(w_i < 50\right)\left(\varepsilon_i\left(a_i\right) + \beta E\left[V\left(w_i + a_i, x'_{p(i)}, \varepsilon'_i(1)\right) | w_i, x_{p(i)}\right]\right)\right\}
\end{aligned}
$$

Each period, borrowers decide whether or not to make one additional installment payment, denoted $a_i \in \{0,1\}$. Borrowers receive a per period utility of $\pi_i\left(a_i, w_i, x_{p(i)}\right)$, which is a function of the action a_i and the state variables $\left(w_i, x_{p(i)}\right)$. The borrower's continuation value is captured by $\beta E\left[V\left(w_i + a_i, x'_{p(i)}, \varepsilon'(1)\right) | w_i, x_{p(i)}\right]$. The variable, $w_i \in [0, 50]$ indexes the total number of loan installments previously paid by individual i and advances deterministically as a function of the previous state and action variables, $w'_i = w_i + a_i$. If $w_i = 50$, then the loan has been fully repaid. The vector $x_{p(i)}$ contains state variables describing individual i 's peer group, $p(i)$. In addition to the per period utility, borrowers also receive an additive, time varying, i.i.d. private utility shock $\varepsilon_i(a_i)$ where $\varepsilon_i(0) = 0$ [30] and $E[\varepsilon_i(1)] = 0$. This cost $\varepsilon_i(1)$ can be thought of as a liquidity shock to the borrower if $\varepsilon_i(1) < 0$. The additive error structure assumption follows Rust (1987) and much of the dynamic discrete choice literature. I assume that all terms in the value function scale linearly with loan size, I_i and omit loan size from the state space.

The per period utility function is assumed to take the following form:

$$
\begin{aligned}
& \pi_i\left(a_i, w_i, x_{p(i)}\right)\\
= & -\mathbf{1}\left(w_i < 50\right)\kappa a_i + \mathbf{1}\left(w_i = 50\right)V_{new}\\
& + E\left[-\mathbf{1}\left(w_i < 50\right)\rho\left(w_i + a_i, x'_{p(i)}\right) + \mathbf{1}\left(w_i = 50\right)\left[\Psi\left(x'_{p(i)}\right)\right] | w_i, x_{p(i)}\right]
\end{aligned}
$$

If the borrower does make a payment $(a_i = 1, w_i < 50)$, then she pays κ. If the loan was fully repaid last period $(w_i = 50)$, then the borrower receives a one-time continuation value V_{new}, representing all of the benefits from borrowing from the MFI in the future.[31]

The model incorporates peer repayment into each individual's value function. Every period, borrowers receive $E\left[\rho\left(w_{it} + a_{it}, x'_{p(i)}\right) | w_i, x_{p(i)}\right]$ (which could be positive or negative) from differing from their peers. Because all borrowers in a peer group make repayment decisions simultaneously, individuals must take the expectation over the peer group's actions

[30] WLOG, $\varepsilon_i(0) = 0$ is a convenient normalization.

[31] While Bulow and Rogoff (1989) show that the threat of credit denial alone is not sufficient to provide repayment incentives, V_{new} may capture other perks from participating in microfinance. Recall Section 3.1.

when deciding whether or not to repay. To close the model, I assume that borrowers receive a terminal utility payoff that depends on peer repayment actions. This peer-based terminal value is denoted $\Psi\left(x_{p(i)}\right)$. For example, future borrowing from the MFI could be less valuable if individual i finishes the loans when only a small fraction of peers has also finished.

To estimate the model, I make several functional form and error distribution assumptions. First, for tractability I assume that the peer state space can be approximated by three variables, $x_{p(i)} = \left(w_{p(i)}, \sigma^2_{p(i)}, w^{50}_{p(i)}\right)$, where $w_{p(i)}$ is the mean state for i's peer group (excluding i), $\sigma^2_{p(i)}$ is the variance of the peer state distribution and $w^{50}_{p(i)}$ is the fraction of peers who have already fully completed making their loan payments at each point in time.[32] Krusell and Smith (1998) make a similar approximation when estimating macroeconomic models with wealth heterogeneity. Second, I assume that the iid, privately observed utility shocks are distributed such that $\varepsilon_i(1) \sim N(0,1)$. The variance of the individual's liquidity shock is not separately identified from the parameters of the model and is normalized to 1. Third, I model the per period peer value as a function of the distance between the individual's own week in cycle and the average peer group week in cycle. Namely, $\rho\left(w_i, x_{p(i)}\right) = \rho_1\left|w_i - w_{p(i)}\right| + \rho_2\left(w_i - w_{p(i)}\right)^2$. Fourth, I model the peer continuation value as a linear function of the fraction of peers who have also repaid their loans, $\Psi\left(x_{p(i)}\right) = \eta w^{50}_{p(i)}$. These assumptions result in five unknown structural parameters, $\theta = (\kappa, V_{new}, \rho_1, \rho_2, \Psi)$. Finally, because the discount factor is not separately identified in this class of model, I calibrate $\beta = 0.999$ on a weekly basis, corresponding to an annual discount factor of 0.95.

Thus, borrowers with $w_i < 50$ will choose to make a payment ($a_i = 1$) if the value is higher than from not paying.

$$\begin{aligned}\Delta V(w_i, x_p; \theta) \quad &\equiv \quad \pi_i\left(1, w_i, x_{p(i)}\right) + \beta E\left[V\left(w_i + 1, x'_{p(i)}, \varepsilon'_i(0), \varepsilon'_i(1)\right) | w_i, x_{p(i)}\right] \\ &\quad\; -\pi_i\left(0, w_i, x_{p(i)}\right) - \beta E\left[V\left(w_i, x'_{p(i)}, \varepsilon'_i(0), \varepsilon'_i(1)\right) | w_i, x_{p(i)}\right] \\ &> \quad -\varepsilon_i(1)\end{aligned}$$

With standard normal $\varepsilon(1)$, the payment probability can be expressed using the normal cdf:

$$\Pr(a_i = 1 | w_i, x_p) = 1 - \Phi(\Delta V(w_i, x_p; \theta)) \tag{6.2}$$

[32] The full state space of the model would ideally be $\left(w_i, \{w_j\}_{,j\in p(i)}\right)$ and would capture the number of payments made by every peer each period. Each borrower can advance by one payment each week, yielding 51 possible individual states. The average peer group has approximately 30-50 borrowers, so a conservative estimate for the number of required states is $51 \times 51^{29} \approx 9 \times 10^{50}$.

With $\varepsilon_i(1)$ unbounded, there is always a strictly positive probability of an individual making a payment at any state in the state space. Thus, all borrowers will take action $a_i = 1$ in finite time, and every individual will eventually repay her loan. However, this repayment process may require an arbitrarily large number of periods. The model does lead to the possibility of multiple equilibria. For a simple illustration see Online Appendix B.

6.2 Firm Profits

The timing of the stream of loan payments determines the MFI's revenues and profits. The Collector's announcement extended the maturity of all outstanding debt with no increase in the installment payments, forcing the contractual 27% annual interest rates toward zero.

Suppose that the MFI faces a weekly cost of capital, r. Then with no defaults and no delays, the expected revenue from a loan with weekly installment I after w payments is:

$$\Pi_{ND}(w) = I \sum_{t=0}^{50-(w+1)} \frac{1}{1+r}$$

Total profits for each loan are $\frac{1}{1+r}\Pi_{ND}(0) - L$. However, with delays in payment, the profit function depends on the borrower's repayment probabilities. The probability of a borrower paying an additional installment is a function of the state variables and is denoted $p\left(a_i = 1|w_i, x_{p(i)}\right)$. The policy functions of the structural model generate these action probabilities. Thus, the expected revenues on a loan with the possibility of delay are:

$$\begin{aligned} E\left[\Pi_D\left(I_i, w_i, x_p, P\right)\right] &= p\left(a_i = 1|I_i, w_i, x_p\right)\left(I_i + \frac{1}{1+r}E\left[\Pi_D\left(w_i + 1, x_p', P\right)|I_i, w_i, x_p\right]\right) \\ &\quad + \left(1 - p\left(a_i = 1|I_i, w_i, x_p\right)\right)\frac{1}{1+r}E\left[\Pi_D\left(w_i, x_p', P\right)|I_i, w_i, x_p\right] \end{aligned}$$

and the terminal value,

$$E\left[\Pi_D\left(I_i, 50, x_p, P\right)\right] = 0, \forall x_p$$

Thus, delay is costly.[33] The expected revenues of an MFI with N loans outstanding are:

$$\sum_{k=1}^{N} E\left[\Pi_D\left(I_i(k), w_i(k), x_{p(i)}(k), P\right)\right] \tag{6.3}$$

[33] It is easy to show that if for all possible state variable realizations, $\left(I_i, w_i, x_{p(i)}\right) \in \Omega_I \times \Omega_i \times \Omega_p$, $p\left(a_i = 1|I_i, w_i, x_{p(i)}\right) < 1$ then $\Pi_{ND}(j) > \Pi_D\left(j, x_{p(i)}\right) \forall j, x_{p(i)}$.

6.3 Estimation

I solve and estimate the model assuming that individuals play a symmetric, Markov Perfect Equilibrium. All borrowers are ex ante identical and have the same probabilities of making a payment conditional on the state variables. Furthermore, I assume that borrowers have rational expectations about their peers' actions (unbiased beliefs), and all borrowers select the same equilibrium if multiple equilibria exist. The strategies and resulting action probabilities associated with the MPE can thus be considered a fixed point of the best response mapping over the possible choice probabilities. I calculate the two-step PML estimator and follow the dynamic games techniques of Aguirregabiria and Mira (2007). For a detailed discussion of the estimation of the underlying model parameters, please see Online Appendix B.

For the structural analysis, I use high frequency repayment data. I restrict the data to the period from March 12, 2006 to September 30, 2006 where the weekly payment indicators are most widely recorded. The resulting sample contains approximately 1.7 million borrower by week observations. I include all members of the borrowing center in the peer group definition, including the individual's own borrowing group. This peer group definition makes the implicit assumption that the size of the peer effect is fixed within the center.

6.4 Counterfactual Model

The goal of the structural exercise is to determine how much more or less costly the crisis was for the MFI because of peer repayment dependencies. In other words, what would have happened to firm profits if the peer terms in the model could have been "switched off"? Restricting κ and V_{new} to be fixed across models, the relevant model without peer effects is:

$$V\left(w_i, \varepsilon\left(1\right)\right) \;=\; \max_{a_i \in \{0,1\}} \left\{1\left(w_i < 50\right)\left(-\kappa a_i + a_i \varepsilon\left(1\right) + \beta E\left[V\left(w_i + a_i, \varepsilon'\left(1\right)\right)\right]\right) + 1\left(w_i = 50\right) V_{new}\right\}$$

I take $\hat{\theta}_{NP} = \left(-\hat{\kappa}, \hat{V}_{new} + \hat{\Psi}, 0, 0, 0\right)$ as given,[34] and I solve for the expected value functions of the model and $\hat{P}_{NP}$ using backward induction. Because the model without peer effects has a single equilibrium, this is quite straightforward. The resulting action probabilities of this model, $\hat{P}_{NP}$, can be plugged into Equation 6.3 to calculate expected revenues under the no peer effect counterfactual.

[34]Note that the two-step pseudo maximum likelihood estimator cannot be used to evaluate the resulting action probability vector, P_{NP}. That method requires rational beliefs about the transition probabilities, based on the observed state transitions in the data. The observed transitions as a function of the state variables are not consistent with $\hat{\theta}_{NP}$ by construction.

6.5 Results

Model Estimates Table 9 shows the estimated model parameters. As predicted, $-\kappa < 0$, representing a cost of making each additional payment. Furthermore, V, the terminal value of repaying is positive. While these parameters are only identified to scale, they indicate that repayment is costly, but yields some terminal reward such as receiving a new loan.

The peer effect parameters show an interesting pattern. First $\Psi > 0$, indicating that the expectation of peer full repayment positively influences an individual's own repayment behavior. Second, $\rho_1, \rho_2 > 0$, which implies that there are actually flow *benefits* from differing from one's peers, not costs. This benefit cuts in the opposite direction as $\Psi > 0$. Taken together, these parameter values imply that the costs of differing from one's peers are concave. If the peer group is likely to finish making all 50 payments quickly, borrowers respond by accelerating their payments. However, for borrowers ahead of the pack, it may not be worth it to wait for their peers to catch up; these individuals may prefer to quickly finish repaying and to be rematched with a new borrowing group.[35]

Figure 6 plots the individual's action probabilities as a function of average peer weeks, conditioning on the individual's own payment history. These curves are the policy functions for individuals in weeks 15, 30 and 45 respectively.[36] The dashed lines plot the smoothed data underlying the structural estimation. The solid lines plot the transition probabilities under the model with peer effects. Finally, the finely dotted flat lines plot the transition probabilities under the counterfactual model with no peer effects. The x-axes in all three plots correspond to the average number of weeks already paid by the center peer group, while the y-axes correspond to the probability of an individual making one additional payment. Under the counterfactual model, the repayment probability does not depend on the peer group.

Panel A details the relationship between peer weeks and repayment likelihood for individuals in week 15 of their loans. If the peer group is farther ahead of the individual borrower, then there is a virtuous peer effect. Individuals make payments with higher probabilities and try to catch up to their peers. They are motivated by receiving Ψ after paying off the loan. However, a higher payment likelihood also arises when the peer group is significantly behind the individual. For these own week variables, the ρ function dominates and a borrower who is ahead of the pack prefers to repay quickly and surge ahead of the peer group. Panel B

[35] During the defaults, borrowers who repaid and had peers who remained in default were rematched with new groups when new loans were disbursed. The Ψ term captures the fact that there may be continuation costs from finishing ahead of the peer group.

[36] I use the weights from the empirical distribution to project the full state space onto these two dimensions.

shows a similar bowl-shaped relationship for individuals in week 30 of their loan cycles.

Panel C captures the relationship between own and peer repayment for individuals in week 45, who are very close to getting a new loan. In contrast to Panels A and B, the model with peer effects is roughly monotone increasing in peer incentives. Having peers with poor repayment incentives only slows down an individual's repayment progress. This is for two reasons. First, there is very little chance that the week 45 borrower will receive the peer bonus Ψ at the end of the loan. Second, there is some value to stalling for the week 45 individual with low incentive peers. These types of borrowers actually receive a repeated flow benefit from ρ as long as they still have payments outstanding. Once the loan has been completed, the borrower no longer receives these flow benefits.

Counterfactual Results To estimate the net costs or benefits of peer effects under the model, I simulate 200 192-week paths for each borrower. At each week in the simulation, I update each borrower's own and peer states using the simulated actions of all individuals in the borrowing center. First, the model with peer effects does a better job coaxing individuals with low incentives to repay than the model without peer mechanisms. The opposite is true for those with already high incentives. Because the average peer week is bounded between 0 and 50, those individuals with low incentives face peers with incentives that are at least as good. In contrast, individuals at week 49 must have weakly better incentives than the rest of their center. I use the firm profit equation and the simulated repayment paths to calculate this expected revenue for each repayment week. Assuming a cost of capital of 10%, Figure 7 plots the expected revenues under the models with and without peer effects. Note that the greatest revenue differences come from increases among borrowers with otherwise low repayment incentives. The expected revenues from individuals in weeks 0-5 are on aggregate 38.2% higher under the peer effects regime. The peer effect benefit is 17.3% for borrowers in weeks 10-15. The revenues are approximately equivalent for individuals in weeks 30-35, while the net peer effect is slightly negative (-1.4%) for those individuals in weeks 45-50. I also compare Spandana's expected revenues under each model. For its 115,000 loans included in the analysis sample, Spandana's expected revenues are 10% higher in a world with peer effects than in a world without them.

Finally, the peer effects estimates allow me to investigate how different initial loan disbursement policies would have fared in encouraging timely repayment. I do this by assuming that each borrowing center consists of 4 groups of 10 individuals each, and that groups are always fully synchronized. However, the MFI can choose how to space the disbursements between groups within a borrowing center. I then simulate the model for different timing arrangements. Because crises generally occur unexpectedly, I calculate the expected value for each spacing arrangement by averaging over each possible timing of the shock.[37]

I investigate three types of spacing regimes, full synchronization, partial synchronization and full separation. Full synchronization is characterized by simultaneous loan disbursement for all members of a borrowing center. In partial synchronization, 2 groups of borrowers are disbursed loans together while the remaining groups wait some number of weeks before receiving their disbursements. In full separation, each group takes turns receiving loan disbursements over time. Within the partial synchronization and full separation strategies, I also vary the number of weeks between each group's loan disbursements.

In the simulations, full synchronization performs the worst out of all of the spacing arrangements. As a rule, putting as much space as possible between groups yields the best outcomes. The full separation strategy with the maximal number of weeks (12-13) between groups yields the highest expected profits for the MFI. The revenue gain from full spacing over full synchronization is very large, on the order of 25%. While there may be reasons outside of the model for keeping groups closer together in terms of loan timing, this spacing exercise shows that any number of weeks between borrowers is preferable to full synchronization.

7 Mechanisms

There are five candidate mechanisms that could be driving the peer effect: social capital-building with microfinance peers; information about future credit prospects from MFI; borrower's availability of funds for repayment; and MFI collections practices; peer mimicry

My preferred interpretation of the results is that borrowers value the opportunity to build social capital through microfinance meetings. Feigenberg et al. (2013) show that the act of having frequent, regular microfinance group (in Spandana's case, center) meetings forges relationships that may be important for economic outcomes outside of the scope of microfinance such as risk-sharing. The group setting allows individuals to learn about each other and to observe one another's repayment behavior and also forces individuals to meet at a set time and place each week, a characteristic that also might be useful for informal contracting. When in default, borrowers do not continue to have weekly meetings and lose microfinance's venue for mandatory repeated interactions. Borrowers must repay their defaulted loans (or take a refinancing plan) in order to restart the group meetings. The relationship between own and peer incentives in the structural exercise suggests that individuals display the highest repayment rates when the peer group is likely to finish paying

[37]Note that this exercise implicitly requires that individuals play the same equilibrium in these various spacing counterfactuals that is observed in the actual data.

off the defaulted loan and resume future borrowing quickly. Additionally, recall that if the borrower's own peer group is unlikely to repay, there are some cases when the borrower prefers to jump ahead of the peer group and to finish making all of the loan installments before the peer group. This pattern is compatible with peers preferring to quickly rejoin their existing borrowing centers if possible. However, if the existing peer group is not going to repay and restart borrowing activities quickly, the individuals would rather join a new borrowing center and begin the process of building new microfinance relationships quickly. Under this interpretation, borrowers find value in the social interactions fostered by microfinance.[38]

There are, however, several competing mechanisms which should also be discussed. A second driver of the observed effects could be borrowers learning from peers about the availability of future credit. Naturally, borrowers at the beginning of the crisis could have been concerned that there would be severe punishment for defaulting or that Spandana would not have the liquidity to disburse loans in the future.

It is unlikely that learning fully drives the estimated effects. MFI field officers made regular visits to all villages regardless of repayment status, and publicized both the availability of new credit and the take-up of new credit in neighboring areas. Further, all information sessions were held at the village level, not the center level. Because any village peer effect is small (albeit noisy), information probably cannot explain the results. In microfinance more generally, outside observers often keep track of new loan disbursements,[39] so news of new post-crisis loans likely traveled across neighborhoods and even villages. Similarly, during the crisis, the credit officers were under intense local scrutiny and often drew crowds of bystanders. Finally, the offer of refinancing plans in late 2007 should have sent a strong signal to borrowers and increased trust across the district. Under these refinancing plans, defaulters could receive new loans before repaying the defaulted loans.[40]

A third explanation is that after the collector's announcement, individuals became liquidity constrained. If the lowest cost repayers repaid first, then those individuals could have lent the proceeds of their subsequent loans to others, facilitating repayment and generating a peer effect. The liquidity mechanism is likely not driving the results. The full roll-out of refinancing plans completely removed liquidity costs of repayment, as borrowers received loan

[38]A natural corollary to this interpretation is that the peer effects are strongest coming from those relationships that are more important to the borrower. Unfortunately, I do not have the data to test this.

[39]Robberies of MFI staff are not uncommon. Individuals know when field staff carry large sums of money.

[40]Additionally, if learning is driving the effect, then we would expect the majority of the peer effect to come from the period before the refinancing plans were introduced. There is no evidence that this is the case. When I run the peer effect regressions for repayment as of October 2006, the peer effect is smaller than the peer effect as of November 2009, but very noisy.

disbursements before having to make any large payments, dampening any liquidity channel.

The fourth mechanism, MFI behavior, could pose problems for the identification strategy. If the credit officers had an effective means of making borrowers repay as a function of some cost (maybe a time cost) spent on each borrower, then the MFI could have decided to focus collections efforts on centers with high concentrations of borrowers already close to full repayment. This would generate the patterns observed in the data without there actually being a peer effect. However, this is unlikely, because center meetings were not taking place after the crisis, giving credit officers no easy way to meet with entire centers at the same time. Assuming that the main fixed cost of making collections is the cost of traveling to a village (which is likely to be true), then it is rational for a credit officer to visit all week-49 borrowers in a village before attempting to encourage repayment from individuals with weaker incentives. If the credit officer is looking for easy repayers, then it would not be rational to focus on borrowers with weak incentives within an above-average center without some sort of strong peer enforcement component. All of these facts would suggest that the MFI strategy would be a far larger component of a village-level peer effect. However, almost the full village peer effect can be attributed to the borrowing center.

A fifth possible explanation is social mimicry, whereby individuals copy the actions of others. The asymmetric peer effects and concave costs of differing from peers make this explanation unlikely. Peers seem to disproportionately influence repayment decisions in positive ways, not negative ways. However, it is impossible to rule out all possible mimicry functions due to the binary nature of the outcome variable.

8 Conclusion

I analyze repayment data following the universal default of Spandana borrowers during the 2006 Krishna crisis. Using a novel identification strategy that exploits the dynamic incentives embedded in microfinance contracts, I find strong evidence of repayment peer effects. Even without joint liability, the decision of a peer group to repay does significantly impact an individual's own repayment likelihood. The entire peer group switching from default to repayment is equivalent to writing off 10 weeks of a borrower's loan. These peer effects seem to be driven by social connections within the borrowing center, not the village, and are likely cultivated by regular MFI practices themselves. Estimates from a dynamic discrete choice model of repayment indicate that the asymmetric nature of the peer effect implies that the net value of social effects is actually positive for MFI revenues.

In light of these results, several policy implications emerge. With respect to these types of default episodes, continuing to foster peer effects through the format of frequent group meetings has positive effects on the value of assets following widespread defaults. Lenders could also improve the value of their loan portfolios following a crisis by staggering the disbursements of loans within the peer lending center, as demonstrated by the spacing exercise. In terms of collections practices following defaults, the results suggest two policies. First, convincing one person in a local peer group to repay has disproportionately large spillover effects, as demonstrated by the non-linear peer effect shape. Therefore, collections practices that give special incentives to the first repayer could have large repayment spillovers. Second, in settings like the Krishna defaults, eliminating joint liability is probably a good decision. In the aftermath of the crisis, joint liability would have made it much more costly for individuals with low incentive peers to jump ahead and repay.

In sum, the social capital that MFIs work so hard to nurture does help to preserve asset values in the aftermath of crises. The Krishna District experience shows that while in good times microfinance can boast nearly perfect repayment rates, when problems arise the system can become quite fragile. This study provides some of the first evidence that, against this backdrop, peer effects may actually improve repayment rates and act as a stabilizing force.

References

Acemoglu, D. and J. Angrist (2000): "How large are the social returns to education? Evidence from compulsory schooling laws," *NBER Macroeconomic Annual*, 15, 9–74.

Aguirregabiria, V. and P. Mira (2002): "Swapping the nested fixed point algorithm: A class of estimators for discrete Markov decision models," *Econometrica*, 70, 1519–1543.

——— (2007): "Sequential estimation of dynamic discrete games," *Econometrica*, 75, 1–53.

Ahlin, C. and R. Townsend (2007): "Using Repayment Data to Test Across Models of Joint Liability Lending," *Economic Journal*, 117, F11–F51.

Angrist, J. and V. Lavy (1999): "Using Maimonides' Rule to Estimate The Effect of Class Size on Scholastic Achievement," *Quarterly journal of economics*, 114, 533–575.

Banerjee, A., T. Besley, and T. Guinnane (1994): "The neighbor's keeper: the design of a credit cooperative with theory and a test," *The Quarterly Journal of Economics*, 109, 491–515.

Basu, K. (2011): "Hyperbolic discounting and the sustainability of rotational savings arrangements," *American Economic Journal: Microeconomics*, 3, 143–171.

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

Bond, P. and A. Rai (2009): "Borrower runs," *Journal of Development Economics*, 88, 185–191.

Bulow, J. and K. Rogoff (1989): "Sovereign debt: Is to forgive to forget?" *The American Economic Review*, 79, 43–50.

Conley, T. and C. Udry (2010): "Learning about a new technology: Pineapple in Ghana," *The American Economic Review*, 100, 35–69.

Duflo, E. (2004): "The medium run effects of educational expansion: Evidence from a large school construction program in Indonesia," *Journal of Development Economics*, 74, 163–197.

Duflo, E. and E. Saez (2002): "Participation and investment decisions in a retirement plan: The influence of colleagues' choices," *Journal of Public Economics*, 85, 121–148.

Feigenberg, B., E. Field, and R. Pande (2013): "Building social capital through microfinance," *Forthcoming, Review of Economic Studies*.

Fischer, G. and M. Ghatak (2010): "Repayment frequency in microfinance contracts with Present-Biased borrowers," *Economic Organisation and Public Policy Discussion Papers Series*, 21.

Ghatak, M. (1999): "Group lending, local information and peer selection1," *Journal of Development Economics*, 60, 27–50.

Ghatak, M. and T. Guinnane (1999): "The economics of lending with joint liability: theory and practice1," *Journal of Development Economics*, 60, 195–228.

Giné, X. and D. Karlan (2006): "Group versus individual liability: evidence from a field experiment in the Philippines," Yale University Economic Growth Center Working Paper 940.

——— (2009): "Group versus Individual Liability: Long Term Evidence from Philippine Microcredit Lending Groups," *Yale University Working Paper*.

Giné, X., K. Krishnaswamy, and A. Ponce (2011): "Strategic Default in joint liability groups: Evidence from a natural experiment in India," *Work in Progress*.

Guiso, L., P. Sapienza, and L. Zingales (2013): "The determinants of attitudes towards strategic default on mortgages," *The Journal of Finance*.

Imbens, G. and K. Kalyanaraman (2012): "Optimal bandwidth choice for the regression discontinuity estimator," *The Review of Economic Studies*, 79, 933–959.

Imbens, G. and T. Lemieux (2008): "Regression discontinuity designs: A guide to practice," *Journal of Econometrics*, 142, 615–635.

Iyer, R. and M. Puri (2012): "Understanding Bank Runs: The Importance of Depositor-Bank Relationships and Networks," *The American Economic Review*, 102, 1414–45.

Karlan, D. (2007): "Social connections and group banking*," *The Economic Journal*, 117, F52–F84.

Kelly, M. and C. Gráda (2000): "Market Contagion: Evidence from the Panics of 1854 and 1857," *American Economic Review*, 1110–1124.

Krusell, P. and A. A. Smith, Jr (1998): "Income and wealth heterogeneity in the macroeconomy," *Journal of Political Economy*, 106, 867–896.

Manski, C. (1993): "Identification of endogenous social effects: The reflection problem," *The Review of Economic Studies*, 60, 531–542.

Munshi, K. and J. Myaux (2006): "Social norms and the fertility transition," *Journal of Development Economics*, 80, 1–38.

Newey, W., J. Powell, and F. Vella (1999): "Nonparametric estimation of triangular simultaneous equations models," *Econometrica*, 67, 565–603.

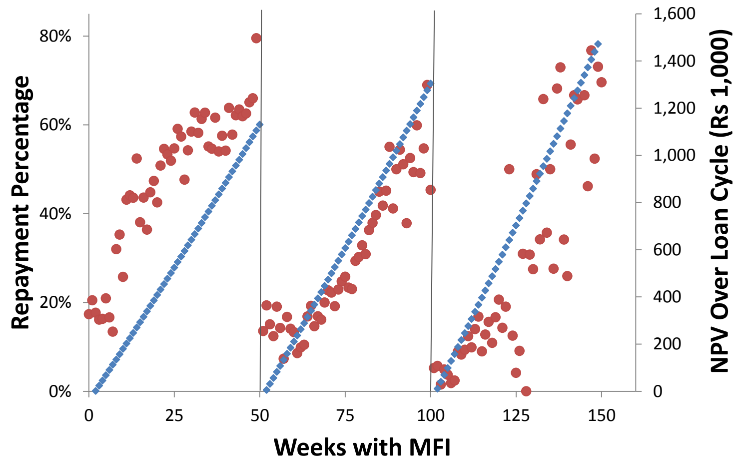
Rust, J. (1987): "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica*, 999–1033.

Stiglitz, J. (1990): "Peer monitoring and credit markets," *The World Bank Economic Review*, 4, 351.

The Hindu (2006): "MFIs Throw Norms to the Winds," April 29.

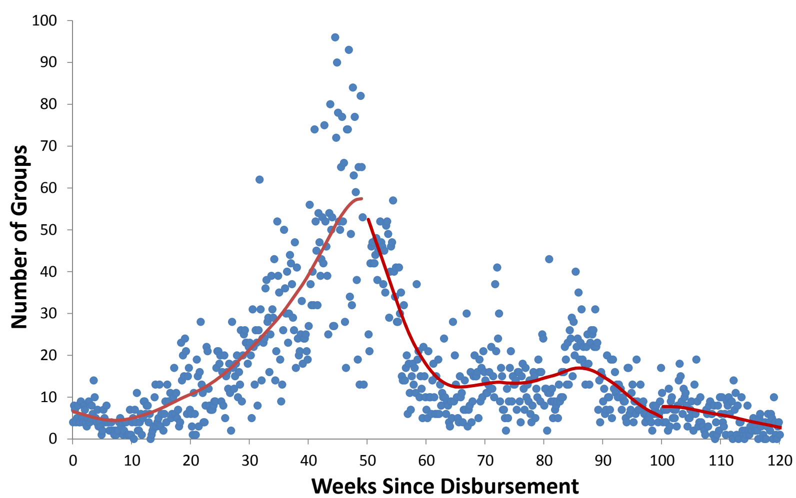
Townsend, R. M. (1994): "Risk and Insurance in Village India," *Econometrica*, 62, 539–591.

Appendix A: Figures



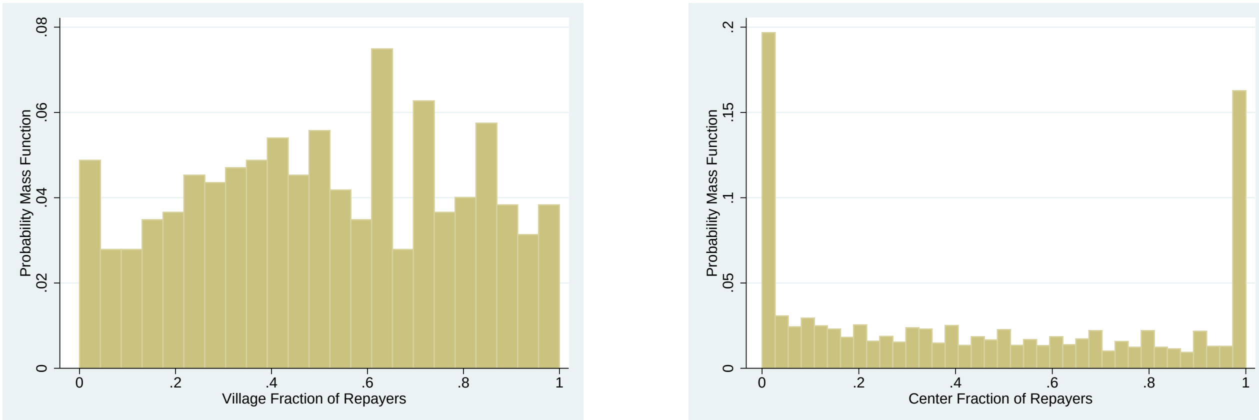
The dotted lines represent stylized repayment incentives across the borrowing relationship. They are derived from simple NPV calculations. The units are in terms of NPV per Rs 1,000 in loan principal initially borrowed and are displayed on the y-axis to the right. The dots represent actual, observed average repayment rates. and units are displayed on the y-axis to the left.

Figure 1: Stylized Repayment Incentives and Actual Repayment Over the Loan Cycle



The scatter plot displays total number of borrowing groups receiving an initial disbursement by week before the crisis. Each dot represents one day. The red solid lines are local linear regressions run separately for groups within each cycle. The bandwidth is chosen using Imbens and Kalayanaraman's optimal RD.

Figure 2: Number of Borrowing Groups by Week with MFI



Panel A: Village Repayment

Panel B: Center Repayment

Figure 3: PMFs of Full Repayment by Village and Center

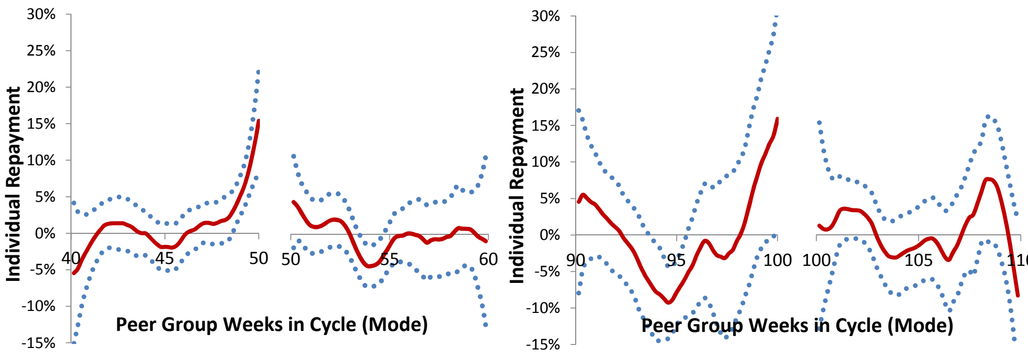


Figure 4: Individual Repayment by Peer Group Incentives across each Cycle Discontinuity

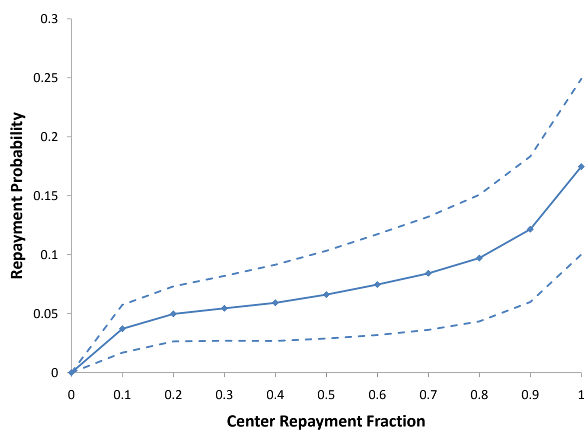
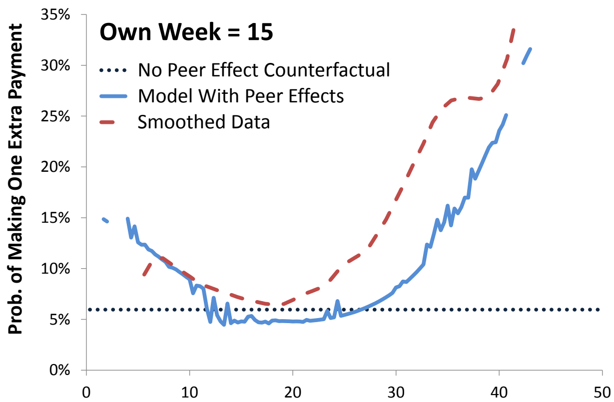
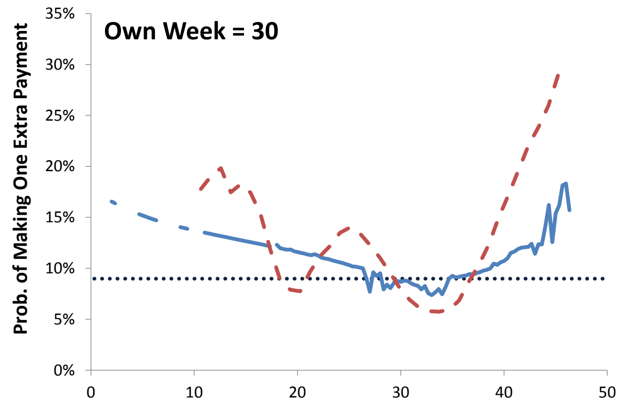


Figure 5: Non-Linear Effects of Peer Repayment on Individual Repayment

Panel A: Own Week = 15



Panel B: Own Week = 30



Panel C: Own Week = 45

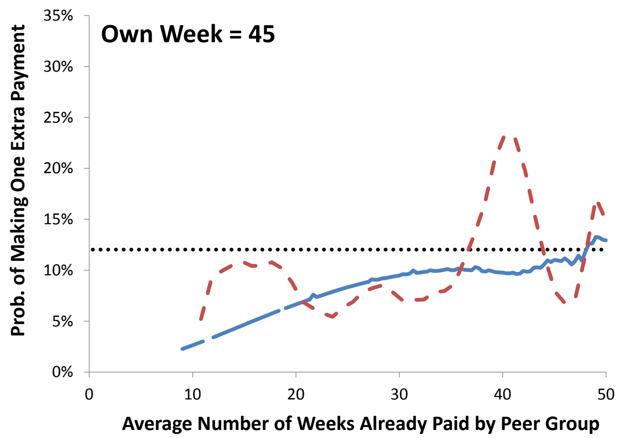


Figure 6: Model and Actual Transition Probabilities by Average Peer Weeks for Individuals at 15, 30, and 45 "Own" Weeks

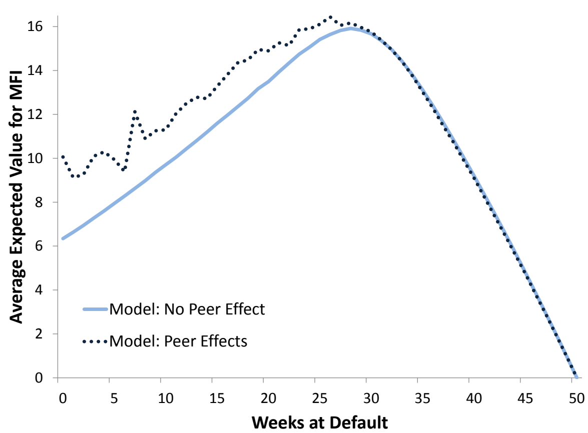


Figure 7: Expected Loan Value Calibration by Weeks Completed when the Defaults Occurred

Appendix B: Tables

Table 1: Summary Statistics

		Std. Dev.
As of 3/9/2006		
Mean Loan Size (Rs)	7,644	1,911
Mean Loan Outstanding (Rs)	3,621	3,120
Mean Date of Disbursement	9/24/2005	111 days
Number of Loans	114,943	
Number of Groups	13,437	
Number of Centers	5,340	
Number of Villages/Slums	574	
As of 11/20/2009		
Percent of Loans Still in Arrears	56.89%	
Most Common Stated Loan Purposes		
Livestock	26.33%	
Textiles	16.81%	
Retail Shop	11.43%	

Table 2: Average Village Characteristics by Average Borrower Week

	Average Week in Cycle	Fraction in ***First*** 5 Weeks	Fraction in ***Last*** 5 Weeks	Significant Difference
Population	-22.36	812	-580	No
	(17.41)	(695)	(474)	
Population: Scheduled Caste and Tribe	-2.067	194.8	-50.25	No
	(3.908)	(156.8)	(114.57)	
Cultivation Area per Capita (000s)	1.679	-37.68	59.54	No
	(1.567)	(39.69)	(81.17)	
Irrigated Area per Capita (000s)	1.834	-77.24	19.01	No
	(1.882)	(49.11)	(92.56)	
Distance to Town (km)	-0.313*	20.93***	13.58***	No
	(0.158)	(5.680)	(4.878)	
Primary Schools per Capita (000s)	0.00770	-0.121	0.435	No
	(0.00853)	(0.185)	(0.467)	
Medical Facilities Indicator	-0.00395**	0.0691	-0.201*	**
	(0.00156)	(0.0929)	(0.117)	
Health Centers per Capita (000s)	-8.49e-05	0.0230*	0.0102	No
	(0.000445)	(0.0127)	(0.0283)	
Health Subcenters per Capita (000s)	-0.00155	0.135**	0.0386	No
	(0.00172)	(0.0558)	(0.0777)	
Number of Banks per Capita (000s)	-0.000412	0.0103	0.0309	No
	(0.000786)	(0.0250)	(0.0435)	
Railway Access Indicator	0.00132	-0.112**	-0.000257	*
	(0.00128)	(0.0444)	(0.0419)	
Paved Roads Indicator	0.00140	-0.0395	0.0985**	No
	(0.00122)	(0.0780)	(0.0406)	
Weeks with MFI Controls	Yes	Yes	Yes	Yes

Notes: Each row represents a separate set of regressions. Column 1 reports a regression where the independent variable is the average week in the loan cycle in the village. Columns 2 and 3 represent a second set of regressions, where the independent variables are the fraction of individuals in the first 5 and last 5 weeks of their loan cycles. Column 4 indicates the significance of the difference in the coefficients in columns 2 and 3. Controls for functions fo village average weeks with MFI. Standard errors are clustered at the village level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 3: OLS Regressions of Own Repayment on Peer Repayment

	(1)	(2)	(3)
Village Repayment ex Group	0.813***		
	(0.0232)		
Village Repayment ex Center		0.302***	
		(0.0221)	
Center Repayment ex Group		0.556***	0.636***
		(0.0158)	(0.0156)
Indvidual and Peer Controls	Yes	Yes	Yes
Fixed Effects	Branch	Branch	Branch
Observations	107734	107734	107734
R-squared	0.345	0.401	0.394

Notes: The dependent variable is an indicator for whether an individual fully repaid her loan by November, 2009. The regressors of interest represent the average repayment in the peer group. Standard errors are clustered at the village level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 4: Individual Determinants of Loan Repayment

	(1)	(2)	(3)	(4)
Week in Cycle	0.0119***	0.0111***	0.0106***	0.0107***
	(0.000430)	(0.000434)	(0.000412)	(0.000395)
Number of Weeks with MFI	-0.00265***	-7.82e-05	0.000219	-0.000977**
	(0.000254)	(0.000148)	(0.000154)	(0.000483)
Number of Weeks with MFI Squared				9.44e-06**
				(4.00e-06)
Loan Amount (Rs 1000s)	-0.00536*	0.00167	0.00229	-0.00797
	(0.00281)	(0.00230)	(0.00221)	(0.00882)
Loan Amount (Rs 1000s) Squared				0.000674
				(0.000524)
Village Peer Controls	No	Yes	Yes	Yes
Fixed Effects	No	No	Branch	Branch
Observations	114943	114943	114943	114943
R-squared	0.170	0.213	0.283	0.289

Notes: The dependent variable is an indicator for whether an individual fully repaid her loan by November, 2009. Village peer controls include the following average village level variables: loan size, loan size squared, and second order polynomials of number of weeks with the MFI. Standard errors are clustered at the village level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 5: Aggregate First Stage and Reduced form: Average and Extreme Weeks Instruments

Panel A: Average Weeks Instrument	Aggregate First Stage: Peer Group Repayment				Reduced Form: Individual Repayment		
	Village	Village	Center	Center			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Village Average Week in Cycle ex Group	0.0110***				0.00122		
	(0.00107)				(0.00112)		
Village Average Week in Cycle ex Center		0.0112***	0.0000			0.000357	
		(0.000946)	(0.000962)			(0.00101)	
Center Average Week in Cycle ex Group		-0.000141	0.0112***	0.0113***		0.00105***	0.00123***
		(0.000196)	(0.000428)	(0.000468)		(0.000336)	(0.000383)
Week in Cycle					0.0107***	0.0103***	0.0103***
					(0.000385)	(0.000383)	(0.000382)
Individual and Peer Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Branch	Branch	Branch	Branch	Branch	Branch	Branch
Observations	107734	107734	107734	107734	107734	107734	107734
R-squared	0.673	0.656	0.490	0.478	0.292	0.296	0.289

Panel B: Extreme Weeks Instrument	Aggregate First Stage: Peer Group Repayment				Reduced Form: Individual Repayment		
	Village	Village	Center	Center			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fraction of Village (ex g) in ***First*** 5 Weeks	-0.268***				-0.0229		
	(0.0499)				(0.0520)		
Fraction of Village (ex g) in ***Last*** 5 Weeks	0.244***				0.0492		
	(0.0446)				(0.0468)		
Fraction of Village (ex c) in ***First*** 5 Weeks		-0.270***	-0.0694			-0.0162	
		(0.0423)	(0.0449)			(0.0467)	
Fraction of Village (ex c) in ***Last*** 5 Weeks		0.269***	0.0446			0.0106	
		(0.0379)	(0.0397)			(0.0416)	
Fraction of Center (ex g) in ***First*** 5 Weeks		-0.00505	-0.237***	-0.245***		0.00327	-0.00577
		(0.00747)	(0.0151)	(0.0167)		(0.0138)	(0.0160)
Fraction of Center (ex g) in ***Last*** 5 Weeks		0.000244	0.212***	0.205***		0.0428***	0.0392***
		(0.00685)	(0.0167)	(0.0176)		(0.0138)	(0.0149)
Week in Cycle					0.0107***	0.0105***	0.0106***
					(0.000401)	(0.000405)	(0.000413)
Individual and Peer Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Branch	Branch	Branch	Branch	Branch	Branch	Branch
Observations	107734	107734	107734	107734	107734	107734	107734
R-squared	0.647	0.622	0.438	0.429	0.292	0.296	0.289

Notes: Panel A presents results using the average weeks instrument. Panel B presents results using the extreme weeks instruments. Columns 1-4 in both panels present aggregate first stage regressions, where the dependent variable is the fraction of individuals in the relevant peer group who fully repaid their loans by November, 2009. Columns 5-7 present reduced form regressions, where the dependent variable is an indicator for whether an individual fully repaid her loan by November, 2009. Note that columns 1 and 5 define the peer group as the village ex group. Columns, 2,3 and 6 analyze two levels, of the peer group: village ex center and center ex group. Column 2 shows the first stage regression for the village ex center peer group, while column 3 shows the first stage regression for the center ex group peer group. Columns 4 and 7 focus on only the center ex group peer group. All specifications include the following individual-level controls: loan size, loan size squared, and fifth order polynomials of weeks with the MFI. Peer controls are defined at the relevant level and include: average loan size and loan size squared, fifth order polynomials of the average number of weeks with the MFI, and the minimum and maximum values for weeks with MFI within the peer group. Standard errors are clustered at the village level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 6: IV Regressions of Own Repayment on Peer Repayment

	Average Weeks			Extreme Weeks		
	(1)	(2)	(3)	(4)	(5)	(6)
Village Repayment ex Group	0.111			0.159		
	(0.0929)			(0.110)		
Village Repayment ex Center		0.0327			0.0471	
		(0.0800)			(0.0937)	
Center Repayment ex Group		0.0967***	0.112***		0.141***	0.145***
		(0.0298)	(0.0329)		(0.0473)	(0.0488)
Week in Cycle	0.0107***	0.0102***	0.0103***	0.0106***	0.00997***	0.0101***
	(0.000383)	(0.000386)	(0.000384)	(0.000414)	(0.000458)	(0.000459)
Individual and Peer Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effects	Branch	Branch	Branch	Branch	Branch	Branch
Observations	114943	107734	107734	114943	107734	107734
R-squared	0.306	0.328	0.323	0.311	0.341	0.332

Notes: In all specifications, the dependent variable is an indicator for whether an individual fully repaid her loan by November, 2009. In columns 1-3, the instrument is the average week in the loan cycle for the relevant peer group definition. In columns 4-6, the instruments are the fraction of individuals in the relevant peer group who are in the first 5 weeks of their loan cycles and the fraction of the peer group in the last 5 weeks of their loan cycles. All specifications include the following individual-level controls: loan size, loan size squared, and fifth order polynomials of weeks with the MFI. Peer controls are defined at the relevant level and include: average loan size and loan size squared, fifth order polynomials of the average number of weeks with the MFI, and the minimum and maximum values for weeks with MFI within the peer group. Standard errors are clustered at the village level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 7: IV Regressions of Own Repayment on Peer Repayment: Restricted Sample

	Sample Restriction: ± 5 weeks around Discontinuities		
	Mode	> 50%	> 75%
	(1)	(2)	(3)
Center Repayment ex Group	0.143***	0.136***	0.143***
	(0.0506)	(0.0491)	(0.0552)
Week in Cycle	0.0107***	0.0106***	0.0111***
	(0.000569)	(0.000569)	(0.000661)
Individual and Peer Controls	Yes	Yes	Yes
Fixed Effects	Branch	Branch	Branch
Observations	28757	30646	21892
R-squared	0.433	0.403	0.418

Notes: In all specifications, the dependent variable is an indicator for whether an individual fully repaid her loan by Nov., 2009. The specifications are restricted versions of column 3 in Table 7. Column 1 restricts the sample to peer groups where the mode week with MFI is close to a discontinuity. Columns 2 and 3 restrict the sample based on the fraction of individuals in the peer group close to the discontinuity. All specifications include the following individual-level controls: loan size, loan size squared, and fifth order polynomials of weeks with the MFI. Peer controls are defined at the center level and include: average loan size and loan size squared, fifth order polynomials of the average number of weeks with the MFI, and the minimum and maximum values for weeks with MFI within the peer group. Standard errors are clustered at the village level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 8: Center Peer Effects: Fuzzy RD

	Fraction of Peers in Extreme Weeks		
	0.75	0.80	0.85
	(1)	(2)	(3)
A. Reduced Form			
Peer Group Has High Repayment Incentive	0.0705**	0.0631**	0.0551*
	(0.0306)	(0.0308)	(0.0321)
Week in Cycle	0.0114***	0.0117***	0.0119***
	(0.000641)	(0.000661)	(0.000677)
Individual and Peer Controls	Yes	Yes	Yes
Fixed Effects	Branch	Branch	Branch
Observations	21188	20561	19857
R-squared	0.378	0.380	0.378
B. Instrumental Variables			
Center Repayment ex Group	0.144**	0.127**	0.111*
	(0.0585)	(0.0586)	(0.0616)
Week in Cycle	0.0112***	0.0115***	0.0117***
	(0.000662)	(0.000684)	(0.000710)
Individual and Peer Controls	Yes	Yes	Yes
Fixed Effects	Branch	Branch	Branch
Observations	21188	20561	19857
R-squared	0.424	0.422	0.416

Notes: The dependent variable is an indicator for whether an individual fully repaid her loan by Nov., 2009. All regressions restrict the sample to peer groups with very high or very low repayment incentives. The instrument is an indicator for whether the peer group has very high incentives. Panel A presents reduced form estimates. Panel B presents IV results. All specifications include individual-level controls: loan size, loan size squared, and fifth order polynomials of weeks with the MFI. Peer controls are defined at the relevant level and include: average loan size, loan size squared, fifth order polynomials of the average number of weeks with the MFI, and the minimum and maximum values for weeks with MFI within the peer group. Standard errors are clustered at the village level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 9: Structural Parameter Estimates

Description	θ	Estimate	Counterfactual
Flow Values			
Repayment Amount	$-\kappa$	-2.52	-2.52
		(0.18)	
Peer - Linear	ρ_1	0.04	0
		(0.05)	
Peer - Squared (* 1,000)	ρ_2	0.050	0
		(0.029)	
Continuation Values			
Individual	V_{new}	65.02	65.02+2.42
		(6.54)	
Peer	Ψ	2.42	0
		(0.67)	

Standard Errors are calculated by bootstrapping the full estimation procedure.

Online Appendix A: The Reflection Problem

Suppose that all peer groups are of size n and that the peer effect operates through the average repayment in the peer group. Then the key structural parameter to identify is α_2. Note that the problem is symmetric for all individuals in the same peer group, so

$$\begin{aligned}
repay_1 &= \alpha_0 + \alpha_1 date_1 + \alpha_2 \sum_{j \neq 1} \frac{repay_j}{n-1} + \epsilon_1 \\
repay_2 &= \alpha_0 + \alpha_1 date_2 + \alpha_2 \sum_{j \neq 2} \frac{repay_j}{n-1} + \epsilon_2 \\
&\ldots \\
repay_n &= \alpha_0 + \alpha_1 date_n + \alpha_2 \sum_{j \neq n} \frac{repay_j}{n-1} + \epsilon_n
\end{aligned}$$

So, first sum equations 2 through n

$$\sum_{j \neq 1} \frac{repay_j}{n-1} = \frac{n}{n-1}\alpha_0 + \alpha_1 \sum_{j \neq 1} \frac{date_j}{n-1} + \alpha_2 \frac{1}{n-1} \sum_{i=2}^{n} \sum_{j \neq i} \frac{repay_j}{n-1} + \sum_{j \neq 1} \frac{\epsilon_j}{n-1}$$

where

$$\begin{aligned}
&\alpha_2 \frac{1}{(n-1)^2} \sum_{i=2}^{n} \sum_{j \neq i} repay_j \\
= \; &\alpha_2 \left[\frac{repay_1}{(n-1)} + \frac{(n-2)}{(n-1)} \sum_{j \neq 1} \frac{repay_j}{(n-1)} \right]
\end{aligned}$$

So

$$\sum_{j \neq 1} \frac{repay_j}{n-1} = \frac{1}{1 - \alpha_2 \frac{(n-2)}{(n-1)}} \left(\frac{n}{n-1}\alpha_0 + \alpha_1 \sum_{j \neq 1} \frac{date_j}{n-1} + \alpha_2 \frac{repay_1}{(n-1)} + \sum_{j \neq 1} \frac{\varepsilon_j}{n-1} \right)$$

Plugging this back into the first equation gives:

$$\begin{aligned}
repay_1 &= \alpha_0 + \alpha_1 date_1 + \frac{\alpha_2}{1 - \alpha_2 \frac{(n-2)}{(n-1)}} \left(\frac{n}{n-1}\alpha_0 + \alpha_1 \sum_{j \neq 1} \frac{date_j}{n-1} + \alpha_2 \frac{repay_1}{(n-1)} + \sum_{j \neq 1} \frac{\varepsilon_j}{n-1} \right) + \varepsilon_1 \\
&= \tilde{\alpha}_0 + \frac{1}{\left(1 - \frac{\alpha_1 \alpha_2}{(n-1) - \alpha_2 (n-2)}\right)} \left(\alpha_1 date_1 + \frac{\alpha_1 \alpha_2}{1 - \alpha_2 \frac{(n-2)}{(n-1)}} \sum_{j \neq 1} \frac{date_j}{n-1} \right) + \tilde{\varepsilon}
\end{aligned}$$

Now, let's go back and look at the average peer repayment equation, since this is in essence, the first stage of my regressions

$$\begin{aligned}
\sum_{j\neq 1}\frac{repay_j}{n-1} &= \frac{1}{1-\alpha_2\frac{(n-2)}{(n-1)}}\left(\frac{n}{n-1}\alpha_0+\alpha_1\sum_{j\neq 1}\frac{date_j}{n-1}+\alpha_2\frac{repay_1}{(n-1)}+\sum_{j\neq 1}\frac{\varepsilon_j}{n-1}\right)\\
&= \phi+\frac{1}{1-\alpha_2\frac{(n-2)}{(n-1)}}\left(\alpha_1+\frac{\alpha_2^2}{n-1}\right)\sum_{j\neq 1}\frac{date_j}{n-1}
\end{aligned}$$

where ϕ includes all of the other terms. So the coefficient on average date in this regression (excluding $date_1$ which is orthogonal to the other date variable) is

$$\frac{1}{1-\alpha_2\frac{(n-2)}{(n-1)}}\left(\alpha_1+\frac{\alpha_2^2}{n-1}\right)$$

So the ratio of the reduced form coefficient over the first stage coefficient is what IV gives us, so

$$\begin{aligned}
\frac{\frac{\alpha_1\alpha_2}{1-\alpha_2\frac{(n-2)}{(n-1)}}}{\frac{1}{1-\alpha_2\frac{(n-2)}{(n-1)}}\left(\alpha_1+\frac{\alpha_2^2}{n-1}\right)} &= \frac{\alpha_1\alpha_2}{\alpha_1+\frac{\alpha_2^2}{n-1}}\\
&= \frac{\alpha_2}{1+\frac{\alpha_2^2}{\alpha_1(n-1)}}\\
&\approx \alpha_2
\end{aligned}$$

If anything, the small sample bias makes this estimate too low. IV gives a consistent estimate of the peer effect.

Online Appendix B: Structural Estimation Details

Multiple Equilibria

This model allows for the existence of multiple equilibria. To illustrate a simple static case of multiple equilibria, suppose that there are two members of each borrowing center. Let the payoffs of each borrower be $\pi_i = a_i V + \rho_S 1(a_1 = a_2) - \rho_D 1(a_1 \neq a_2)$. The possible payoffs are:

	$a_2 = 0$	$a_2 = 1$
$a_1 = 0$	ρ_S, ρ_S	$-\rho_D, V - \rho_D$
$a_1 = 1$	$V - \rho_D, -\rho_D$	V, V

Note that if $\rho_S > V - \rho_D$ there are multiple equilibria in the stage game. This same logic carries forward to the full dynamic game. Under a model with multiple equilibria, the estimation procedure will in fact deliver a consistent estimate of the underlying model parameters under the assumption that conditional on the state, all individuals and borrowing centers are playing the same equilibrium.

Estimation Details

The estimation of the PML estimator of Aguirregabiria and Mira (2007) follows the following two-step procedure:

1. As a first stage, I estimate both peer and own repayment probabilities for each state in the state space. These estimated action probabilities serve as each individual's beliefs about future peer actions.

2. Given these beliefs, I update each individual's transition probabilities using the model. Then using maximum likelihood, I select the primitives of the model, θ, that best match the individual's observed transition probabilities.

Empirical States To discretize the state space, I bin the possible values as follows: $w_i \in \{0, 1, ..., 50\}$, $w_p \in \left\{0, \frac{1}{3}, \frac{2}{3}, ..., 49\frac{2}{3}, 50\right\}$, $\sigma_p^2 \in \{low, high\}$, $w_p^{50} \in \left\{0, \frac{1}{3}, \frac{2}{3}, 1\right\}$. This results in $151 \times 2 \times 4 = 1208$ values of x_p and $1208 \times 51 = 61608$ total states. While this may seem like a very large state space, the structure of the model puts restrictions on the allowed transitions from state to state: $w_i$and w_p are only able to increase by at most 1 week or stay the same. This yields $2 \times 4 \times 2 \times 4 = 64$ admissible transitions for each state, greatly reducing the computational burden. I denote the full state space as $\Omega_i \times \Omega_p$.

First Stage For the first stage, I follow Aguirregabiria and Mira (2002) and use a multinomial sieve logit to estimate the probabilities of transitioning to one of the 2 possible own repayment states or 32 possible peer repayment states from any state in the state space. Alternately, I could use the empirical distribution of transitions conditional on state. However, doing so would be very noisy. There are some transitions that are never observed and others that are observed rarely. The multinomial logit smooths the transition probabilities, helping to fill gaps in the data. The logit for the agent's own repayment probability is:

$$\Pr\left(a_i = a | w_i, x_p\right) = \frac{\exp\left(\phi_a \mathbf{q}\left(w_i, x_p\right)\right)}{\sum_{a' \in \{0,1\}} \exp\left(\phi_{a'} \mathbf{q}\left(w_i, x_p\right)\right)}$$

where $\mathbf{q}\left(w_i, x_p\right)$ is a vector of third degree polynomials of the state variables. ϕ_a is the vector of coefficients for action choice a. The logit for the peer group's state transition is:

$$\Pr\left(x'_p = \chi | w_i, x_p\right) = \frac{\exp\left(\phi_{p,\chi} \mathbf{q}\left(w_i, x_p\right)\right) \mathbf{1}\left(\chi \in B\left(x_p\right)\right)}{\sum_{\chi' \in \Omega_p} \exp\left(\phi_{p,\chi'} \mathbf{q}\left(w_i, x_p\right)\right) \mathbf{1}\left(\chi' \in B\left(x_p\right)\right)}$$

where $\phi_{p,\chi}$ is a vector of coefficients for new state χ and $B\left(x_p\right)$ is the set of permissible transition states. Note that $\chi \in B\left(x_p\right)$ if $\chi = \left(\chi_1, \chi_2, \chi_3\right) : \chi_1 \leq x_p + 1, x_p^{50} \leq \chi_3 \leq x_p^{50} + 1$. $\hat{P}^0$ is the resulting vector of estimated transition probabilities evaluated at each state.

Second Stage: Model Solution Recall that the transition probabilities can be calculated given equation 6.2. This problem can be reformulated using the beliefs, P estimated in the first stage, to evaluate all expectations. For $w_i < 50$:

$$\begin{aligned} \Psi_i\left(1 | w_i, x_{p(i)}, P\right) &= \Phi(\pi^P\left(1, w_i, x_{p(i)}\right) - \pi^P\left(0, w_i, x_{p(i)}\right) \\ &+ \beta \sum_{x'_{p(i)} \in \Omega_p} \left[\Gamma_i\left(w_i + 1, x'_{p(i)}, P\right) - \Gamma_i\left(w_i, x'_{p(i)}, P\right)\right] f^P\left(x'_{p(i)} | w_i, x_{p(i)}\right)) \end{aligned}$$

$\Gamma_i\left(w'_i, x'_{p(i)}, P\right)$ is the expected value function at state $\left(w'_i, x'_{p(i)}\right)$, where all beliefs and expectations are taken over the transition probabilities, P. The probability mass function $f^P\left(x'_{p(i)} | w_i, x_{p(i)}\right)$ gives the peer transition probabilities for state $\left(w_i, x_{p(i)}\right)$ according to P.

Because $\pi^P\left(a_i, w_i, x_{p(i)}\right)$ is linear in the model's parameters, it can be rewritten

$$\pi^P\left(a_i, w_i, x_{p(i)}\right) = z_i^P\left(a_i, w_i, x_{p(i)}\right) \theta$$

Similarly, the expected value functions can be decomposed into observable and stochastic components:

$$\Gamma_i\left(w_i', x_{p(i)}', P\right) = \Gamma_i^Z\left(w_i', x_{p(i)}', P\right)\theta + \Gamma_i^\lambda\left(w_i', x_{p(i)}', P\right)$$

Finally, combining:

$$\Psi_i\left(1|w_i, x_{p(i)}, P\right) = \Phi\left(\tilde{z}^P\left(w_i, x_{p(i)}\right)\theta + \tilde{\lambda}^P\left(w_i, x_{p(i)}\right)\right)$$

where

$$\begin{aligned}
\tilde{z}^P\left(w_i, x_{p(i)}\right) &= z_i^P\left(1, w_i, x_{p(i)}\right) - z_i^P\left(0, w_i, x_{p(i)}\right) \\
&\quad \beta \sum_{x_{p(i)}' \in \Omega_p} \Gamma_i^Z\left(w_i + 1, x_{p(i)}', P\right) - \Gamma_i^Z\left(w_i, x_{p(i)}', P\right) f^P\left(x_{p(i)}'|w_i, x_{p(i)}\right) \\
\tilde{\lambda}^P\left(w_i, x_{p(i)}\right) &= \beta \sum_{x_{p(i)}' \in \Omega_p} \Gamma_i^\lambda\left(w_i + 1, x_{p(i)}', P\right) - \Gamma_i^\lambda\left(w_i, x_{p(i)}', P\right) f^P\left(x_{p(i)}'|w_i, x_{p(i)}\right)
\end{aligned}$$

Second Stage: Pseudo-Likelihood Maximization Aguirregabiria and Mira (2007) demonstrate that the vector of equilibrium transition probabilities, P^* represents a fixed point of $\Psi_i\left(1|w_i, x_{p(i)}, P\right)$. Thus given a consistent estimate $\hat{P}^0$ from the first stage, the two step estimator is $\hat{\theta}_{2S} = \arg\max_\theta Q_M\left(\theta, \hat{P}^0\right)$:

$$Q_M(\theta, P) = \sum_{t=1}^{T} \sum_{i=1}^{N(g)} \ln \Psi_i\left(a_{it}|w_{it}, x_{i,t}, P; \theta\right)$$

where i indexes individual borrowers and t indexes weeks.

Online Appendix C: Supplemental Figures and Tables

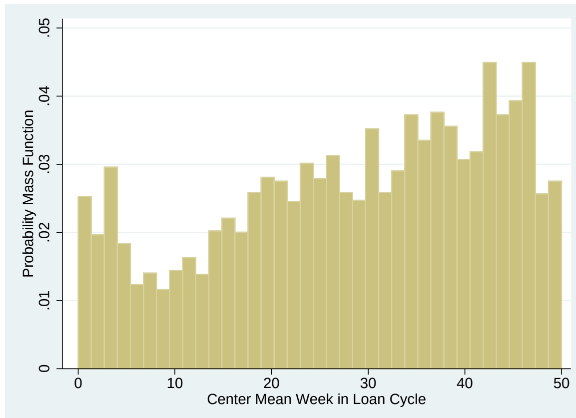
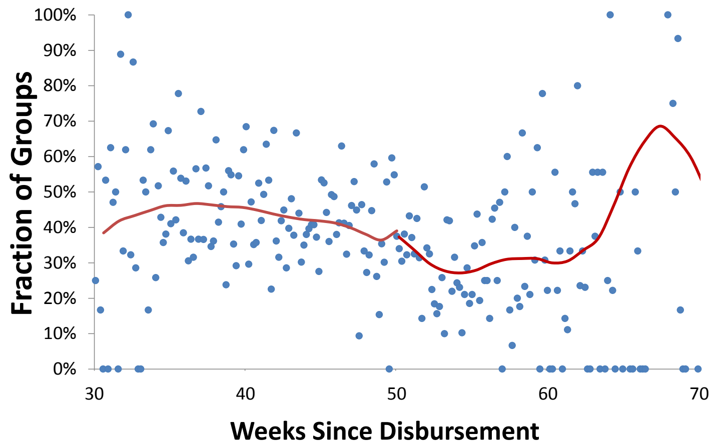
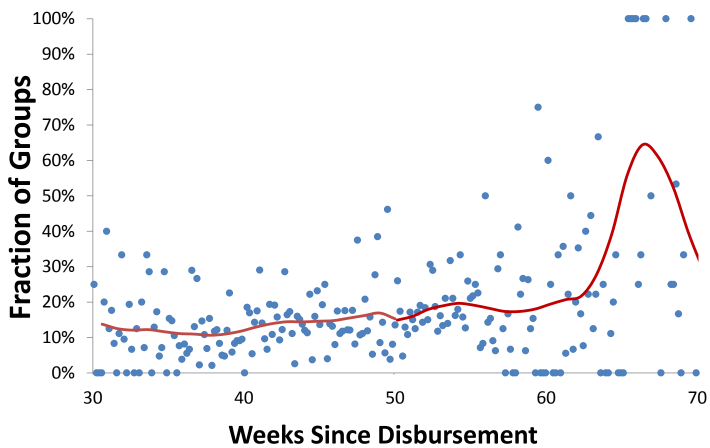


Figure 8: Distribution of Average Weeks in Cycle Across Centers

Panel A: Livestock



Panel B: Textiles



Panel C: Retail Shop

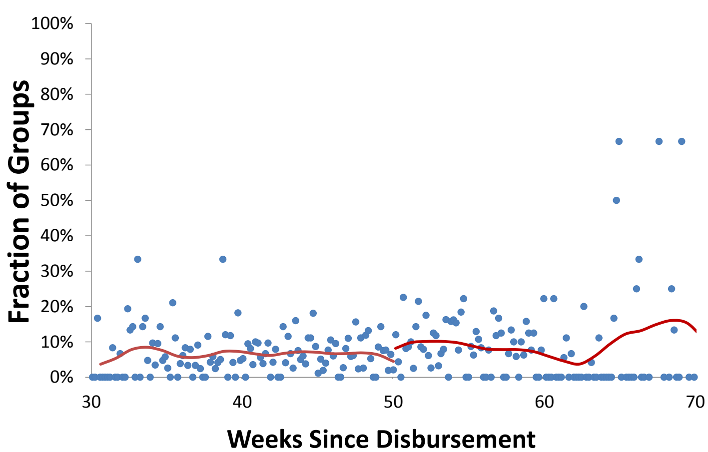


Figure 9: Placebo RD Across Loan Thresholds by Loan Purpose

Table 10: IV Regressions of Partial Repayment on Peer Partial Repayment

	Partial Repayment			# Peers Repaid
	(1)	(2)	(3)	(4)
Village Repayment ex Group	0.356**			
	(0.165)			
Village Repayment ex Center		0.174		
		(0.140)		
Center Repayment ex Group		0.222***	0.301***	0.0115***
		(0.0708)	(0.0722)	-0.00143
Week in Cycle	0.00466***	0.00420***	0.00429***	0.0102***
	(0.000378)	(0.000386)	(0.000401)	-0.000378
Individual and Peer Controls	Yes	Yes	Yes	Yes
Fixed Effects	Branch	Branch	Branch	Branch
Observations	107734	107734	107734	107734
R-squared	0.249	0.299	0.296	0.317

Notes: In columns 1-3, the dependent variable is an indicator for whether an individual partially repaid her loan by November, 2009. The endogenous regressors of interest are partial repayment rates of the peer groups. In column 4, the dependent variable is full repayment by November 2009, and the endogenous regressors are the number of peers who have fully repayed by November 2009. In all specifications, the instrument is average peer group weeks in the cycle. All specifications include the following individual-level controls: loan size, loan size squared, and fifth order polynomials of weeks with the MFI. Peer controls are defined at the relevant level and include: average loan size and loan size squared, fifth order polynomials of the average number of weeks with the MFI, and the minimum and maximum values for weeks with MFI within the peer group. Standard errors are clustered at the village level. *** significant at 1%, ** significant at 5%, * significant at 10%.