Peer Monitoring and Microcredit: Field experimental evidence from Paraguay

Jeffrey Carpenter* Tyler Williams†

April 12, 2010

Abstract

Given the substantial amount of resources currently invested in microcredit programs, it is more important than ever to accurately assess the extent to which peer monitoring by borrowers faced with group liability contracts actually reduces moral hazard. We conduct a field experiment with women about to enter a group loan program in Paraguay and then gather administrative data on their repayment behavior in the six month period after the experiment. In addition to the experiment which is designed to measure individual propensities to monitor under incentives similar to group liability, we collect a variety of the other potential correlates of behavior and repayment. Controlling for other factors, we find a very strong causal relationship between the average monitoring propensity of one's loan group and repayment. Our lowest estimate suggests that borrowers in groups with above median average monitoring are 36 percent less likely to have a problem repaying their portion of the loan. In addition, to confirming some previous results, we also find some evidence that risk preferences, social preferences and cognitive skills affect repayment.

1 Introduction

It is now generally agreed upon that one of the major impediments to climbing out of poverty in the developing world is the lack of access to credit. In fact, economists now expect that even small amounts of financial capital, if used to start or expand ventures, can ultimately give the poor the needed "leg up." This vision, developed to a large extent at the Grameen Bank in Bangladesh, Bank Rakyat in Indonesia and Banco Sol in Bolivia, among many other places, has been fine tuned and implemented across the developing world (Armendariz de Aghion and Morduch, 2005).

For a long while, however, there has been a dearth of solid evidence on whether these "microcredit" programs actually help borrowers or, for that matter, whether the loan programs work in accordance with the incentive structures embedded in the contracts. Although

*Department of Economics, Middlebury College and IZA; jpc@middlebury.edu.
†Department of Economics, MIT; tkwillia@mit.edu.

the availability of data on loan programs has not been a major issue, self-selection into programs has, and, as a result, investigators have been forced to make only guarded statements about the impact of specific microcredit programs. This problem is, to some extent, on the verge of being "solved," given new data from a number of recent randomized trials. In one study (Banerjee et al., 2009), micro finance institutions were opened randomly in half the slums of Hyderabad India, and, after a year and a half, there appeared to be little effect of access to credit on per capita expenditures. However, borrowers' expenditures on durables did increase for households with existing businesses, indicating the fulfillment of expansion plans.

While there has been more work to resolve whether the loan programs work as theory suggests, many studies suffer from other data limitations. In some of the most straightforward models (for example, Stiglitz, 1990, Besley and Coate, 1995, Armendariz de Aghion, 1999, Rai and Sjostrom, 2004), microcredit "works" because the moral hazard and adverse selection problems faced by bankers with poor customers are pushed onto groups of borrowers, which also lowers the cost of lending. Specifically, if one starts from the premise that credit is not extended to the poor because these borrowers provide no collateral to assure that they act prudently once given a loan, then by insisting that each member of a group of borrowers will be held responsible for the loans taken out by the others, bankers incentivize each group member to monitor the activities of her peers and to threaten social sanctions (including group expulsion) in response to observed moral hazard. In other words, "peer monitoring," with the threat of social sanctions, solves the banker's moral hazard problem. Indeed, Gomez and Santor (2008) do find lower default rates in Canadian group lending programs but their data do not allow them to cleanly identify whether the effect is driven by the incentives of peer monitoring or by the differential selection of "good" borrowers into the group lending program. At the same time, the randomized intervention by Gine and Karlan (2008) which controls for any possible selection effects allows the authors to conclude that Philipino banks need not bother with groups because individual loans perform just as well.

Given that the amount currently invested in microcredit programs is quite large and that even the best evidence is not crystal clear, it is more important than ever to determine whether group liability actually attenuates the moral hazard problem: *does peer monitoring reduce the likelihood of default?*

We are interested in identifying a direct causal link between the propensity to monitor one's peers in a group lending program and the rate at which members of the group default. While there already exists a literature that attempts to estimate this link, the results are not conclusive. One strand of this literature, which is perhaps higher in external validity, measures peer monitoring by proxy, though it does so in actual group lending programs. Another strand, much higher in internal validity, uses laboratory experiments to test whether the incentives of group lending are strong enough to cultivate peer monitoring. Considering the first approach, Wydick (1999) made a valuable early contribution by showing that Guatemalan loan-group members who work farther away, on average, from the other members are less likely to insure each other against negative shocks and have lower repayment rates. Further, when group members know the sales of the other members, insurance is more likely to be offered and there are fewer defaults. More recently, Ahlin and Townsend (2007) study group lending in Thailand, where they use the fraction of the group living in the same village and the number of members with a relative in the group as measures of peer monitoring. Their estimates indicate that the bank is less likely to have penalized the group by raising the interest rate if more of the group lives in the same village and if there are fewer relatives in the group, the latter result being interpreted as it potentially being harder to discipline close relatives. Lastly, working in Peru, Karlan (2007) also shows that the physical distance between group members affects repayment: the higher the fraction of the group living within a ten minute walk of each other, the less in default the members are at the end of the loan cycle.[1]

In the laboratory, the incentives of group lending have been simulated to more cleanly identify the effects of peer monitoring. In addition to the contribution of Abbink et al. (2006) who showed how "social ties" can affect loan performance via selection into loan groups, Cason et al. (2009) show that group members are willing to monitor each other, even at a cost, and when the cost to group members is lower than the cost to the banker, group lending is more profitable. More still is learned when the lab is brought to the field. Gine et al. (2009) set the stage with experiments conducted in a large market in Peru. They show how group loan programs can raise repayment because of the embedded mutual insurance arrangement that allows some borrowers to invest in riskier, but more rewarding projects. Somewhat surprisingly, working in Vietnam with poor inhabitants of Ho Chi Minh City, Kono (2006) shows that group lending performs worse than individual lending even after group members can monitor and penalize each other. In South Africa and Armenia, Cassar et al. (2007) also simulate the incentives of group lending and find that experimental measures of trustworthiness between group members predict experimental repayment rates.

Despite all this interesting work, the evidence that peer monitoring affects real loan repayment rates is still circumstantial. While informative, the survey work relating the physical distance between group members to loan performance is, at best, testing whether a potential cost of monitoring predicts repayment; this is not the same as linking the act of monitoring to loan outcomes. At worst, survey measures may proxy for social preferences relevant to repayment (for example, altruism, trust, or reciprocity) that may have less directly to do with monitoring. Likewise, although we learn a lot about the decision to monitor and strategically default in the lab, ultimately we also want to know how monitoring affects loan performance outside the lab.

We contribute by combining aspects of the two previous stands of literature to directly test whether peer monitoring predicts loan performance. Borrowing from the behavioral literature, we develop an experiment to measure individual propensities to monitor one's peers in a social dilemma with incentives similar to group lending. We then test whether the monitoring propensities of women about to enter a real group lending program in Paraguay predict loan performance six months later.

There are many advantages to our approach. First, instead of relying on a proxy for the cost of monitoring, we directly measure the behavioral propensity of individuals using an experiment in which monitoring is costly. Second, because our participants did not know the exact identity of the people that they chose to monitor in the experiment, our measures of peer monitoring are inherent; that is, they could not be conditioned on individual characteristics like being a friend or being a bad credit risk. In this sense our measures are

[1] Other valuable contributions to the survey-based literature include Hermes et al.(2005), Kritikos and Vigenina (2005), Barboza and Barreto (2006), Simtowe and Zeller (2007) and Feigenberg et al. (2009).

much less likely to be endogenous. Third, inspired by Karlan (2005), we estimate the effect of peer monitoring on subsequent loan performance. Because we ran the experiment before the groups received their first loan (we collected loan data six months after we measured the behavioral data), simultaneity bias is also less likely to affect our results, and we can be more confident that we are estimating a causal relationship between peer monitoring and repayment. Fourth, since overt default rates tend to be very low in group lending, like Wydick (1999), our loan performance measure is broader and indicates whether or not an individual had trouble repaying her loan during the cycle. However, while broader, our measure is still constructed from administrative data and corroborated by cross-reports from individual interviews. Fifth, in addition to measures of peer monitoring, our protocol also allowed us to gather behavioral measures of time, risk, and social preferences that we can also use to predict repayment problems. Lastly, we collect a large set of controls that include standard demographics and a number of other potential correlates of default (for example, the number of family members in the loan group, an objective measure of default risk, and a measure of non-verbal IQ).

Although we find many interesting results, at this point we focus on the question that motives our research. Our data suggest that there is a significant link between peer monitoring and group loan performance. Specifically, we find that individuals in groups populated by inherently "nosey" monitors are approximately ten percent less likely to have problems repaying their loans. Further, our estimates are robust to differences in the formulation of our peer monitoring measure and the inclusion of a number of other significant and important factors. In fact, when the controls are added, our point estimates increase substantially. These results suggest that, regardless of whether or not group lending leads to measureable reductions in poverty, it is the case that moral hazard in the groups is attenuated by peer monitoring.

We proceed by first describing our participants and the loan program in which they participated. We then describe the design of our peer monitoring experiment and the methods that we used to gather our other behavioral measures. Before estimating the link between peer monitoring and loan performance, we first describe how we created individual propensities to monitor. Considering our results, we begin by focusing on our main results described above and then we look at some of the other important factors that affect loan performance. In the final section, we offer a few concluding remarks.

2 Loan Program Details and Participant Characteristics

Our participants are women in a group loan program run by the Fundación Paraguaya de Cooperación y Desarrollo (Paraguayan Foundation for Cooperation and Development). The Fundación is a non-profit organization headquartered in Asunción, Paraguay. It also has many branch offices throughout the Eastern half of the country that administer the Fundación's many programs. The goal of the Fundación is to empower the low-income citizens of Paraguay by helping them develop entrepreneurial skills and by giving them the resources necessary for them to apply these skills in their lives.

Three of their programs are focused on the development of entrepreneurial skills: (i) a self-sustaining agricultural high school, (ii) the Junior Achievement program, which focuses on business education in schools, and (iii) a business incubator which helps entrepreneurs learn new business techniques. Their fourth program is a microcredit program called "Bancomunal" (that is, "Community Bank"), which helps give low-income entrepreneurs the capital that they need to start and maintain small businesses. As with many microcredit programs, this program offers much lower interest rates than those offered by banks and loan sharks.

Traditionally, all of the Fundación's microloans had been made to individuals. Also, these individual loans, which the Fundación continues to make, require borrowers to provide some sort of physical capital that can be seized in lieu of payment. However, just prior to the start of our project, the Fundación began a group loan program that does not require its participants to offer collateral to receive a loan. In the program, a loan is made to a group of women that is formed through a mix of recruitment by the Fundación and by the women themselves. When at least 15 (and no more than 20) women have been identified to form a group, these women, by group decision, approve each potential borrower's membership.

After a group has formed, the group members decide on the size of the loan that they will request from the Fundación. To decide the amount, they determine an individual loan amount for each woman in the group. This amount is primarily determined by how much each woman would like to borrow, but also by the borrowers' (and the Fundación employees') opinions of how much each woman can afford to borrow. For their first loan, group members may only request amounts between 100,000 PGY (about \$17) and 400,000 PGY (about \$67).[2]

Each group decides whether they would like to make loan payments weekly for two months or biweekly for two or three months, and each group member is responsible for repaying her portion of the group loan as well as the interest on her portion. The interest rate charged depends on the duration of the loan and on the payment frequency (which are chosen by each group), but all rates are less than 50 percent annually.

Although the group loan program does not employ physical collateral, the Fundación does use joint liability and sequential loaning mechanisms to help motivate repayment, as do many other microfinance institutions. Specifically, they require that all borrowers repay their portion of the group loan in order for any borrower in the group to receive part of a second group loan with the Fundación. Thus, the group as a whole is liable for any defaulting group member's unmade payments. If the group does not cover this liability, then the group cannot request another loan.

After a loan has been completely repaid, there are three possible changes before the next cycle begins: (i) borrowers may remove themselves from the group or be expelled by group decision; (ii) borrowers may "take a break" and choose not to request a loan but still remain in the group (this outcome can also be enforced by the group); and (iii) borrowers may request a loan amount that is higher by up to 50 percent of their previous loan. However, it is important to note that new borrowers cannot join groups after the first cycle. Also, if borrowers decide to take a break for one (or more) loan(s), they are still expected to help repay defaulters' loans if they want to receive loans as part of the group in the future.

[2] The Paraguayan currency is Guaraníes; the exchange rate at the time was about 1 PGY = \$0.00017. All conversions in the paper use this rate.

The participants in the Bancomunal program are all women from the lower economic strata of Paraguayan society. However, their characteristics vary widely, as can be seen in the top two panels of Table 1. The mean age of our participants was 37 years, but the women varied in age between 17 and 60. While a few of our participants had graduated from high school and even attended college, well over half (57 percent) stopped their education after primary school. Lastly, 60 percent of our participants were married, although long-term cohabitation without a formal marriage is also common in Paraguay.

Considering their socio-economic characteristics, 26 percent of our participants classify themselves as the "head of the household." The minimum monthly income in the sample is only 100,000 PGY (about $17), while the maximum is over 6 million PGY. Median monthly income in the study is 1.5 million PGY, while the legal minimum monthly salary in Paraguay is about 1 million PGY. Given their income, it is no surprise that our participants find it hard to save substantial amounts. In fact, while mean savings are 24,650 PGY, more than 80 percent of our participants report having no savings.

The women participate in a variety of business activities, though most are small, entrepreneurial efforts run out of the women's homes. Some examples of these ventures include food preparation, delivery, and sales; very small convenience stores/stalls; clothing production; and used clothing sales. On average, our participants have between 7 and 8 years of experience in their businesses. Although a few women work for wages in an outside business, they are encouraged to invest their loans in an entrepreneurial effort. The women in the study all live in two neighborhoods that are uniformly poor. One is a neighborhood of Asunción, Paraguay's capital and largest city, while the other is farther away and part of a suburb of Asunción.

3 Designing an Experiment to Measure Peer Monitoring

We refined a social dilemma experiment that has already been used extensively in the field so that it better suited our purposes. In Carpenter and Seki (2010), Japanese fishermen participated in a voluntary contribution experiment in which non-monetary sanctions are used to control free riding. The experiment has also been used in Carpenter et al. (2004a) and Carpenter et al. (2004b) to examine the behavior of poor people in Southeast Asian urban slums. While it is possible to create measures of individual propensities to cooperate and punish other group members using this experiment, we decided to make one subtle change that allows us to focus more directly on the decision to monitor one's peers. After the contribution stage, but before punishment was allowed, participants were asked if they wanted to have access to the contribution decisions of the other members of their experimental group. If the participant paid a small fee, she was shown, in random order to protect anonymity, the contribution levels of all the participants in her experimental group. Only those who paid the monitoring fee were eligible to socially sanction the other participants. Because monitoring is at the heart of much of the theory of group lending, in our analysis we focus on this decision.

In our Paraguayan implementation, 58 participants in 8 sessions were randomly split into anonymous groups of four where they stayed for the entire experiment.[3] In all but the first session, the experiment lasted eight rounds.[4] The detailed instructions for the experiment appear in the appendix and as a result we discuss only the important highlights. At the beginning of each round, participants were given fifteen 100 PGY coins as an endowment and then asked how many they would like to contribute to a "group project," keeping the residual. They were told that the total amount contributed by the group would be increased by 50 percent and then redistributed evenly to the group. Hence, the marginal per capita return from the public good is 37.5 PGY for each 100 PGY coin contributed. The incentives are those of a standard social dilemma: each coin contributed to the group project returns only 0.375 of a coin to the contributor while the same person also receives 0.375 for all those coins contributed by the other group members. As a result, one can always earn more by not contributing but the socially efficient outcome occurs when all group members contribute fully.[5]

After the contribution stage, each participant was shown her gross income for the round and the group total contribution. At this point participants were asked if they wanted to monitor the rest of the group.[6] If an individual paid 50 PGY, she was shown the individual contribution data and then could send messages of disapproval (unhappy faces) to other individuals in the group for an additional 50 PGY per message. Notice that because monitoring and sanctioning are costly only for the monitor, they should not influence the standard, free-riding, prediction based on egoistic preferences. Participants who contribute nothing should not care if their contributions are seen, especially given the anonymity of the experiment. Further, since punishment does not reduce one's payoff, it should be ignored.

Despite the incentives, as one can see in Table 2, the participants contributed an average of 8.4 coins, which is 56 percent of their endowment. This fraction is comparable to both previous studies using this experimental design and to the related experiment of Masclet et al. (2003). Despite the costs involved, we also see that the probability of monitoring on any given round was 0.47 and that, while low, an average of 0.14 messages of disapproval are sent per round. To get a better sense of the dynamics, Figure 1 presents a time series of the average experimental behavior. Unlike standard voluntary contribution games (see for example Ledyard, 1995), contributions are relatively flat over time, as is the likelihood that the average participant monitors. Looking closer we see that monitoring does fall off to some extent; however, this is probably explained by the fact that contributions increase slightly over time and therefore there is less reason to monitor. In section 5 we use the

[3] Clearly 58 participants can not be evenly divided into groups of four. Instead of turning away people from our limited subject pool, we relied on the fact that participants could not know who the other members of their group were and formed groups with "shadow members." These randomly chosen shadow members contributed to their own group but their behavior was also counted in another group to get the total up to four persons.

[4] The first session lasted 10 rounds. Because this took longer than our time allocation, in all subsequent sessions eight rounds were played. Our analysis uses all the available data.

[5] Average earnings in the experiment were 15,400 PGY, or about \$2.60 at the time of the study. Considering 18 percent of Paraguayans live off of less than \$2 per day and participants' median daily earnings were 50,000 PGY (\$8.50), the incentives appeared to be salient.

[6] To save time and to eliminate end-game effects, the monitoring and punishment ended after round seven. This was not announced until contributions had been made for the eighth round near the end of the experiment.

individual level data from this experiment to create monitoring propensity measures for our participants.

4 Gathering Other Behavioral Controls

While focusing our analysis on peer monitoring, we decided that it would be important to control for some of the other behavioral reasons why people may or may not repay their loans. To this end, along with asking the demographics and socio-economic survey questions discussed above, we also had the women participate in four other tasks that followed our peer monitoring experiment. Summary statistics of the responses to these tasks and a few other relevant survey questions appear in the bottom two panels of Table 2.

Because cognitive skills may determine the degree to which people are able to properly analyze costs and benefits and make sound business decisions (Burks et al., 2009), we decided to implement a very short version of a standard non-verbal IQ test. We borrowed three questions from the 60-question Raven's Progressive Matrices examination (Raven et al., 2003), in which people are asked to complete a pattern. There are a number of reasons why our participants might not do well (for example, lack of familiarity with this sort of task or low levels of formal education), which were confirmed. Out of three questions, nobody got all three correct but there is still some variation in performance that we can exploit: just under one-tenth got two correct and about one-quarter got one right.

Risk preferences may determine how the women invest their loans and, therefore, we conducted the same binary lottery experiment used in Cardenas and Carpenter (2009) to assess levels of risk aversion. Six binary lotteries in which the odds are 50-50 are arrayed in a circle, and the participant is asked to pick the lottery in which she would most like to participate. Once the lottery was chosen, a coin was tossed to determine the payoff, and this amount was added to the participant's final earnings. The lottery in the 1 o'clock position is the sure thing: both outcomes are 33,000 PGY. At 3 o'clock is a 25,000|47,000 lottery, followed by 18,000|62,000 at 5 o'clock, 11,000|77,000 at 7 o'clock, 4,000|91,000 at 9 o'clock, and 0|95,000 at 11 o'clock. The expected values for the lotteries in this exercise were well over half the median daily earnings of the subjects, and subjects were paid based on their choices and the flip of a coin.

Both the expected value and the variance in payoffs increase as one moves clockwise around the circle, with the exception of the last lottery. Here, while the variance in payoffs continues to increase, the expected value plateaus. Given this design we can infer that anyone choosing the last lottery at 11 o'clock must be risk seeking (or possibly risk neutral). According to Table 1, 14 percent of our participants choose the most risky lottery. At the same time, another 10 percent can be easily classified as extremely risk averse because they chose the sure thing.

People may also struggle to repay their loans because of impatience (see, for example, Meier and Sprenger, 2009). To gather information on the discount rates of our participants, in the survey we asked them to (hypothetically) pick between four pairs of payments. One payment in each pair was always 18,000 PGY to be paid immediately, and the other, to be paid in one month, was taken from the set $\{18{,}300, 18{,}750, 19{,}500, 21{,}000\}$. If the participant always choose the smaller/sooner payment we know that her monthly discount rate must be at least 16.7 percent. Returning to Table 2, we see that 17 percent of our participants always chose the sooner payment and can therefore be categorized as "high discounters." Likewise, if the participant always chose the larger/later payment her monthly discount rate can be no larger than 1.67 percent. These women (26 percent) were classified as "low discounters."

Repayment behavior might also depend on one's social orientation to the group. For example, altruists might be less willing to impose a negative externality on the rest of the group by defaulting on their portion of the group loan. In a hypothetical Dictator Game (Forsythe et al., 1994), each participant was endowed with 30,000 PGY, any portion of which she could give to an anonymous stranger. As indicated in Table 2, the mean amount "given" was 43 percent of the endowment. While this seems like a lot compared to student versions of the Dictator Game, it is not when compared to other non-students (Carpenter et al., 2005).

There are other avenues through which one's social orientation might affect repayment. Those people who are more engaged in the community might be less likely to default because doing do so would tarnish their image or reputation. It might also be the case that the composition of the group matters. As indicated in Ahlin and Townsend (2007), the number of family members or friends might affect repayment. With these pathways in mind, we asked participants, as part of the survey, to list the community groups with which they were affiliated. We also asked each woman for the number of family members and friends in her loan group. Summary statistic from these survey responses appear at the bottom of Table 2.

5 Creating Behavioral Measures of Peer Monitoring

Based on the data from our social dilemma experiment, our goal is to create behavioral measures of the propensity of each of our participants to incur some cost to monitor the behavior of the other women in her group. Our hope is that, since the experiment is anonymous and players could not condition their monitoring choices on any observable characteristics specific to the other participants, the propensities that we create are "inherent" and therefore capture the instinct to monitor other people in a situation similar in incentives to group lending. In other words, we seek to capture the basic "nosiness" of our participants.

The obvious place to start with our data is to look at the raw frequency with which participants chose to monitor the group. As seen in Figure 2(a), there is considerable variation in this frequency. Slightly more than 15 percent of the participants never monitor, while about the same fraction always monitor. There is a weak mode at monitoring once but, overall, there is a considerable amount of monitoring: the average monitoring frequency is 0.46.

One obvious problem with looking at the raw monitoring frequency is that it does not account for the fact that some groups are generally more cooperative than others. Therefore, there may be more need to monitor in some groups than in others. To a great extent, this means that it is hard to make "apples to apples" comparisons across individuals using the raw monitoring frequencies. One intuitive way to create meaningful comparisons across individuals is to ask how each individual would react to a common stimulus. In our context we ask, at the average level of cooperation in the experiment how likely are you to monitor?

Recall that contributions for the current round are only revealed after paying the monitoring costs so it is sensible to model monitoring choices as depending on the lag of the other group member's contributions. Regardless of whether or not a participant monitored on round $t-1$, knowing the group total contribution from round $t-1$ and her own contribution allows her to infer how cooperative the other members of her group were last round. With this formulation in mind, we regress the decision to monitor in one round on the sum of the amount kept by the other members of one's experimental group in the previous round. We define $\Pr(M_{i,t}|\sum(15 - C_{-i,t-1})$ as the probability that participant i monitors in round t conditional on the amount kept by the rest of the group in the previous round and then estimate:

$$\Pr(M_{i,t}|\sum(15 - C_{-i,t-1}) = \beta_i^0 + \beta_i^1 \sum(15 - C_{-i,t-1}) + \epsilon_{i,t}, \tag{1}$$

where $\epsilon_{i,t}$ is a disturbance term. To get a sense of the pooled population response to the amount kept by the other group members, in the first two columns of Table 3, we report a linear probability model that controls to some extent for individual heterogeneity by including random effects. As one can see, overall, monitoring choices do appear to be affected by the level of cooperation in the group: each additional coin kept by a team member last round leads to an almost 1 percent increase ($p<0.01$) in the likelihood that the representative player monitors on the next round.

While it is interesting that the pooled regression confirms that monitoring does depend on lagged free-riding, for our purposes the individual heterogeneity of responses to free-riding is more interesting. Considering that rho, the test statistic for the fraction of the variation explained by the random effects, is clearly not zero ($p<0.01$), there does appear to be considerable heterogeneity in the response of individuals to free-riding. The question is, what is the best way to extract this heterogeneity from the data?

Given that we have seven rounds of monitoring choices for each player, one simple way is to run the regression separately for each individual (as in Carpenter and Seki, 2010). However, this method seems inefficient given we have responses for all 58 participants. In other words, running individual regressions allows us to take advantage of knowing the within subject variation in responses but it disregards the potential importance of the among subject variation. A standard way to incorporate both sources of variation is the random coefficients regressor developed in Swamy (1970).

In the second set of columns in Table 3 we report the results for the "typical" participant using the random coefficients model. Notice first that the results are similar to those generated by the random effects model. However, what is particularly important is that the Wald test for parameter consistency clearly indicates that responses are heterogeneous. Indeed, there are many participants who react strongly to the amount contributed by others—the more kept in round $t-1$, the more likely these people are to monitor in round t. However, at the same time, there are people who react oppositely by reducing their monitoring when the others keep more, as well as people who always or never monitor.

To summarize the heterogeneity in monitoring behavior, we use the individual estimates generated by the random coefficients model to predict the response to a common stimulus, the experiment average level of free-riding. Across all sessions and groups, on average, the other members of one's group kept 19.77 out of 45 possible coins. The distribution of the predicted probability of monitoring given this level of free-riding is given in Figure 2(b). Compared to the raw monitoring frequencies in Figure 2(a), we see that there is still considerable variation, but now there is a noticeably stronger mode of not monitoring. Despite the nuances of the second method, however, the two measures of the propensity to monitor are highly correlated (ρ=0.9, p<0.01).

6 Does Peer Monitoring Affect Loan Repayment?

In this section, we present our main results from testing whether behaviorally generated measures of inherent monitoring propensities predict group loan performance. Conceptually, it is important to first clarify the details of our loan data and the hypothesized link between our experimental data and these loan data. First, one advantage of our study is that endogeneity is less likely to affect our results than in previous studies. Not only do we examine loan performance in the six months *after* we collected the behavioral data, so that simultaneity should not bias our results, we also use monitoring propensities generated in an experimental vacuum (in that it is hard to imagine that individuals' unobservables could have affected these measures) and we collect many other factors that could affect repayment. One might also worry that shocks at the loan-group level will generate omitted variables bias when estimating the effect of group members' average monitoring propensity on individual repayment behavior.[7] However, it is equally unlikely that common group shocks are correlated with group members' monitoring in the experiment *and* with loan repayment behavior. Thus, there should be little or no omitted variables bias in our results.

Second, while we generated individual monitoring propensities in the previous section, we are not interested in the relationship between these, per se, and loan repayment. We are, however, interested in the relationship between the repayment behavior of borrower i and the propensity to monitor of the *other* people in her loan group. That is, we want to test whether the average monitoring propensity of the other group members predicts borrower i's chances of default.

Table 4 summarizes the loan data, some of which we collected during a follow up trip to Asunción six months after the experiment. During these six months, the four loan groups to which our participants belonged went through two or three loan cycles. This process generated data on 136 loans. The smallest loan during this period was for 150,000 PGY, the largest was for 900,000 PGY, and there was a strong mode at loans for 300,000 PGY; overall, the average loan was for 379,320 PGY which is equivalent to a little more than $64. Our interviews suggested that 90 percent of these loans were used for business purposes but the remaining loans were used for a variety of reasons including to smooth unexpected shocks, such as illnesses. In anticipation of this possibility, we also asked borrowers to tell us the number of unanticipated shocks that occurred during the loan cycle.

Another advantage of our study is that we were able to get some sense of the intrinsic creditworthiness of our participants. This information should also serve as a control for many unobservables. According to administrative data reported in Table 4, 26 percent of the loans were taken out by people registered in Paraguay's national bad debtor database. Because not everyone took out a loan in each period, the fraction of the women in the database is actually 23 percent.

[7] Manski (1993) outlines this common shocks problem in peer effects estimation

Lastly, because overt defaults are rare (and, as a consequence, there is little variation to explain), we adopted a method similar to Wydick (1999) to create our dependent variable. Based on administrative records and cross-reports from our participant interviews, we created an indicator, *Repayment Problem*, which is one if the borrower had trouble repaying the loan. This not only includes overt defaults, it also includes instances in which people needed extensions or had to have other members of the group help with repayment.

As mentioned above, we test the extent to which the monitoring propensities of the other members of the loan group affect the remaining member's behavior. To operationalize this test we created, for each loan, a variable that is the average monitoring propensity of everyone in the loan group except the borrower. To examine the robustness of our results we create three versions of this variable. The first is based on the raw monitoring frequencies of our participants. The second is based on the predicted probability of monitoring at the experiment average level of free-riding. The third also uses the predicted monitoring probability from the random coefficients estimates but is a non-linear transformation of the second version. Specifically, we decided that it also would be interesting to examine the relationship between loan repayment and the predicted probability of not being monitored, which is just the product of one minus the predicted monitoring probability for each of the other loan group members.[8]

As a first pass at the analysis, in Figure 3 we break our three versions of the intensity of peer monitoring variable at the median to see if being in a loan group with inherently nosier people has an affect on the chances of incurring repayment problems. As panel (a) suggests, those borrowers who are likely to be highly monitored are approximately 10 percent less likely to have repayment problems. However, as one can see from the 95 percent confidence intervals and according to a simple two-sided t-test, the difference is not significant (t=1.20, p=0.23). Moving to panel (b), which is based on the predicted probability of monitoring, we see that the gap widens slightly and the confidence intervals shrink. Here the 13 percent difference in the raw probability is now marginally significant (t=1.78, p=0.07). Lastly, as expected, panel (c) indicates that those borrowers who are predicted to be the most likely to not be monitored are more likely to have repayment problems. Again the 11 percent difference is on the verge of significance (t=1.59, p=0.11).

While simple t-tests can be informative, unobserved shocks could affect both monitoring and repayment in a particular group or during a particular loan cycle, and these shocks might bias our estimates of the effect of peer monitoring. With this in mind, Table 5 records the results of probit estimates of the effect of being in highly monitored groups. As one can see, using the first two measures of peer monitoring, the point estimates roughly double in size compared to the t-tests and are significant. According to our simple measure, borrowers in highly monitored groups are 26 percent less likely to have a repayment problem (p=0.07). Using instead the predicted probability of monitoring leads to a similar estimate: here the highly monitored are 22 percent less likely to have a problem (p=0.06). It does not appear, however, that unobserved shocks have much of an effect on our estimates of the effect of being unmonitored. In the last two columns of Table 5 we see that the estimate remains

[8]Of course we had to deal with the fact that some people monitor always and therefore the residual is zero. We assumed that there was some small error (0.01) to perturb our estimates. Experimenting with different small error values does not affect the predicted probabilities of not being monitored appreciably.

near 10 percent.

When we further "ratchet up" the analysis in Table 6, we see that the results strengthen considerably. To be consistent with Figure 3, we continue to report the marginal effects from probit regressions of the *Repayment Problem* indicator on the three indicators for the intensity of peer monitoring (Table 7 shows that the results are very similar when we allow for continuous monitoring intensity measures). In addition, we not only include loan cycle and loan group fixed effects, we also add all the other factors that might affect repayment.

Considering only the first three rows of Table 6 (i.e., postponing discussion of all the other interesting results until the next section), we see that our estimates again increase substantially in both magnitude and significance. Now, based on the raw monitoring frequencies, those borrowers in groups with greater than median levels of expected monitoring are approximately half as likely to have repayment problems ($p<0.01$). The point estimate falls slightly when we use the predicted monitoring probabilities but it is still substantial: highly monitored borrowers are now 36 percent less likely to get in trouble ($p<0.01$). Lastly, when we "invert" the analysis and ask what happens to repayment when the chances of not being monitored increase, we find that those with higher than median chances of not being monitored are 40 percent more likely to incur repayment problems ($p<0.01$). In sum, our results suggest that there is a strong and robust empirical relationship between the propensity of peers to monitor each other and the performance of the group's loan.

7 Other Factors Affecting Loan Performance

Although our main purpose is to test the link between peer monitoring and loan repayment, our data allow us to explore a number of other interesting facts about group loan performance. Returning to Table 6, we see that in addition to a smaller number of interesting, but less robust correlations there are quite a few important and robust results. Because our data is measured on a number of difference scales, to foster comparisons of the magnitude of the point estimates, we have standardized all the continuous marginal effects.[9]

To begin, we discuss the effect of different properties of the loan. For example, perhaps because of selection (we control for experience), we see that larger loans are less likely to run into trouble, but the effect is not significant. The number of reported adverse "shocks" also appears to have no significant effect. What is significant, however, is that business loans are less likely to have problems than loans made, for example, in emergency situations. The effect also appears to be large. In model (2), where the effect is smallest and not quite significant, business loans are 13 percent less likely to have repayment problems; this estimate grows to 27 percent in model (1) ($p<0.05$). By far, the strongest predictor among the loan characteristics is being in the bad debtor database. Borrowers who have defaulted before are between 48 and 67 percent more likely to have recurring repayment problems.

There are fewer demographic predictors of loan repayment. As measured by our survey, age, education, home ownership, and years of business experience all seem to have little effect on loan performance. That said, there are a number of factors that suggest that having more financial resources does mitigate repayment problems. Married women, for

[9] Although we realize that linear approximations in probit models are often a problem, we decided to report standardized marginal effects because the effects were otherwise very hard to compare.

example, are between 22 and 35 percent less likely to have repayment problems, perhaps because they can draw on two sources of income.[10] It is also the case that having more family income and savings reduces the chances of repayment problems.

Moving to the cognitive and preference-based influences, we first see a robust effect of IQ. A standard deviation increase in our IQ measure is associated with between a 0.10 and a 0.23 standard deviation decrease in the probability of having trouble repaying the loan. While we do not see significant effects of time preferences, we do see that very risk averse borrowers (those who choose the safe option in our experiment) are between 9 and 13 percent less likely to get in repayment trouble, In addition, our measure of altruism, giving in the dictator game, also plays some role in repayment. As hypothesized, altruists may worry more about imposing a negative externality on the rest of the loan group, which is borne out in our data. A standard deviation increase in dictator giving is associated with between a 0.10 and a 0.20 standard deviation reduction in the probability of repayment being a problem.

Lastly, we can also explore the effect of the characteristics of the loan group. There appears to be no effect of community engagement on repayment behavior so either one's standing in the community plays no role or we have not measured it precisely enough. We also see that having more friends in the group is associated with fewer repayment problems, but the effect is small and insignificant. What is somewhat surprising, however, is that there appears to be two levels of the moral hazard problem in groups with multiple family members. Our estimates suggest that a standard deviation increase in the number of family members in the group increases the chances that one of them will have a repayment issue by as much as 0.58 of a standard deviation.[11] It might be the case that borrowers know that they can depend more on other family members to repay their loans and this affects the investment and repayment choices that they make.

8 Concluding Remarks

Our results come closest to testing the key assumption of many models of group lending: peer monitoring can solve the moral hazard problem in lending to people who have no appreciable collateral. We find that the more inherently "nosey" people there are in your loan group, the less likely you are to have troubles repaying your loan. We prefer the simple interpretation that nosey people look over the fences at each other and discourage moral hazard. To a large extent, this simple interpretation may also be warranted given our research design. Because the loan program women who participated in our field experiment did not know who they were monitoring, their decisions could not be conditioned on individual characteristics and, in this sense, our measures of the propensity to monitor are inherent. On top of this we control for many of the other reasons why people might get into trouble or why one might want to monitor them and still find a strong effect of inherent "nosiness."

Although the purpose of this study is to establish an empirical link between peer monitoring and loan performance after loan groups have been formed, unfortunately we are

[10] Recall that long-term cohabitation is also common in Paraguay, and, therefore, marriage may also proxy for other characteristics, such as religiosity.

[11] It is interesting that our "family member" effect is similar to the results described in Ahlin and Townsend (2007).

somewhat silent on the possibility that groups form in anticipation of monitoring. Perhaps, for example, it is the case that people self-select into groups with people that they think will monitor more or less. While it will surely be interesting to learn more about group selection, it is not clear that this possibility necessarily diminishes our results. To begin, notice that such behavior does not negate the fact that peer monitoring appears to be a strong mechanism that reduces moral hazard. On top of this, if people were so good at selecting into "optimal" groups, they would anticipate the large negative effect of joining groups with people in the bad debtors' database or groups with other family members, to name just two examples. Given that we see plenty of women joining groups with their kin it is not clear that selection is that sophisticated.

9 Appendix A – Instructions for the Social Dilemma Experiment

(Note: back translations from Spanish)

Instructions for the participants:

Thank you very much for your participation today. I will pay you for participating. The amount that you receive will depend on your decisions and the decisions of the others in the study.

All your decisions and answers to questions in the experiment will be completely confidential. In order to ensure total confidentiality, I ask that you do not speak amongst yourselves before the end of the experiment.

Instructions for the game:

In order to understand this experiment, think about how you use your time. You use one part for activities that only help yourself and/or your family. You use another part to do things that help everyone in your community.

This experiment should be similar to a situation where you have to decide between doing something only for yourself and doing something to help the whole community. For example, imagine that you have a little extra money, and you can use it to help pay for a business training program in the community center or to rent a better space for your business. If you spend the money on the training program, everyone in the community will benefit from the program, whether they help pay for it or not. But, if instead of this, you rent the new space, only you will benefit. This activity is similar to this decision.

There will be eight rounds of decision making. There are three other women in a group with you. You will not know their identities either during or after the activity.

At the beginning of each round, each person will receive fifteen 100 Guaraní coins. This money is yours to be kept after the game. However, there is also a decision involving the money. Each person in the group will decide how much money she wants to give to a group project (like in the example above), and how much she wants to keep. The whole group will benefit equally from the money contributed to the group project, but the money you keep will only benefit you.

When the four members in each group have decided how much money (of the 1500 PGY possible) they want to contribute to the group project, I will add these contributions together. When I know the total, I will add 50%. For example, if the total contribution from the group is 2000 PGY, then I will add 1000 PGY to make a final total of 3000 PGY. After this, each person in the group will receive an equal part of the final amount (in this case, each person would receive 750 PGY). If contributed evenly (i.e., 500 PGY each), then each person would receive 1000 PGY plus 750 PGY for their final income (1750 PGY).

For another example, if the four group members contribute 0, 0, 0, and 1500 PGY, respectively, then I will add 750 PGY to make the final total 2250 PGY. After this, each person in the group will receive an equal part, 563, of the final amount. Therefore, the person who contributes all their money will receive 563 PGY for their total income, while the other three will receive 2063 PGY (which is 563 PGY plus 1500 PGY).

One last example: if all four group members contribute 1500 PGY, then I will add 3000 PGY to make a final total of 9000 PGY. After this, each person in the group will receive an equal part, 2250 PGY, of the final amount. Since everyone contributed everything they had, this will be everyone's final total income as well.

In order to contribute money to the group project, I will give you two envelopes. One envelope will say "only for you," on it and the other will say "for the group project." To begin, all of your 1500 PGY will be in the "only for you" envelope. Each one of you will remain at your seat, and from this envelope, you will remove your contribution to the group project for that round and place it in the "group project" envelope. After, I will count how much money was given for the group project (from all four players). Then, I will calculate how much you each will receive from the project. Finally, I will add this to the money that you kept for yourself. This is your total income for the round. To make things easier, I will keep all the money, but mark down everyone's total. Then you will receive your 8-round total at the end.

After I calculate your income in each round, each person will meet with me individually. In this meeting, I will show you the total for the group project and your total income. Also, you can pay 50 PGY to see a card with the contributions and total income from the round (in random order) for the others in your group. If you decide to spend it, I will subtract this 50 PGY from your income for the round. Lastly, if you are unhappy with the contribution of someone, you can pay (from your income for the round again) 50 PGY per message to send a message of disapproval (*show and explain the unhappy face message*).

That is, if you want to send a message to one person that you are unhappy with their contribution, you have to pay 50 PGY; to send messages to two people, 100 PGY; for three, 150 PGY. In order to send a message to someone, simply indicate the person(s) when you are shown the card with everyone's contributions and incomes. There is no need to write a message or deliver a message yourself (I will perform these tasks) and you can only send a message if you chose to pay to see the contributions of the others.

At the beginning of the next round, I will give everyone back their envelopes, and for the next round, you may change your contribution decision by taking coins out of the "group project" envelope or by adding more from the "only for you" envelope or you may leave it the same. Inside the "group project" envelope, there will also be a card showing any disapproval messages that have been sent to you. If the card is blank, then no one decided to send you a message. All of you will have a number in your group, the same number for all eight rounds. The messages will say, for example, "From player 2 to you" (*write on the board and show with unhappy face*). Only the sender and the receiver will know that there was a message, and there is no obligation to change your decision if you receive a message.

It is very important that you understand all these instructions. Are there any questions about how the activity will work? You can ask me for any clarifications during the activity as well.

10 References

Abbink, K., Irlenbusch, B., and Renner, E., 2006. Group Size and Social Ties in Microfinance Institutions. *Economic Inquiry.* 44(4), 614–628.

Ahlin, C., and Townsend, R., 2007. Using Repayment Data to Test Across Models of Joint Liability Lending. *The Economic Journal.* 117(517), 11–51.

Armendariz de Aghion, B., 1999. On the Design of a Credit Agreement with Peer Monitoring. *Journal of Development Economics.* 60(1), 79–104.

Armendariz de Aghion, B., and Morduch, J., 2005. *The Economics of Microfinance.* The MIT Press: Cambridge, MA, and London.

Banerjee, A., Duflo, E., Glennerster, R., and Kinnan, C., 2009. The Miracle of Microfinance? Evidence from a Randomized Evaluation. Jameel Poverty Action Lab working paper.

Barboza, G., and Barreto, H., 2006. Learning by Association: Micro Credit in Chiapas, Mexico. *Contemporary Economic Policy.* 24(2), 316–331.

Besley, T., and Coate, S., 1995. Group Lending, Repayment Incentives and Social Collateral. *Journal of Development Economics.* 46, 1–18.

Burks, S.V., Carpenter, J.P., Goette, L., and Rustichini, A., 2009. Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment. *Proceedings of the National Academy of Sciences.* 106(19), 7745–7750.

Cardenas, J.C., and Carpenter, J., 2009. Risk Attitudes and Well-Being in Latin America. Middlebury College, Department of Economics, working paper.

Carpenter, J., Burks, S., and Verhoogen, E., 2005. Comparing Students to Workers: The Effects of Social Framing on Behavior in Distribution Games, In: J. Carpenter, G. Harrison, and J. List (Eds.), *Field Experiments in Economics, Research in Experimental Economics.* JAI/Elsevier: Greenwich, CT, and London, pp. 261–290.

Carpenter, J., Daniere, A., and Takahashi, L., 2004a. Cooperation, Trust, and Social Capital in Southeast Asian Urban Slums. *Journal of Economic Behavior & Organization.* 55(4), 533–551.

Carpenter, J., Daniere, A., and Takahashi, L., 2004b. Social Capital and Trust in Southeast Asian Cities. *Urban Studies.* 41(4), 853–874.

Carpenter, J., and Seki, E., 2010. Do Social Preferences Increase Productivity? Field Experimental Evidence from Fishermen in Toyama Bay. *Economic Inquiry.* Forthcoming.

Cason, T., Gangadharan, L., and Maitra, P., 2009. Moral Hazard and Peer Monitoring in a Laboratory Microfinance Experiment. Purdue University, Department of Economics, working paper.

Cassar, A., Crowley, L., and Wydick, B., 2007. The Effect of Social Capital on Group Loan Repayment: Evidence from Field Experiments. *The Economic Journal.* 117, 85–106.

Feigenberg, B., Field, E., and Pande, R., 2009. Do Social Interactions Facilitate Cooperative Behavior? Evidence from a Group Lending Experiment in India. Jameel Poverty Action Lab working paper.

Forsythe, R., Horowitz, J., Savin, N.E., and Sefton, M., 1994. Fairness in Simple Bargaining Experiments. *Games and Economic Behavior*. 6, 347–369.

Gine, X., Jakiela, P., Karlan, D., and Morduch, J., 2009. Microfinance Games. *American Economc Journal: Applied Economics*. Forthcoming.

Gine, X., and Karlan, D., 2008. Peer Monitoring and Enforcement: Long Term Evidence from Microcredit Lending Groups With and Without Group Liability. Working paper.

Gomez, R., and Santor, E., 2008. Does the Microfinance Lending Model Actually Work? *Whitehead Journal of Diplomacy and International Relations*. 9(2), 37–56.

Hermes, N., Lensink, R., and Mehrteab, H., 2005. Peer Monitoring, Social Ties and Moral Hazard in Group Lending Programs: Evidence from Eritrea. *World Development*. 33(1), 149–169.

Karlan, D., 2005. Using Experimental Economics to Measure Social Capital and Predict Financial Decisions. *American Economic Review*. 95(5), 1688–1699.

Karlan, D., 2007. Social Connections and Group Banking. *The Economic Journal*. 117(517), 52–84.

Kono, H., 2006. Is Group Lending a Good Enforcement Scheme for Achieving High Repayment Rates? Evidence from Field Experiments in Vietnam. Institute of Developing Economies, Japan External Trade Organization working paper No. 61.

Kritikos, A., and Vigenina, D., 2005. Key Factors of Joint-Liability Loan Contracts: An Empirical Analysis. *Kyklos*. 58(2), 213–238.

Ledyard, J., 1995. Public Goods: A Survey of Experimental Research. In: J. Kagel and A. Roth (Eds.), *The Handbook of Experimental Economics.* Princeton University Press: Princeton, NJ, pp. 111–194.

Manski, C.F., 1993. Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*. 60, 531–542.

Masclet, D., Noussair, C., Tucker, S., and Villeval, M.-C., 2003. Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism. *American Economic Review*. 93(1), 366–380.

Meier, S., and Sprenger, C., 2009. Present-Biased Preferences and Credit Card Borrowing. *American Economic Journal: Applied Economics*. Forthcoming.

Rai, A.S., and Sjostrom, T., 2004. Is Grameen Lending Efficient? Repayment Incentives and Insurance in Village Economies. *Review of Economic Studies*. 71(1), 217–234.

Raven, J., Raven, J.C., and Court, J.H., 2003. *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Harcourt Assessment: San Antonio, TX.

Simtowe, F., and Zeller, M., 2007. Determinants of Moral Hazard in Microfinance: Empirical Evidence from Joint Liability Lending Programs in Malawi. MPRA Paper No. 461.

Stiglitz, J., 1990. Peer Monitoring and Credit Markets. *World Bank Economic Review*. 43, 351–366.

Swamy, P.A.V.B., 1970. Efficient Inference in a Random Coefficient Regression Model. *Econometrica*. 38(2), 311–323.

Wydick, B., 1999. Can Social Cohesion Be Harnessed to Repair Market Failures? Evidence from Group Lending in Guatemala. *The Economic Journal*. 109(457), 463-475.

11 Tables and Figures

TABLE 1: Descriptive Statistics on the Participants

Variable	Description	N	Mean	Std. Dev.
Demographics:				
Age	Participant's age	58	37.07	11.35
Elementary (I)	1 if elementary school is the highest level of education attained	58	0.57	0.50
Married (I)	1 if married	58	0.60	0.49
Socio-Economic:				
Head of Household (I)	1 if the participant has the most influence in the household	58	0.26	0.44
Family Income	Participant's monthly family income (in thousand Guarani)	58	1725.42	1332.29
Savings	Participant's savings (in thousand Guarani)	58	24.65	132.49
Own Home (I)	1 if participant (or family) owns her home	58	0.91	0.28
Business Experience	Participant years of experience in current work	58	7.70	9.17
Cognitive and Behavioral:				
IQ	The number (out of 3) non-verbal questions answered correctly	58	0.41	0.65
Risk Seeking (I)	1 if picked the 6th (i.e., most risky) lottery in the risk task	58	0.14	0.35
Risk Averse (I)	1 if picked the 1st (i.e., safe) lottery in the risk task	58	0.10	0.31
Patient (I)	Choices in the discounting task consistent with low discounting	58	0.26	0.44
Impatient (I)	Choices in the discounting task consistent with high discounting	58	0.17	0.38
Altruism (Dictator Giving)	The amount (out of 30k Guarani) allocated to a stranger	58	12.84	5.22
Lending Group Related:				
Community Engagement	The number of community groups affiliated with	58	0.94	1.13
Family Members	The number of family members in the participant's loan group	58	0.98	1.26
Friends	The number of friends originally in the participant's loan group	58	4.15	3.29

TABLE 2: Descriptive Statistics from the Social Disapproval Experiment
(pooling over all rounds)

Variable	Description	N	Mean	Std. Dev.
Contribution	Number of 100 Guarani coins contributed (15 coin endowment)	448	8.40	3.82
Monitor (I)	1 if the participant chose to monitor during the round	399	0.47	0.50
Messages Sent	Number of disapproval messages sent during the round (3 possible)	399	0.14	0.39

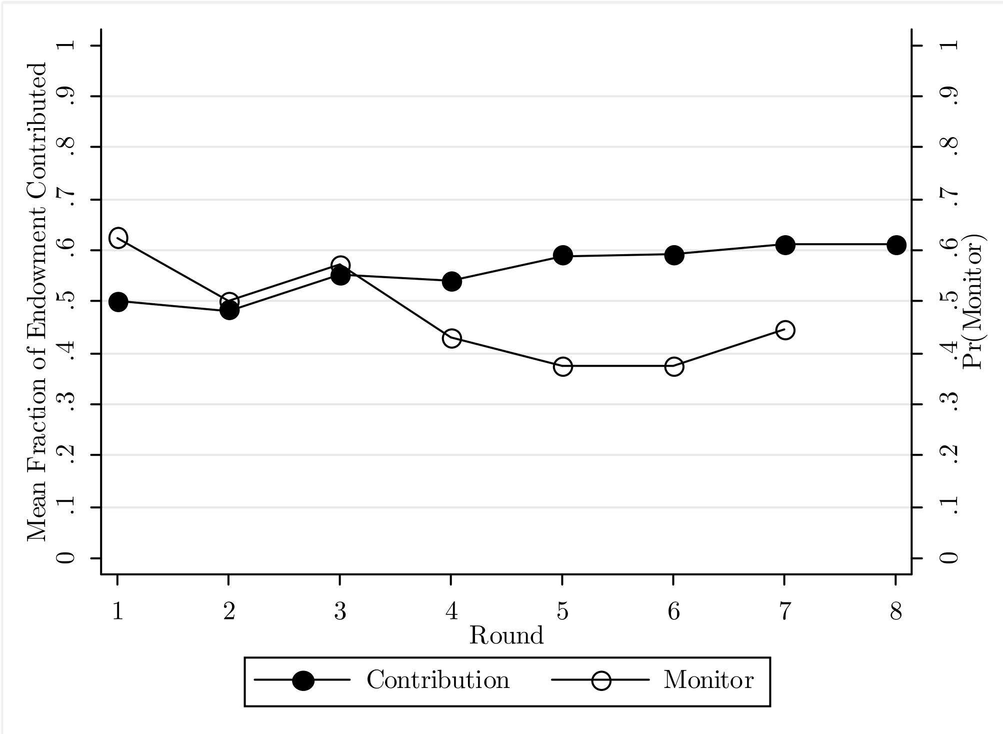

Figure 1: Experimental behavior (by round).

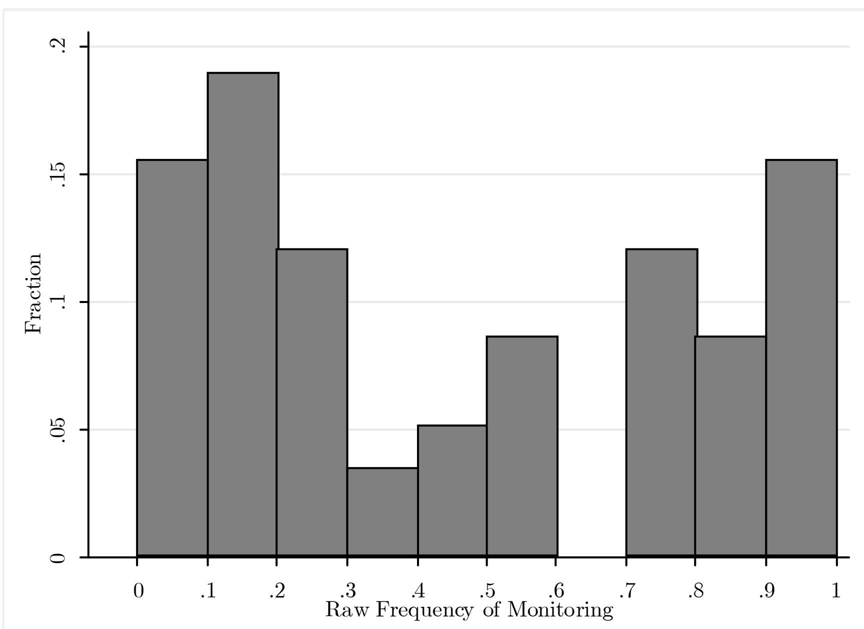

Figure 2(a): Raw monitoring frequencies.

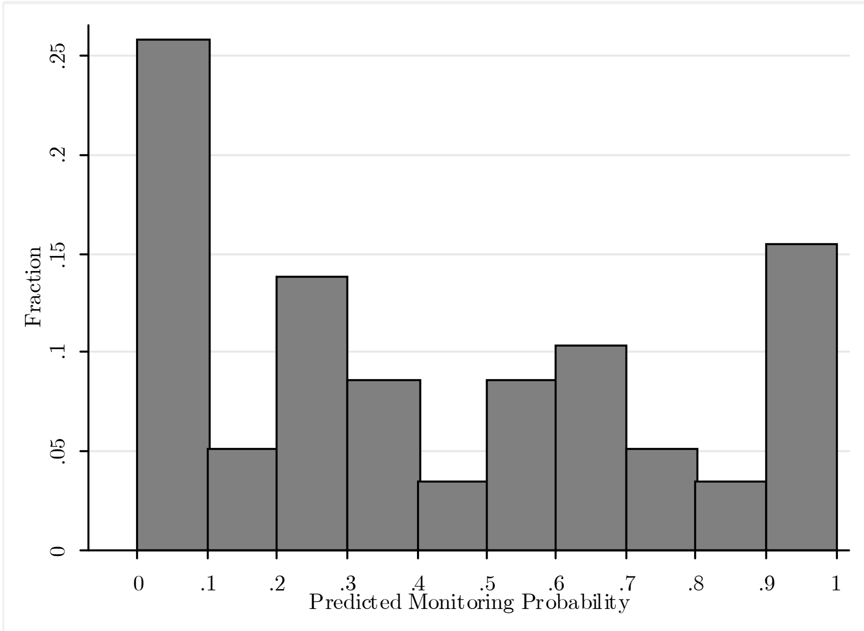

Figure 2(b): Predicted monitoring probabilities.

TABLE 3: Creating Peer Monitoring Preference Measures

	Random Effects		Random Coefficients	
	coef.	s.e.	coef.	s.e.
$\sum \text{kept}_{-i,t-1}$	0.008	(0.003)***	0.007	(0.008)
intercept	0.241	(0.078)***	0.254	(0.203)
random effects	yes		no	
N	400		400	
rho	0.43			
Wald chi^2, p-value	7, p<0.01		180, p<0.01$^+$	

Notes: Results are from linear probability models; rho is the fraction of the variance accounted for by the individual random effects; *** significant at 1%, ** 5%, *10%. $^+$Test for parameter constancy.

TABLE 4: Descriptive Statistics on Loan Activity				
Variable	Description	N	Mean	Std. Dev.
Loan Amount	The amount borrowed (in thousand Guarani)	136	379.32	160.86
Business Loan (I)	1 for loans to enhance one's business (versus emergency loans)	136	0.90	0.31
Adverse Shocks	Number of unexpected costly events during the cycle (e.g., illness)	136	0.73	1.07
Informconf (I)	1 for borrowers in Paraguay's national loan default database	136	0.26	0.44
Repayment Problem (I)	1 for borrowers with repayment problems (based on administrative records)	136	0.25	0.43

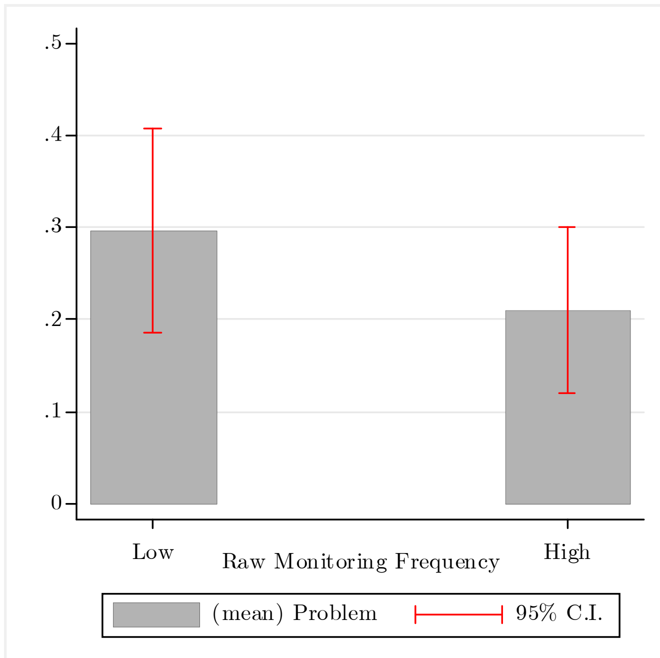

Figure 3(a): Repayment problems and monitoring frequency.

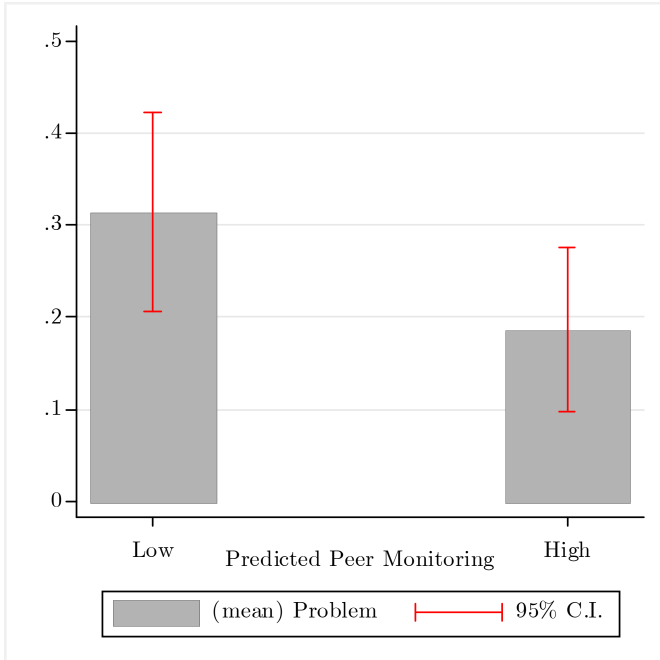

Figure 3(b): Repayment problems and the predicted probability of monitoring.

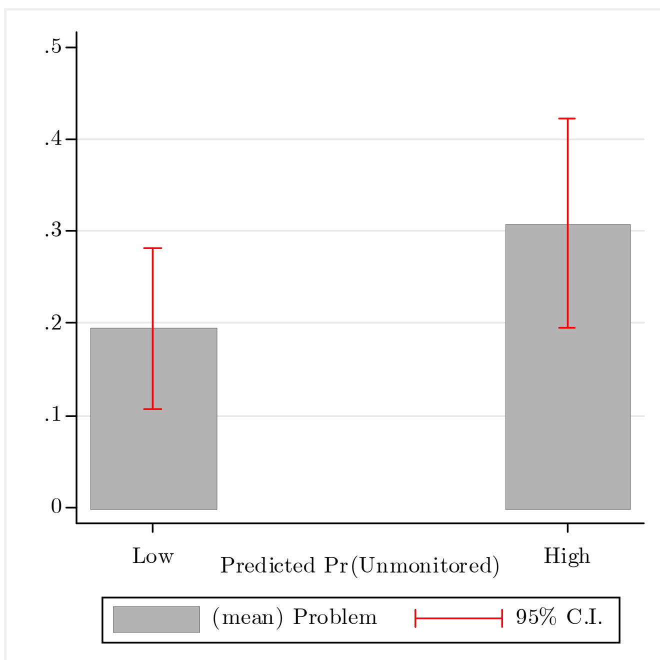

Figure 3(c): Repayment problems and the probability of being unmonitored.

TABLE 5: Loan Repayment Problems Fall with more Peer Monitoring						
	(1)		(2)		(3)	
High Peer Monitoring Frequency (I)	-0.263*	(0.151)				
High Predicted Pr(Peer Monitoring) (I)			-0.219*	(0.115)		
High Predicted Pr(Unmonitored) (I)					0.092	(0.168)
Loan Cycle Fixed Effects	Yes		Yes		Yes	
Loan Group Fixed Effects	Yes		Yes		Yes	
Wald Chi^2	4.34		5.36		1.42	
Pseudo R^2	0.05		0.05		0.01	
Observations	136		136		136	

Notes: The dependent variable is whether or not the borrower had a repayment problem; marginal effects reported from probit estimates; (robust standard errors clustered on the individual); * p<0.10, ** p<0.05, *** p<0.01.

TABLE 6: Peer Monitoring and Loan Payment Problems (with indicator monitoring measures)						
	(1)		(2)		(3)	
High Peer Monitoring Frequency (I)	-0.454***	(0.121)				
High Predicted Pr(Peer Monitoring) (I)			-0.357***	(0.102)		
High Predicted Pr(Unmonitored) (I)					0.401***	(0.152)
Loan Amount	-0.170	(0.000)	-0.111	(0.000)	-0.122	(0.000)
Business Loan (I)	-0.273**	(0.151)	-0.135	(0.119)	-0.175*	(0.133)
Adverse Shocks	-0.010	(0.019)	-0.029	(0.015)	-0.014	(0.022)
Informconf (I)	0.570***	(0.224)	0.671***	(0.214)	0.478**	(0.219)
Age	-0.022	(0.004)	0.018	(0.003)	-0.001	(0.005)
Elementary Education (I)	-0.091	(0.098)	-0.074	(0.080)	-0.087	(0.125)
Married (I)	-0.350**	(0.136)	-0.223**	(0.133)	-0.301**	(0.147)
Head of Household (I)	0.207*	(0.150)	0.153	(0.138)	0.228	(0.163)
Family Income	-0.186**	(0.000)	-0.187***	(0.000)	-0.266***	(0.000)
Savings	0.204**	(0.001)	0.123**	(0.001)	0.071	(0.001)
Own Home (I)	-0.095	(0.186)	0.041	(0.037)	-0.024	(0.160)
Business Experience	0.001	(0.004)	-0.109	(0.003)	-0.031	(0.006)
IQ	-0.176**	(0.062)	-0.098*	(0.047)	-0.235**	(0.076)
Risk Seeking (I)	-0.060	(0.070)	0.010	(0.077)	-0.034	(0.108)
Risk Averse (I)	-0.125***	(0.053)	-0.087***	(0.049)	-0.134**	(0.053)
Patient (I)	0.067	(0.095)	0.084	(0.103)	0.109	(0.115)
Impatient (I)	0.055	(0.124)	0.194	(0.175)	0.102	(0.147)
Altruism (Dictator Giving)	-0.198***	(0.008)	-0.109**	(0.007)	-0.105	(0.008)
Community Engagement	0.080	(0.041)	-0.004	(0.026)	-0.018	(0.058)
Family Members in Loan Group	0.579***	(0.054)	0.462***	(0.062)	0.544***	(0.065)
Friends in Loan Group	-0.040	(0.013)	-0.091	(0.010)	-0.005	(0.018)
Loan Cycle Fixed Effects	Yes		Yes		Yes	
Loan Group Fixed Effects	Yes		Yes		Yes	
Wald Chi^2, p-value	111, <0.01		151, <0.01		84, <0.01	
Pseudo R^2	0.44		0.50		0.35	
Observations	136		136		136	

Notes: The dependent variable is whether or not the borrower had a repayment problem; marginal effects reported from probit estimates; the effects of the continuous variables have been standardized to ease comparisons; (robust standard errors clustered on the individual); * p<0.10, ** p<0.05, *** p<0.01.

TABLE 7: Peer Monitoring and Loan Payment Problems (with continuous monitoring measures)								
	(1)		(2)		(3)		(4)	
Frequency of Peer Monitoring	-0.340*	(1.343)						
Predicted Pr(Peer Monitoring)			-0.553**	(1.279)				
Predicted Pr(Unmonitored)					-0.055	(60.769)	0.115*	(725.691)
Loan Amount	-0.126	(0.000)	-0.138	(0.000)	-0.125	(0.000)	-0.159	(0.000)
Business Loan (I)	-0.232*	(0.146)	-0.252**	(0.152)	-0.165	(0.130)	-0.210*	(0.127)
Adverse Shocks	-0.018	(0.021)	-0.017	(0.021)	-0.016	(0.022)	-0.027	(0.022)
Informconf (I)	0.492***	(0.207)	0.506***	(0.198)	0.473**	(0.218)	0.616***	(0.229)
Age	0.020	(0.005)	0.053	(0.005)	-0.004	(0.006)	-0.023	(0.005)
Elementary Education (I)	-0.108	(0.120)	-0.125	(0.122)	-0.086	(0.126)	-0.036	(0.124)
Married (I)	-0.323**	(0.146)	-0.309**	(0.147)	-0.293**	(0.148)	-0.319**	(0.152)
Head of Household (I)	0.266*	(0.161)	0.308**	(0.169)	0.221	(0.164)	0.118	(0.145)
Family Income	-0.240**	(0.000)	-0.259***	(0.000)	-0.264**	(0.000)	-0.301***	(0.000)
Savings	0.112	(0.001)	0.128	(0.001)	0.083	(0.001)	0.078	(0.001)
Own Home (I)	-0.042	(0.150)	-0.070	(0.156)	-0.023	(0.161)	0.067	(0.081)
Business Experience	-0.082	(0.005)	-0.101	(0.005)	-0.030	(0.006)	-0.050	(0.005)
IQ	-0.242**	(0.068)	-0.215**	(0.064)	-0.239**	(0.078)	-0.181*	(0.079)
Risk Seeking (I)	0.014	(0.141)	-0.002	(0.119)	-0.036	(0.107)	0.079	(0.172)
Risk Averse (I)	-0.124**	(0.052)	-0.118**	(0.051)	-0.134**	(0.054)	-0.129**	(0.058)
Patient (I)	0.078	(0.105)	0.112	(0.114)	0.105	(0.116)	0.064	(0.100)
Impatient (I)	0.138	(0.162)	0.210	(0.184)	0.098	(0.146)	0.083	(0.142)
Altruism (Dictator Giving)	-0.169*	(0.008)	-0.174*	(0.008)	-0.105	(0.008)	-0.096	(0.007)
Community Engagement	0.008	(0.056)	0.007	(0.053)	-0.013	(0.059)	-0.024	(0.056)
Family Members in Loan Group	0.569***	(0.060)	0.566***	(0.059)	0.540***	(0.064)	0.612***	(0.066)
Friends in Loan Group	-0.005	(0.016)	0.007	(0.015)	-0.003	(0.018)	-0.155	(0.020)
Loan Cycle Fixed Effects	Yes		Yes		Yes		Yes	
Loan Group Fixed Effects	Yes		Yes		Yes		Yes	
Wald Chi^2, p-value	102, <0.01		122, <0.01		85, <0.01		162, <0.01	
Pseudo R^2	0.37		0.39		0.35		0.38	
Observations	136		136		136		132	

Notes: The dependent variable is whether or not the borrower had a repayment problem; marginal effects reported from probit estimates; the effects of the continuous variables have been standardized to ease comparisons; (robust standard errors clustered on the individual); * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Because of the nonlinear process used to create the "unmonitored" variable, two participants were calculated to have probabilites that are orders of magnitude higher than the others. These outliers are removed in model 4.