ECONOMIC GROWTH CENTER
YALE UNIVERSITY
P.O. Box 208629
New Haven, CT 06520-8269
http://www.econ.yale.edu/~egcenter/

CENTER DISCUSSION PAPER NO. 913

Social Connections and Group Banking

Dean S. Karlan
Princeton University
and
Yale University

June 2005

Notes: Center Discussion Papers are preliminary materials circulated to stimulate discussions and critical comments.

I thank my advisors Abhijit Banerjee, Dora Costa, Esther Duflo, and Sendhil Mullainathan for their guidance throughout this research. I also thank Alexander Aganine, Beatriz Armendariz de Aghion, Robin Burgess, Ben Jones, Leigh Linden, Norman Loayza, Cade Massey, Jonathan Morduch, Ashok Rai, Laura Schechter, Richard Thaler, Ashley Timmer, Chris Udry, Justin Wolfers and Eric Zitzewitz, and participants at the 2001 NEUDC, the 2001 LACEA PEG/NIP Conference, and 2004 ASSA, and at seminars at Harvard/M.I.T., Princeton, Yale, Maryland, Texas, Williams, Johns Hopkins, UC-Irvine, Berkeley, UCLA, University of Washington, Michigan, Georgetown, Miami, and the Univeristy of Natal, South Africa. Last, but not least, thanks to Iris Lanao, Aquiles Lanao, and Morena Lanao from FINCA and my field team, particularly Alcides Medina, Fatima Oriundo, and Jeny Yucra. The research reported herein was supported by the SSRC, the Russell Sage Foundation, the M.I.T. George Shultz Fund, and the Center for Retirement Research at Boston College pursuant to a grant from the Social Security Administration. All views and errors are mine.

Social Connections and Group Banking

Dean S. Karlan

Abstract

Lending to the poor is expensive due to high screening, monitoring, and enforcement costs. Group lending advocates believe lenders overcome this by harnessing social connections. Using data from FINCA-Peru, I exploit a quasi-random group formation process to find evidence of peers successfully monitoring and enforcing joint-liability loans. Individuals with stronger social connections to their fellow group members (i.e., either living closer or being of a similar culture) have higher repayment and higher savings. Furthermore, I observe direct evidence that relationships deteriorate after default, and that through successful monitoring, individuals know who to punish and who not to punish after default.

Keywords: Microfinance, Group lending, Informal savings, social capital

JEL Codes: O12, O16, O17, Z13

Lending to the poor is a difficult task throughout the world, as attested to by the many projects that experience high default rates. Starting with the Grameen Bank in Bangladesh and FINCA village banking in Latin America, development policymakers have embraced group lending as a possible alternative for lenders to provide credit to the poor. Group lending typically links the fate of borrowers by stipulating that if one borrower within a group fails to repay her loan, the others in her group must repay it for her. This potentially works for a few reasons (which all rely on social connections): individuals are able to select creditworthy peers, are able to monitor each others' use of funds and ability to repay, are able to enforce repayment, or perhaps are more likely to repay merely because of altruism towards those in their group.

I test whether groups that are more connected socially perform better, and specifically whether this is a *causal* relationship from *ex-post* contract monitoring and/or enforcement. The empirical tests employed rule out selection or unobserved dimensions, such as economic opportunities, that coincide with social connections. FINCA-Peru has a group formation process that generates a natural experiment in which some groups are endowed with stronger social connections than others. I find that stronger social connections of the group lead to improved *ex-post* repayment and savings behavior by the clients.

While theoretical models have described the potential of group lending, little empirical evidence has been found to understand if and how group lending actually improves repayment rates [see Banerjee, Besley and Guinnane (1994), Besley and Coate (1995), Ghatak (2000), Stiglitz (1990) and Varian (1990)]. Advocates of group lending not only argue that it does, but offer an explanation as to how this is accomplished: by taking advantage of the social networks and relationships. As Varian writes in a 2001 *New York Times* article, "Peer pressure can be an immensely strong force, and the Grameen Bank has figured out how to make it work in the cause of economic development."

Group lending uses the borrower's personal reputation in much the same way as physical collateral is used under ordinary lending: to raise the borrower's cost of defaulting [see van Bastelaer (1999)]. Furthermore, the more valuable the social connections are, the larger the stakes, and thus the higher the repayment rate is.[1] Individuals with stronger or more extensive social connections also can collect better information about other group members. With this information advantage, they are in an even better position to determine who is creditworthy, to monitor ability to repay, and to punish in the case of default without cause. Rai and Sjostrom (2004) develops a model that highlights how group lending can create incentives to observe negative shocks, and hence know who to punish and who to forgive. I find empirical support for this particular model. Lastly, groups with stronger social connections may simply have stronger feelings of altruism towards each other. Group liability may embrace pre-existing social relationships, or might create incentives for borrowers to form such relationships in order to have a successful group. Either way, social connections are helping to generate

[1]As van Bastelaer [1999] discusses, these organizations provide credit on the basis of "social collateral," through which borrowers' reputations, or the social networks to which they belong, take the place of traditional physical or financial collateral.

successful lending outcomes.

In this paper, I define social connections as the links and commonalities that bind a group of people together and determine their social interactions.[2] Social connections in this context can be thought of as a broader form of social capital, one that encompasses the transaction costs of monitoring successfully, gathering information on each other, and/or punishing in the case of default, or perhaps even just the presence of stronger altruistic motives towards each other. The strength of these connections might merely be a function of living closer to someone else, whereas social capital typically refers to the depth of a given relationship or the level of trust and/or information between individuals. This paper analyzes the extent to which social connections facilitate the monitoring and enforcement of loans in a group liability arrangement. Typically, showing that higher social connections cause higher loan repayments is a difficult task due to selection and group formation issues. Since most group lending programs rely on peers to screen each other and form groups, fundamental endogeneity problems exist when analyzing the impact of social connections on lending outcomes. For instance, if groups are formed around neighborhoods, and neighborhoods with stronger social networks also have more economic opportunities, then empirically one should observe a correlation between the social connections of a group and its likelihood to repay. Indeed prior studies have found correlations, but no causal link, between social connections and repayment.[3] With specific evidence on the causal link between social connections and credit markets, one could design better development credit policies. This paper's main contribution is its ability to solve this selection issue and show such a link: social connections, through better monitoring and enforcement, causes higher repayment and savings for participants in this group lending program.

Peer-selection also typically prevents the econometrician from distinguishing between *ex-ante* (selection) and *ex-post* (monitoring and enforcement) paths through which social connections cause better (or worse) lending outcomes. The analysis here successfully isolates the monitoring and enforcement path. It is important to note that this does not imply anything about the effectiveness (or ineffectiveness) of peer selection on group lending.

I collected data from FINCA-Peru, a group lending organization, to investigate whether geographic and cultural concentrations make peers more likely to both repay their loans and save more. FINCA-Peru's process for assigning individuals to groups creates a natural experiment with quasi-random group formation. This quasi-random process provides the strategy for identifying social connections. When lending groups are formed, the initial members neither select each other nor are neighborhood-based, as is common in other group-lending organizations. Instead, when individuals seeking a loan come to FINCA, they are put on a list. Once this list contains thirty names, a group is formed. Group meetings take place in the FINCA office in the city center and not in the various neighborhoods, allowing groups to contain members from all over the city. This unique assignment process creates groups with exogenous levels of initial social

[2]This definition is similar to Adler and Kwon's [2000] internal social capital.
[3]See Zeller [1998], Wydick [1999] and Ahlin and Townsend [2004].

connections. Since each group has fewer than 30 members, chance alone produces some groups with higher levels of social connections, i.e. they are more geographically or culturally dense. In addition, because individuals do not screen each other beforehand, improved enforcement and monitoring, and not selection, explains the impact of social connections on group outcomes.

I find that individuals who live closer to one another and are more culturally similar to others in the group are more likely to repay their loans and save more. There are many reasons to believe that this is a result of their ability to better monitor and enforce the loans. I present direct evidence of monitoring, such as knowledge of each other's default status, as well as direct evidence of punishment, such as deterioration of relationships. Monitoring and enforcement could lead to improved repayment rates directly by increasing the cost of defaulting, or more indirectly, by inspiring stronger group solidarity, more diligent work ethics and hence better business outcomes. I also find evidence that better connected individuals are more likely to be forgiven after defaulting, suggesting that their peers were able to distinguish between default due to moral hazard and default due to true negative personal shocks. These findings provide important insights into the factors that drive the success of group banking projects.

The measures I use (geographic proximity and cultural similarity), although more general than standard measures of social connections (or social capital), have three distinct advantages in this context. First, they are unlikely to be influenced by participation in the credit program, and hence are not endogenous with respect to outcomes of the lending group. Second, they can be measured accurately even on a recall basis. Third, they are easily observable, making it simpler to formulate and present policy recommendations. Furthermore, as shown in the Data Appendix, I find that the geographic and cultural concentration indices are correlated with several direct measures of social interaction, such as whether individuals have bought or sold from each other, know each others' homes, borrow directly from each other, and sit next to each other in group meetings. Furthermore, Karlan [2004] finds that both the cultural and the geographical concentration indices are correlated with cognitive social capital measures, as measured by behaviors in a trust game and a public goods game.[4]

However, the broadness of the geographic proximity measure allows for alternative explanations for the lower observed default rates, such as reduced transaction costs to conduct monitoring activities. For this reason, I refer to the measures as measures of *social connections* rather than *social capital* as is often done in the group lending literature. Regardless, whereas this broadness limits the ability to interpret the results as evidence of social capital *per se* influencing lending outcomes, the primary findings provide solid evidence that peer monitoring and enforcement effectively reduce default rates. This is an important finding for policymakers and microfinance institutions, as well as academics interested in testing contract theories. We have observed a plethora of lending schemes to the poor, some with more success than others. Little empirical work has shown why some designs seem to work better than others; the findings here provide valuable insight into these important questions regarding lending to the poor [see

[4]See Krishna and Shrader [1999] and Uphoff and Wijayaratna [2000] for a discussion of social capital measures.

Banerjee (2002)].

This paper proceeds as follows. Section 1 discusses joint liability mechanisms and FINCA-Peru, the source of the data for this research. Section 2 discusses the survey procedures and summarizes the data collected. Section 3 discusses the identification strategy employed in the analysis. Section 4 presents the central results on the lending and savings outcomes. Section 5 presents results on direct observations of monitoring and enforcement activities. Section 6 concludes.

1. Joint Liability Mechanisms

1.1 FINCA-Peru

FINCA-Peru uses a village banking lending methodology, first introduced by FINCA International in 1984 and now used by over 80 organizations in 32 countries. A village bank is a group of 30 women who meet weekly at the FINCA office both to borrow and to save, simultaneously. Most members have two loans, one from FINCA (the external loan) and one from their own pool of savings (the internal loan). Interest rates on both external and internal loans are 3 percent per month. In the case of default on either loan, the group's savings is used to pay back the loan. Each week the members make an installment payment on their external loan. In addition to the installment on their external loan, all members must make a savings deposit such that at the end of the four-month loan cycle they will have saved at least 20 percent of the amount borrowed under their external loan. Operationally, the loan installment and savings deposit are made together, as one payment. Clients also are encouraged to make additional voluntary savings deposits. The savings deposits (both mandatory and voluntary) do not lie idle. Each week, the savings are accumulated and lent out to some of the group members as one month internal loans. At the end of the loan cycle, interest earned on the internal loans is paid out to the members proportionally, by the amount of savings each has amassed.[5] FINCA earns the interest on the external loans. The savings and internal loan structure is very similar to a rotating savings and credit association since all members make small weekly deposits, and then each week a small fraction of the members receives large loans from the savings of everyone[6].

Empirically, FINCA has perfect repayment on its loan to the group. When there has been default, it always has been on the individual level and fully covered by the individual's own savings or by the other women's savings. Regardless, in weekly meetings FINCA employees emphasize to the clients the need to monitor and enforce each other's loans, even if they are fully collateralized to FINCA. FINCA does this for two reasons. First, although their rate of return is not directly affected by internal default, groups with higher internal default pose a higher risk of eventual default to FINCA. Second, groups with higher internal default have higher dropout rates, and the acquisition of new clients is costly for FINCA, particularly since new clients start out at lower loan

[5]Hence, the interest received by an individual is equal to her pro-rata share of the net interest earned by the group. The net interest is equal to the sum of all interest earned by the group on the internal loans minus the sum of all defaults.
[6]See Besley, Coate and Loury [1993] for a discussion of rotating savings and credit associations, or ROSCAs.

sizes than tenured clients.

When individuals want to join a bank, they typically arrive on their own or by invitation of a member of another group.[7] FINCA does not advertise in the community, nor explicitly ask individuals to seek out new applicants. Most people in the community, in fact, are aware of FINCA and know where the office is. Clients do not come in already formed groups.[8] If these individuals meet the basic criteria (have a business, understand the rules, and want a loan), their names are placed on a waiting list. When 30 names are on the list, a group is formed and individuals receive their first loan. This process happens quite quickly, typically in a week or two. FINCA claims to follow this methodology for two reasons, despite its potential drawbacks (i.e., it does not use the peers to help select the best clients). First, they believe that it is the fastest way to create new banks. Asking clients to go out and find others would inevitably take much longer. Thus, individuals do not feel compelled to seek out others in order to speed up the group formation process. Second, FINCA's mission includes building new social connections, hence they prefer initial group members not to know one another. FINCA hopes that through participation in its program it not only provides credit to the poor, but also helps the poor develop new relationships, both social and business, and in so doing strengthens the social fabric of the community as a whole.

Each week the clients are required to attend a meeting at the FINCA office located in the town center. Several activities occur at this weekly meeting, including loan payments, savings deposits, issuing new loans, training in group operations and the importance of group solidarity, and monitoring of loan repayment by all members. Attendance at meetings typically exceeds 90 percent, although poorly performing groups often experience lower attendance.[9] For some groups, monitoring activities are very regimented. After all payments are recorded, the group "board" (with supervision by a FINCA employee) reviews all the default situations. It then assigns specific individuals to visit the person in default and to inquire as to the cause of the default or late payment.[10] When members leave a group, either voluntarily or involuntarily, their place often is filled by a friend or relative of another group member through direct invitation.

FINCA's operating philosophy encourages clients to develop solidarity or social capital. While this is evident from the meeting hall posters propagating the values of camaraderie, trust and teamwork, it is even more evident in the training materials provided to the employees and clients. In these materials, FINCA emphasizes that the clients themselves are responsible for monitoring the group members in order to ensure that loan proceeds are used properly and for enforcing repayment and attendance.

1.2 Why Group Lending?

---

[7] This occurs when there is no current opening in one's own bank. So an individual may be referred by a client of FINCA, but placed into a group without the referring member.

[8] In the entire sample, I observed only three instances of individuals coming in groups of three, and no groups larger than that.

[9] Clearly, this could be causally related in either direction.

[10] Some groups do this more carefully than others. Anecdotally, groups with higher overall repayment rates were more likely to follow through with such proactive monitoring activities.

Poor individuals lack formal credit because lenders have little means to screen clients, monitor the use of funds, or enforce repayment. In recent years many development organizations have used group lending to deliver credit to poor individuals. Group lending purports to pass off the screening, monitoring and enforcement of the loans to the peers [see Banerjee, Besley and Guinnane (1994), Diamond (1984), Ghatak and Guinnane (1999), Stiglitz (1990), and Varian (1990)]. Furthermore, group loans help formal lenders overcome the prohibitively high fixed cost of delivering small loans.

Monitoring and enforcement are distinct, although difficult to distinguish empirically. Monitoring itself does not guarantee repayment, but it allows a lending organization to know whom to punish for not repaying. Although a commercial bank can attempt to monitor business and life outcomes for individuals, it is both difficult and costly to do so. Group lending mechanisms provide incentives to the borrowers to monitor each other to see who can pay and who can not. Monitoring can take on several forms, such as observing repayment of the loan, visiting another's business to verify that it is in operation, showing receipts to demonstrate that inventory was purchased with the loan proceeds,[11] and talking to others in the community to confirm negative shocks like illness. In these examples, the extent of someone's social networks is critical and positively related to the ability to monitor or be monitored.[12] Armendariz de Aghion and Gollier [2000] and Armendariz de Aghion [1999] show theoretically how peer monitoring alone, with random formation of groups, can help overcome adverse selection problems when monitoring is costly for the lending institution itself. Stronger social networks have lower monitoring costs, which results in more credit being extended.

To enforce lending contracts, lending institutions typically resort to legal options, such as seizing property of the borrower or garnishing wages directly from the employer. In most poor communities, such punishments fail for one of two reasons, either the legal infrastructure does not support such action, or the borrower has no seizable assets or wages. De Soto [2000] and Besley and Coate [1995] discuss these issues at length. Group lending overcomes these failures by taking advantage of people's desire to protect their social connections (and social capital) and avoid any possible repercussions. Such repercussions could be economic and result in reduced trading partners for one's business, social and lead to loss of friends, or psychological and damage one's self-esteem.

Group lending does not unambiguously facilitate repayment through monitoring and enforcement. Three issues in particular could cause group lending to generate higher default than individual lending, and cause groups with higher social connections to have higher default than groups with lower social connections. First, if social connections are strong enough to permit the monitors to distinguish between personal negative shocks and mere reneging, then punishment could be made contingent upon the observations of the monitor. This effectively would be an insurance as well as a lending mechanism and would weaken the incentive to repay after personal negative shocks. Second, Besley and Coate [1995] present a strategic default model: as good individuals observe others

[11] The fungibility of money potentially makes this particular monitoring action no better than observing that they are working.

[12] In the extreme, family members have been shown consistently to overcome information asymmetry problems, for example, in the used car market. See Pollack (1985).

defaulting, they themselves default as well since they will not receive a new loan even if they repay and they will suffer no scorn from others for defaulting. If borrowing individually, these individuals might have repaid. In both of these theories, higher social connections should generate higher default.[13] Third, the presence of the insurance and possible risk-sharing arrangements could encourage ex-ante moral hazard, or shifts into riskier project choice by the clients. Whereas this may be optimal for the clients, this does pose a greater risk to the lending organization (which may be compensated in that higher interest could be charged). Hence, the theoretical relationship between social connections and repayment is ambiguous.

The existing empirical research on the relationship between social connections and repayments is also inconclusive, partly due to the endogeneity problems discussed earlier. For instance, Sharma and Zeller [1997] using credit groups in Bangladesh, and Ahlin and Townsend [2004] using data from Thailand, find that groups with high levels of family relations have higher default. These findings could be because family members are unable to screen effectively. Ahlin and Townsend [2004] and Wydick [1999] find that groups that report threats of social sanctions for failure to repay have higher repayment; however, why some groups decide to have such policies is not understood, and potentially endogenous (or potentially creating omitted variable problems for drawing causal inferences). Also, such reports do not indicate whether higher levels improve or worsen the ability of social connections to cause better outcomes. Sadoulet and Carpenter [1999] analyzes the structure of a Guatemalan peer mechanism and finds that by design it lends itself to risk-sharing as well as enforcement of repayment. Most recently, La Ferrara [2003] studies kin groups in Ghana and finds that punishment is exacted not only on those who default, but also on the kin of those who default, and that the threat of such punishment induces compliance in the short run. These studies demonstrate that the relationship between social connections and group lending outcomes is complicated and worthy of further study. This paper builds on that research by using a natural experiment to show that having stronger social connections *causes* higher repayment and savings by facilitating monitoring and enforcement of group lending contracts.

2. Data

This research uses data from participants in the Ayacucho[14] program in Peru from 1998 to 2000. For this study, I divide participants into two groups, those that were invited by a member of their own group and those that were not. The analysis is conducted on the latter, i.e. the uninvited. The primary analysis will regress loan default, savings and

[13]A third concern involves the formation of small groups within the larger group and then collusion among the members of the smaller group. Suppose a bank has many small, well-connected groups. Suppose a small group decides to collude whereby one member does not repay while the others report that indeed she has no capacity to repay due to some calamity. In an individual setting with imperfect monitoring, this individual might repay. However, in this setting, the promise of false monitoring by her immediate peers in fact guarantees that she is not monitored. The entire small group could not go into default because then there would be no "good" client to report back to the group. Naturally, if the entire bank divides into mini-groups with each mini-group using this strategy, this could lead to the unraveling of the group as a whole. I found no anecdotal evidence to support this possibility at FINCA-Peru. See Genicot and Ray (2003) for a theoretical discussion of such dynamics.

[14]Ayacucho is a town in the Andes with a population of 150,000. The Shining Path, the communist-oriented faction from the 1980's civil war, was started in Ayacucho.

attrition on geographic and cultural dispersion.[15] The default, savings, and attrition data come from FINCA-Peru's internal records. These records also contain certain basic demographic information, such as marital status, number of children, and age. For this project, I employed a team of 10 surveyors from January through June 2000 to collect data on cultural identity, social connections amongst group members, method of their arrival to FINCA (i.e., invited or uninvited), location of their home, and other demographic information not already collected by FINCA. Three types of surveys were conducted in this phase: group interviews to collect publicly known information (such as who invited whom), individual surveys conducted privately, and individual surveys conducted in the homes or businesses of former members. See the Data Appendix for a description of the data collection process.[16] Further data about monitoring and enforcement activities were collected in 2001 and are discussed in Section 5.

The primary dataset for this project contains 2,054 individuals over 6,874 loan cycles, or an average of 3.3 loans per individual. For the primary analysis, the dependent variables are the outcome for each uninvited individual's first loan, and the key independent variables are that person’s connection to the *original* members of her group. I have data on the selection method (i.e., uninvited or invited) for 1,719 of the 2,054 individuals. Twenty percent of the uninvited individuals had some default on their first loan, whereas only 16.0 percent of the invited individuals had some default on their first loan. The average savings deposits made during the 4-month loan cycle was $59 for both the uninvited and invited. Tables 1 and 2 show the summary statistics for individuals, and Table 3 shows the summary statistics for groups. The summary statistics are shown separately for the invited versus the uninvited individuals since this is a crucial distinction for the econometric identification of social connections.

To measure social connections, I examine the cultural and geographic proximity of each individual to the original members of the group. Research at both macroeconomic and microeconomic levels suggests that cultural heterogeneity influences the societal norms that dictate how economies and political bodies organize themselves. For instance, Alesina, Baqir and Hoxby [2004] find evidence for explicit tradeoffs between racial and income heterogeneity and economies of scale in the formation of local jurisdictions. Alesina and La Ferrara [2000] find that cultural heterogeneity negatively influences participation in community and civic activities. Glaeser et al [2000] discuss the determinants of trust in the United States, with strong findings for cultural heterogeneity negatively influencing trust.

Most people in Ayacucho, Peru are a blend of indigenous and Western heritage. Individuals of either extreme can be identified easily by their language, dress, and hair style. For instance, indigenous individuals wear black hats with large rims, keep their hair in braids, and speak only Quechua, whereas Western individuals have short, styled hair, speak only Spanish, and wear jeans and other Western clothing. Using the above characteristics, I create a culture score from zero to eight for each individual. I then

[15] The Data Appendix discusses the formulation of these measures and provides evidence supporting the relevance of these as social capital measures.
[16] See http://www.karlan.net for copies of the survey instruments.

calculate the probability that a given individual has the same culture score as a randomly chosen individual from the original group. This is analogous to a standard cultural fragmentation index [see Alesina and La Ferrara (2000)], which calculates the probability that two individuals randomly drawn from a group are of the same cultural background.

Geographic distance between members captures social connections for many reasons. Monitoring costs are reduced when individuals live closer to each other. Individuals with more common acquaintances or friends will procure information more easily about each other. Also, the threat of reputation loss is potentially more effective among those who live closer to each other since such individuals will have more frequent future interactions and more acquaintances in common. In order to quantify geographic concentration, I employ two measures: the average distance of an individual's home to those of the original members, and the percentage of original members who live within a 10-minute walk of the individual. The first is similar to a metric used by Busch and Reinhardt [1999] to calculate geographic concentration of industries. The second measure recognizes that it is costly, perhaps exceedingly so, for everyone to monitor everyone else. Therefore, it is more sensible for individuals to be responsible for monitoring those who live close to them. For reasons discussed in the next section, both measures relate distance to the original, not current, members of the group. For group-level analysis for both cultural similarity and geographic concentration, I use the average of the individual measures.[17]

3. Identification Strategy

The identification strategy exploits the institutional fact that FINCA-Ayacucho forms initial groups with little self-selection. This solves an endogeneity problem fundamental to group lending, that peers select their own group members [see Ghatak (1999) and Ghatak (2000)]. Peer selection and group formation in this context create two empirical issues: the first issue is about establishing a causal link from social connections and group outcomes and the second issue is about distinguishing between selection and ex-post monitoring and enforcement stories. Peer selection might generate omitted variable problems (e.g., individuals assortatively match into groups on characteristics unobservable to the econometrician, yet correlated with both social connections and business success) or simultaneity problems (successful groups help create better social connections). Such omitted variable and simultaneity problems make it difficult to argue that observed correlations between social connections and repayment (or other group outcomes) are causal in nature, rather than spuriously correlative. Second, it prevents the econometrician from identifying the impact of social connections on effective monitoring and enforcement of loans, as distinct from the effective selection of trustworthy individuals.

As discussed in Section 1, when individuals want to receive a loan from FINCA, they typically arrive on their own or by invitation of a member in a group without an

[17]Again, due to the endogeneity of the social connections for invited members, the group-level measure is best calculated by averaging the uninvited person's connection to the original members of the group, rather than by measuring the connections of current group members to each other.

opening. Their name is then put on a list and once 30 names have been collected, a new group is formed. As individuals leave the group, openings are typically filled by invitation of a member of that group. Out of the 1,078 individuals who came by invitation of a member of the same group, only eight reported coming by invitation of two others, and only one reported coming by invitation of three others. Hence, even when individuals come by invitation, few cases exist of even a small portion of the group forming prior to arrival to FINCA.

I divide participants into two groups, invited and uninvited. I claim that the social connection between the current, uninvited members and the original, uninvited members is exogenous (whereas that of the invited is endogenous). I examine this key assumption below. Since the uninvited members can invite members, I want to measure the social connections between each uninvited member and the original, not current, members of the group. This solves another problem as well, that the dropout process may homogenize groups at different rates depending on the prior success of the group. Furthermore, by only analyzing the uninvited members, I can eliminate peer selection as a possible explanation of the findings. This issue has been difficult to overcome in prior studies, such as Sharma and Zeller (1997).

This analysis then takes advantage of small sample variation. Since each group has on average 15 uninvited individuals, the idiosyncratic variation proves sufficient to conduct an analysis of the impact of social connections on financial outcomes[18]. Table 3, for instance, shows the means and standard deviations of the group-level measures of social connections.

The basic model I estimate is of the form:

$$Y_i = \alpha + \beta_1 X_i + \beta_2 Z_i + \varepsilon_i, \tag{1}$$

where Y is a financial outcome (either default, savings or dropout), X is one of the social connections measures (either geographic proximity or cultural similarity), and Z is a matrix of neighborhood and cultural dummies and other demographic information.

Using invited individuals poses at least two endogeneity problems to the above specification. First, there is an unobservable selection problem. For example, more sophisticated individuals might be more likely to have successful businesses and repay their loan, as well as more likely to invite their peers into the group. Hence, since individuals tend to invite those who live closer to them, geographic proximity would be correlated with repayment, but not because of improved monitoring or enforcement of the loans. Second, a simultaneity problem exists. Most group lending financial institutions claim to provide two key benefits: higher or smoother consumption by resolving credit market failures and greater social cohesion or empowerment. If this second benefit is true, the correlation between social connections and group outcomes easily could be

[18]To verify that the observed variation was consistent with a random process, I conducted a monte carlo simulation in which 500 sets of 42 similarly sized groups were formed randomly from the entire sample. I then verified that the actual mean and standard deviation for both the geographic and cultural measures fell within the middle 95% of the distribution of each statistic in the monte carlo simulation. Furthermore, the size of the group is not correlated with the measures of geographic and cultural concentration, so the measure does not appear to be a construct of, for example, endogenous group size or missing data.

causal from the other direction.

I use two tests to confirm that the group members identified as uninvited are placed randomly into groups. These tests show that there was no assignment to groups on observables, but cannot prove this absolutely, as assignment could have been on unobservables. However, interviews with FINCA and the participants support the claim that the uninvited truly were uninvited, and assignment to groups can be considered random.

First, I use a test similar to Ellison and Glaeser [1997] to determine whether the observed geographic dispersion is different from what one would expect to arise randomly,[19] as if location were chosen using a dartboard. Ellison and Glaeser uses the following measure of geographic dispersion:[20]

$$GD_{group} = \sum_{neighborhoods} (s_i - x_i)^2 \quad (2)$$

where s_i is the share of the group from neighborhood i and x_i is the share of the general population from neighborhood i.[21] E(GD), given random selection, is given by:

$$E(GD_{group}) = [1 - \sum_{neighborhoods} (x_i)^2]/n \quad (3)$$

where n is the number of members in each group. The results of this test support the claim that uninvited, but not invited, individuals select into groups randomly. Table 3 shows these results. The mean of GD is not significantly different than the E(GD) for uninvited (0.147 versus .127), but is significantly more for invited (0.252 versus 0.203, significant at 95 percent). This supports the claim that uninvited individuals are grouped together in a random process with respect to geographic concentration, and this also supports the omission of invited individuals from the analysis since they do not pass this test. I conduct a parallel test for the cultural dispersion of each group. For both uninvited and invited, the difference between actual and expected cultural concentration is insignificant (0.119 versus 0.106 for uninvited, and 0.184 versus 0.167 for invited).[22]

This measure of geographical concentration incorporates the dispersion across neighborhoods, but does not take into account distance between neighborhoods. To capture distance between individuals, I test whether the percentage of individuals in one's group who live within a 10-minute walking circle is greater than the percentage of

[19] This is a test of the exogeneity of the social capital measures, whereas the monte carlo simulation referred to in the above footnote verifies that the small sample of each group was not sufficiently large as to remove any variation across groups. Such variation is necessary to identify generate an interesting enough range of observed values for the exogenous variable.

[20]Since this measure does not incorporate distance between neighborhoods, I do not use it for the primary analysis.

[21]Without data on population by neighborhood, I use the total sample of all banks to generate general population estimates. The area was then broken into a grid with 43 different neighborhoods.

[22]Similar to geographic dispersion, the measure of cultural dispersion is

$$CD_{group} = \sum_{i = 0 \, to \, 8} (s_i - x_i)^2$$

where s_i is the share of the bank with culture score i, and x_i is the share of the general population with culture score i. Similarly, E(CD), given random selection, is given by

$$E(CD_{group}) = [1 - \sum_{i = 0 \, to \, 8} (x_i)^2]/n.$$

individuals in the entire sample who live within this same 10-minute circle. Table 2 shows this comparison: 22.4 percent of fellow group members live within a 10-minute circle of each uninvited member, whereas 21.1 percent of the total sample live within these same 10-minute circles. These are statistically the same. However, for the invited members, the difference is 20.5 percent versus 17.2 percent, significant at 99 percent. This suggests both that the uninvited are randomly located and that this is a powerful test.[23] It found that invited individuals are not randomly located, and in fact are more clustered geographically. I conclude that the allocation of uninvited individuals into groups appears random, allowing the idiosyncratic variation to identify the social connections within the groups.

4. Empirical Results on Lending and Savings Outcomes:

4.1 Default Rate

The default rate is perhaps the single most focused on outcome for both researchers and practitioners in analyzing the effectiveness of a particular mechanism design. To the extent that default, or specifically the risk of default, leads to credit market failures, default is harmful to social welfare. However, over-monitoring or over-punishing might yield higher repayment rates but not maximize social welfare. Regardless, microfinance institutions focus intensely on repayment rates as one of the key, if not only, metrics of financial health and sustainability.[24]

Social connections could facilitate monitoring and enforcement through reduced cost, increased accuracy of information or higher reputation values. In Ayacucho, monitoring means visiting clients or neighbors of delinquent clients to verify their stories. If someone says they have not repaid due to illness or death in the family, a simple house check or conversation with their neighbors typically can confirm this. Hence, group members who are physically close should be better at monitoring one another. Furthermore, with more mutual acquaintances, the information garnered through the monitoring is likely to be more accurate. This may also cause the threat of enforcement to be more effective since reputation matters more among one's peers. Stories of repossession of assets are rare; most enforcement activities involve moral suasion via frequent visits to the person's home or place of business. Cultural homogeneity captures the expected level of these social connections between individuals, as well as the likely extent of mutual acquaintances.[25]

The dependent variable, default as a percentage of potential loan amount,[26] is truncated at zero since most individuals fully repay. Furthermore, almost all individuals pay part of their loan. Default typically begins somewhere in the middle of the loan term, at which point the client stops attending meetings and making her weekly loan payments.

[23]Note that uninvited individuals are more likely than invited individuals to live near those in other groups. This indicates that uninvited individuals are more likely than invited individuals to come from the center of town, but does not imply any bias in the group formation process itself.
[24]See Morduch (1999a, b).
[25]It does not, however, capture direct travel costs.
[26]Potential loan amount is equal to the client's last loan amount plus their accumulated savings.

The estimating model uses a tobit specification as follows:

$$Default_i = X\beta + Z\gamma + \varepsilon_i \tag{4}$$

$$Default_i^* = \{0 \; if \; default_i \leq 0; \;\; default_i \; if \; default_i > 0\} \tag{5}$$

$Default_i$ is a latent variable for person *i*'s default, X is either geographic concentration or cultural similarity, Z is a matrix of control variables, including neighborhood dummies, year and tenure of group, and education.[27] I include neighborhood dummies in order to account for a potential correlation between density of a neighborhood and business profitability. For similar reasons, I control for distance to the town center, where the main market and FINCA office are located. Each measure of social connection is included in a separate specification. For bank-level specifications, the default is calculated as the average default for individuals in that group. Similarly, most control variables are calculated as the average of the group. When examining the impact of geographic dispersion, I control for average distance to town center and the percentage of the group that lives within 5 minutes of the town center. This accounts for the possibility that higher-concentrated groups are closer to the town center where the most economic activity takes place. When examining the impact of cultural similarity, controls for the percentage of each group that are indigenous and Western are also included. Table 4 shows the results for the specifications with the individual as the unit of observation. Table 7 column 1 shows the results for the specifications with the group as the unit of observation. The Data Appendix Table 2 shows the typical relationships between the control variables and outcomes of interest.

Of the 616 uninvited individuals in the sample, 125 had defaulted at the end of their first loan. Of the 245 group observations, 44 had individuals with default at some point in the sample. The default only occurred on the internal loans made from the members' savings. FINCA had perfect repayment on its loans to the groups.

For individual-level analysis, I use the initial loan cycle for each client and not the entire history since both an attrition bias and an attenuation bias exist if the entire history is used. When expanding the analysis to each client's full history with the project, I weight each individual equally. However, many of those who dropped out without default perhaps were close to default and left because they feared repercussions from failure to repay the next loan or found the pressure exerted from the first loan too unpleasant. This attrition should understate the predictive power of the social connection measures since these are the individuals for whom social connections potentially matter more. Conducting the analysis on the initial loan cycle avoids this bias. Furthermore, since the independent variable is a measure of distance (either geographic or cultural) to the original members of the group, attenuation bias suggests that as the group ages, this becomes a noisier measure of the enforcement and monitoring capabilities of the group.

[27]Control variables also include distance to FINCA (town center), Ayacucho versus Huanta dummy, age, age-squared, marital status, siblings, children, and number of persons in household.

To allow for clearer interpretation, each measure of social connection is included in its own specification and is presented separately as a cell in Table 4.[28] Columns 1-3 show the OLS, tobit and probit results, respectively. Both cultural similarity and geographic concentration negatively predict default (significance ranges from 99 percent to marginally insignificant). The second geographic concentration measure, which captures the number of individuals within a 10-minute walk, is significant statistically and economically. The first measure, average distance to the original members, is signed intuitively but not significant statistically. This is likely due to the irrelevance of the distance of the further individuals for effective monitoring and enforcement. The economic magnitude of these findings is significant: a shift from the 25th percentile (6 percent) to the 75th percentile (32 percent) of the second geographic concentration measure suggests a 7.2 percentage point decrease in the probability of default. Similarly, a shift in the cultural dispersion measure from the 25th percentile (8 percent) to the 75th percentile (28 percent) decreases the probability of default by 3.9 percentage points. Comparing column 4 to column 1, column 5 to column 2, and column 6 to 3 shows how the attrition and attenuation bias leads to underestimating the impact of social connection, e.g. in the tobit model the coefficient on cultural similarity falls from -4.23 to -1.46 and the coefficient on the second measure of geographic concentration falls from -6.08 to -3.75.

The group level specifications in Table 7 show that both measures of geographic concentration predict default, significant to 95 percent for the average distance of all members and 99 percent for the percentage that live within a 10-minute walk. The cultural concentration, although signed intuitively, is not significant statistically.

4.2 Savings

All individuals are required to make weekly savings payments such that over one loan term, the individual has saved 20 percent of their loan from FINCA (e.g., on a $50 loan, at the end of sixteen weeks the client has $10 in savings). In addition, many clients make voluntary savings payments as well. This savings does not lie idle, but rather serve as another source of borrowing for these same members. Thus, for each dollar in savings a member typically has access to two dollars of loans: one dollar from FINCA and one dollar from the savings pool. The return on this savings is the same for the entire group, and is calculated as the profits on loans made minus default, divided by total group savings at the end of the loan cycle. Social connection influences each input into this formula. First, as found above, higher social connection leads to lower default, and since defaults are covered by the group's savings, lower default directly implies a higher return on savings. Second, not all groups lend out all of their savings. Many groups invest their savings if they do not have safe projects. Again, since higher social connections lead to lower default, groups with higher social connections should lend out a higher percentage of their savings. Any savings not lent out remains with the FINCA cashier and does not earn interest.

Table 5 shows the results for the specifications with the individual as the unit of

[28]Results remain similar when all three measures included in each specification.

observation, and Table 7 shows the results with the group as the unit of observation. Again, to allow for clearer interpretation, each measure of social connection is included in its own OLS specification. Geographic concentration, but not cultural similarity, produces higher savings. Table 5, columns 1 through 3 show the results using three different savings variables: total savings, mandatory savings and voluntary savings. All specifications include the same controls as were included in the default analysis. The results for total savings show that individuals who live further from others in the group save less, significant at 90 percent in the individual-level (Table 5, column 1) and insignificant at the group-level (Table 7, column 2). A shift from the 25th percentile to the 75th percentile in the average distance to others in the group implies an increase of $13.20 in savings per client in their first 4-month loan cycle, which is significant given that the mean savings is $58.69. As with default, when the analysis uses the entire tenure of each client, the attrition biases the results downward (see Table 5, column 4 versus column 1).

Since mandatory savings are paid in the same installment along with weekly loan payment, predictors of loan repayment also predict mandatory savings deposits. As such, the percentage of the group which lives within 5 minutes is a stronger predictor of individual default and is also the stronger predictor of mandatory savings. Furthermore, since voluntary savings should be driven by return on savings, measures that predict group-wide return on savings should also predict voluntary savings (see Table 7, columns 4 and 5), significant at 95 percent.

Following this logic, Table 7, column 5 shows that as the group is more concentrated, the return on savings rises (significant at 95 percent). The coefficient of 0.04 suggests that a shift from the 25th percentile to the 75th percentile in geographic concentration would increase the return on savings by 1.31 percentage points per annum. On $100 in savings, such a change in group composition could produce additional interest earnings approximately equal to the daily wage of a poor entrepreneur.

Cultural similarity, although influential on default, does not significantly influence the level of savings.[29] One possible explanation is that cultural similarity inspires empathy within cultural groups, but where empathy is asymmetric in gains versus losses. In other words, empathy inspires repayment on loans because failure to do so would harm peers; however, empathy does not inspire higher savings since that has a positive and second-order benefit to the peers. Indeed, no statistically significant relationship is observed between cultural similarity and voluntary savings.

4.3 Attrition

Since financial outcomes are highly accurate predictors of retention, an attrition bias must be considered when examining the predictors of default and savings. Those who remain in the project for many years are different in many respects than those who leave. For FINCA, length of participation in a group varies widely, with attrition likelihood initially

[29] The results are insignificant but negative, with higher cultural similarity predicting lower savings. When geographic concentration is omitted from the specification, the coefficient for cultural similarity falls to zero when predicting total savings.

high and then falling over time. Attrition falls from 24 percent after the first loan to 16 percent after one year and 11 percent after two years. Default is the strongest predictor of attrition: 71 percent of those with default left while only 13 percent of those without default left.[30] There is neither a firm rule nor a precise process for deciding whether a group member who defaults should remain part of the group. While FINCA influences this decision to some degree, the ultimate judgment lies with the group as a whole. Table 6, column 1 shows a probit model of the dropout decision. Default is highly correlated with attrition, significant at 99 percent and those with higher savings are less likely to leave (insignificant statistically). The probit model is specified as follows:

$$Y_i = \alpha + \beta_1 X_i + \beta_2 D_i + \beta_3 (D_i * X_i) + \beta_4 Z_i + \varepsilon_i, \tag{6}$$

where Y_i =1 if an individual drops out and Y_i =0 if an individual remains in the group, X_i is one of the social connection measures, D_i is default, D_i *X_i is the interaction of default and the social connection measure, and Z_i is a vector of control variables.

I test two hypotheses. First, I examine whether social connections influence the decision to leave the program. I do not observe clearly whether an individual leaves by force or voluntarily (the reality is often murky, so this is not just a data issue). Such an effect can be due to lower utility from attending meetings when there are fewer sociable peers at the meeting. Or it could be the fact that those with higher levels of social connections have more to lose in the case of default, and hence might be quicker to leave. Empirically, all three measures of social connections indicate that higher levels of social connections lead to lower dropout rates. None of these results is significant statistically.

To be able to distinguish between idiosyncratic negative shocks and merely reneging, one needs especially good monitoring. Individuals who are particularly close to each other potentially can arrange such a risk-sharing arrangement.[31] Although no anecdotal evidence exists to suggest that such arrangements are made explicitly ex-ante, both qualitative and quantitative evidence suggest that they take place ex-post. FINCA reports instances where individuals vouch for delinquent members in order to prevent them from being forced out of the group. I test for this empirically using a probit model that finds that those with higher levels of social connections are more likely to remain in the group after default than are those with lower levels of social connections. Table 6, columns 4, 5 and 6 show that the interaction of social connections and default is significant at 99 percent and negative. This suggests that individuals with higher levels of social connection are not being punished after default as much as those with lower levels of social connection. This is exactly consistent with the Rai and Sjöström [2001] model, in which individuals have information that the lender (FINCA) does not. As Rai and Sjöström [2001] discusses, the lending institution (FINCA) provides the framework to facilitate a risk-sharing arrangement, hence uses the information that peers are able to gather (but FINCA is not) regarding each other's ability to repay loans. Another possible story is that those with higher social connections receive alternate, but less severe,

[30]Although, most who leave do so without default. See Panel A of Table 8.
[31]See Rai and Sjöström [2001] for a theoretical discussion of how cross-reporting can efficiently induce repayment.

punishments, or are perhaps unpunishable. These quantitative data cannot distinguish between lower dropout due to the higher cost of punishing someone you know or due to successful identification of individuals with true negative shocks.

Qualitative data support the monitoring explanation for this empirical finding. In a second survey (discussed in more detail below in Section 5), I asked current members about other members still in the program who had default. In each instance, I asked an open-ended question as to why that person was allowed to stay. I recorded the free-form response of each member, and then categorized the answers. Appendix Table 3 shows these results. In 38 out of 44 instances of individuals having default but remaining, at least one of the other current members reported a negative shock that the individual experienced or reported evidence that the person did undergo some process of explanation to her peers in order to be allowed to remain. Conversations with FINCA Peru management also support the story that individuals with negative and observable shocks are forgiven if the shock is verified by someone else in the group. Hence, these results support the hypothesis that group liability, through the social connections of the members, provides incentives to members to monitor each other's ability to repay loans.

5. Monitoring and Enforcement Activities

5.1 Data Collected

In 2002, after the initial data collection reported above had been completed, I collected further data from FINCA-Peru in Ayacucho on the monitoring and enforcement activities of clients. In a private interview, I asked each current client from 28 lending groups about each of the individuals that left their group or defaulted on their loan in the prior two loan cycles (the prior 5-8 months). Specifically, the questions included the following: a) Do you remember the individual?; b) Did the person leave with default?; c) If so, why did the person go into default, and how did you acquire this information; d) Did you know the person before joining the group?; e) Is the person an extended family member?; f) Is person a close friend?; g) Does the person live near your home; and h) Have you ever visited the business of the other person?

In addition, I asked them to compare their current relationship now that the person had left the program to their relationship with that person when they were in the program. Specifically, I asked: a) Do you buy or sell goods from the person more, the same, or less frequently?; b) Has your friendship become stronger, the same or weaker?; and c) Has your trust in this person become stronger, the same, or weaker? Table 8 shows the summary tabulations on these questions. The primary goal in doing this was to observe whether default is correlated with the destruction of social relationships. If it is, this supports the idea that social relationships help to enforce group lending contracts through the threat of reputation loss (or other informal punishment paths).

The empirical analysis on these data contains two tests. First, I examine monitoring activities by looking at how much accurate information current members have about those who recently left or had default. Second, I examine punishment by looking at whether relationships deteriorated after default.

5.2 Monitoring: Knowledge and Awareness of Causes of Default

In a developing country setting, lenders often cannot observe outcomes of borrowers, and hence are not able to assess ability to repay loans. Peer lending transfers responsibility of this task onto the peers; they should have access to better information (presumably through stronger social networks) and thus will be able to assess who can and cannot repay [Banerjee, Besley and Guinnane (1994)]. If peers are monitoring each other, then they should have accurate information about each other's outcomes and reasons for default. I observe this directly, by first asking individuals why their peers left the program and/or went into default.[32] I then examine the accuracy of these answers and whether it improves with stronger social connections. Table 9 presents these results. Next, I ask each respondent whether the individual who left had default when they left and then create a dependent variable for the accuracy of their information. This variable equals to one if the respondent's information was correct. Columns 1 and 2 in Table 9 report these results.

Using a probit specification, I then examine whether those with stronger social connections are more likely to get this question right. Those of similar cultural background are more likely to correctly answer this question (significant at 90 percent or 95 percent, depending on whether cultural similarity is scored as a binary variable if similar, or as the absolute difference between the two culture scores). Living closer to each other is not a predictor of having accurate information once I control for knowing the person before hand. Having visited the business of the borrower and knowing the individual beforehand are strong predictors of correctly knowing whether the individual had default. This supports a monitoring theory as a key mechanism through which group lending works.

Lastly, I ask respondents why an individual did not repay their loan. Restricting the sample to those who did have some default, I find that cultural similarity does correlate with being more likely to know the cause of default (see Table 9, Columns 3 and 4). Moreover, knowing someone before joining is also a positive and stronger (in magnitude) predictor of having this information. Geographic proximity, on the other hand, does not correlate with such knowledge. Again, simple monitoring activities, such as having visited the business of the borrower and knowing the individual beforehand, are strong predictors of knowing why someone had default.

5.4 Punishment: Changes in Relationships with Those Who Dropout of the Program

Peer pressure to repay the loan is often cited as a benefit of group lending [Besley and Coate (1995), Ghatak and Guinnane (1999)]. If this is in fact a mechanism through which group lending generates repayment, then after default one should observe some

[32] Appendix Table 3 shows the qualitative responses received to this question.

destruction of social relationships. This is testable using the survey we conducted of current members regarding those who dropped out of the program. For each individual who left, I asked the current member what has happened to their relationship: Did it remain the same, improve, or worsen?

For each question (business transactions, trust and friendship), there are three possible outcomes: worsen (-1), stay the same (0) or improve (1). Tabulations of these responses are shown in Panel B of Table 8. Whereas 2.8 percent of relationships deteriorated when there was no default, 12.0 percent deteriorated when there was default. Regarding the reported trust, the difference is 1.0 percent versus 5.9 percent, and regarding buying and selling goods from each other, the difference is 0.7 percent versus 2.1 percent.[33] In addition, when the dropout has default, the respondent is far more likely to report having spoken to that person outside of the bank meeting (e.g., at their business or home), 18.5 percent versus 6.3 percent for those without default. This is direct evidence of the monitoring and enforcement activity, since it is those with default that are visited by current members in order to observe their ability to repay and convince them to repay.

Regarding improvements in the relationships, those with default are far less likely to experience improvements in their relationships. In fact, no instances of improvement in friendship or trust exist after an individual leaves in default, whereas 0.3 percent of individuals report an improvement in trust or friendship with an individual after they left without default. Regardless, the small frequency of reported improvements suggests that any short term gain in social networks among current members tends to diminish as individuals leave the program.

6. Conclusion

In response to abysmal repayment rates and unsustainable projects [See Adams et al (1984), Kahlily and Meyer (1993), and Yaron (1994)], the past few decades have seen dramatic changes in the design of credit projects. Four mechanism design changes stand out: (1) the use of group liability to reduce screening, monitoring and delivery costs, (2) the promise of repeat lending as a repayment incentive, (3) the use of regular and more frequent payments, and (4) the offer, or sometimes requirement, of savings. Despite these significant changes, there has been little empirical research conducted to help organizations understand the effect of these innovations [see Banerjee (2002)]. In particular, the decision of whether to impose joint liability on borrowers is a central choice that many organizations face, yet few have studied empirically. This research finds evidence to support one hypothesis behind group lending: that monitoring and enforcement activities do improve group lending outcomes, and that social connections, broadly defined, facilitate the monitoring and enforcement of joint liability loan contracts. Social connections might have this effect simply through lowering the cost of gathering information about each other (i.e., a monitoring story), or through a social

[33] An alternative explanation for the deterioration of buying/selling is that individuals who leave after default are more likely to close their business.

capital story in which more connected individuals trust each other more and value each other's relationships more. Note that this social capital story encompasses both actions taken to protect one’s relationships, and also actions taken merely out of altruism towards those similar to you.

I find that both cultural similarity and geographic concentration lead to improved group lending outcomes (specifically, higher repayment rates, savings rates, and returns on savings). There is also suggestive evidence that social connections help groups distinguish between true negative shocks and mere reneging, and that those who have negative shocks are forgiven and thus allowed to continue borrowing. Furthermore, I find direct evidence of effective monitoring, such as knowledge and awareness of each other's default status and causes, as well as direct evidence of punishment, such as deterioration of relationships. The monitoring activities specifically occur through the same cultural channels found to predict repayment and savings. This further establishes the causal link between cultural similarity and repayment rates and savings.

These findings show that peer lending programs can be more effective if groups are more concentrated geographically and similar culturally. However, the conclusion does not support creating entirely homogenous groups, either geographically or culturally, since extreme situations are not observed in these data. Complete homogeneity might result in collusive activities or may make punishment more difficult. Furthermore, the findings should not be construed as an endorsement of group lending over individual lending, since the sample consists entirely of group borrowers, and those who opt for group lending may be influenced differently by peer pressure.

Although this paper examines the link between informal social connections and repayment of loans, it speaks to a larger issue of how nonmarket institutions and forces can help overcome market failures. These findings support a growing literature on the importance of informal networks for development. Further research to understand how these networks can best be harnessed, or better yet developed, is critical.

7. Data Appendix

7.1 Survey Data Collection Process

The primary survey data were collected from January to May, 2000 by a team of 10 local surveyors, and in 2002 by a team of 2 surveyors. Three surveys were completed in 2000: an individual survey conducted publicly at the weekly meeting, a private individual survey, and a former member survey, and one private individual survey was conducted in 2002.

The public individual survey included questions for which the answers were public information, such as how many homes of the others someone knows, how someone joined the group, and from how many others each person has bought or sold a product or service. These questions were done publicly for three reasons. First, individuals are more likely to speak truthfully for fear of others seeing them be untruthful. Second, other individuals were able to help out with certain answers, such as when respondents had a difficult time understanding the questions. Third, this procedure was significantly faster because each question did not need to be repeated for each and every person. I conducted these surveys with the assistance of one or two employees in order to communicate with the Quechua-speaking respondents.

The private individual survey was conducted privately by one of the 10 surveyors. These questions were more personal and included certain subjective questions for other related research.

The former-member survey sought to gather basic demographic information, such as location of home, cultural characteristics, religious affiliation, and social connections with members of the group. When possible, this information was gathered from current members, but otherwise was conducted in the home or business of the former member.

7.2 Formulation of Cultural Measures

For each individual a simple cultural index was calculated which equally weights four physical characteristics: hair, dress, language, and hat. For each category, the individual receives a zero, one, or two, zero being the most Western and two being the most indigenous. A borrower wearing her hair in braided pigtails receives two points, in a long and flowing style (i.e., probably recently in pigtails or easily put in pigtails) receives one point, and in a short or curled-styled receives zero points. A Spanish only speaker receives zero points, a bilingual speaker receives one point, and a Quechua-only speaker receives two points. A woman wearing an indigenous hat receives two points, while a woman with no hat receives no points. Last, a woman wearing a *pollera*, an indigenous-style skirt, receives two points, a woman wearing Western-style clothing receives zero points, and those in the middle receive one point. In total, each person receives between zero and eight points. Individuals with a score of zero or one are categorized as Western, and individuals with a score of five or more are categorized as indigenous. The results reported in this paper are robust to various formulations and combinations of these cultural measures.

7.3 Relevance of Measures of Social Connections

The cultural and geographic concentration indices are correlated with several direct measures of social connections. First, more indigenous individuals tend to sit together at group meetings. This is also true, but to a lesser extent, of the Western individuals. Similarly, individuals tend to sit next to those who live closer to them. Empirically I test this by comparing the mean probability that the person in the next seat is of the same culture to the mean probability that a randomly chosen person from the group is of the same culture. Table 2 shows this comparison in the Seating Arrangements section. For uninvited individuals, the probability rises from 23 percent to 26 percent (significant to 95 percent). For invited individuals the probability rises from 24 percent to 26 percent (significant to 95 percent). Similarly, the same comparison holds with respect to distance between members. Both uninvited and invited members live one minute and two minutes, respectively, closer to the person seated next to them (insignificant for uninvited, significant at 95 percent for invited).[34]

Second, participants reported several direct measures of social and business interactions, and these responses were correlated with both cultural and geographic dispersion. Five questions were asked: (1) how many homes they knew of others in the group, (2) from how many others they have purchased a good or service, (3) to how many others they have sold a good or service, (4) from how many others they have borrowed directly, and (5) to how many others they have lent directly. These questions are not good measures for the primary analysis, since the current information is endogenous and the questions asked in recall are both suspect and mostly invariate (few people say they knew anyone when they joined). These questions do, however, provide evidence supporting the social connection measures used in the heart of this paper. Geographic dispersion and cultural similarity are correlated with these direct measures of social connection, as shown in Data Appendix Table 1. The first question, how many homes they knew personally, is correlated significantly with both geographic proximity at 99 percent and cultural similarity at 95 percent (column 1). The second and third questions (combined) are correlated significantly with geographic proximity at 99 percent (column 2), but not with cultural similarity. The fourth and fifth questions (combined) on direct borrowing and lending also are correlated significantly with geographic proximity at 95 percent but not with cultural similarity.

[34]The distance between invited members could be less than that for uninvited for one of two reasons. First, individuals tend to invite other household members or neighbors to the bank (more so than they do by culture). Second, for logistical reasons, individuals will walk to the meetings with their neighbors or household members. Then, if walking into the meeting in a group, it would be awkward to then separate and sit apart from each other. If an immediate neighbor or household member is in the bank, then one of them most likely invited the other.

References

Adams, D., Graham, D., and Von Pischke, J., 1984. *Undermining Rural Development with Cheap Credit*. Westview Press, Boulder, CO.

Adler, Paul and Kwon, Seok-Woo, 2000. "Social Capital: The Good, the Bad, and the Ugly," in E. Lesser, ed., *Knowledge and Social Capital: Foundations and Applications*, Boston: Butterworth-Heineman, pp.89-115.

Ahlin, C. and Townsend, R., 2004. "Using Repayment Data to Test Across Models of Joint Liability Lending." Working Paper.

Alesina, A.; Baqir, R.; and Hoxby, C., 2004. "Political Jurisdictions in Heterogeneous Communities." *Journal of Political Economy* Vol. 112(2), 348-96.

Alesina, A. and La Ferrara, E., 2000. "The Determinants of Trust." NBER Working Paper #7621.

Armendariz de Aghion, Beatriz, 1999. "On the Design of a Credit Agreement with Peer Monitoring." *Journal of Development Economics*. Vol. 60(1999) 79-104.

Armendariz de Aghion, Beatriz and Gollier, Christian, 2000. "Peer Group Formation in an Adverse Selection Model." *The Economic Journal*. Vol. 110, No. 465, pp 632-643.

Banerjee, A, 2002. "The Uses of Game Theory: Against a Purely Positive Interpretation of Game Theoretic Results," MIT Working Paper.

Banerjee, A., T. Besley and T. Guinnane, 1994. "Thy Neighbor's Keeper: The Design of a Credit Cooperative with Theory and a Test." *Quarterly Journal of Economics* 109(2): 491-515.

Besley, T. and Coate, S., 1995. "Group Lending, Repayment Incentives and Social Collateral." *Journal of Development Economics* 46(1): 1-18.

Besley, T., Coate, S., and Loury, G., 1993. "The Economics of Rotating Savings and Credit Associations." *American Economic Review* 83: 782-810.

Busch, Marc and Reinhardt, Eric, "Industrial Location and Protection: The Political and Economic Geography of U.S. Nontariff Barriers." *American Journal of Political Science* 43:3.

Coleman, James S., 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology*, 94: S95-S120.

Coleman, James S., 1990. *Foundations of Social Theory*. Cambridge, MA. Harvard University Press.

Collier, Paul, 1998. “Social Capital and Poverty.” World Bank Social Capital Initiative Working Paper No. 4.

De Soto, Hernando. 2000. *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*, New York, NY. Basic Books.

Diamond, D.W., 1984. “Financial Intermediation and Delegated Monitoring.” *Review of Economic Studies* 51(3): 393-414.

Ellison, Glenn and Glaeser, Edward, 1997. “Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach.” *The Journal of Political Economy*. 105(5): 889-927.

Genicot, Garance and Ray, Debraj, 2003. “Endogenous Group Formation in Risk-Sharing Arrangements,” *Review of Economic Studies,* Vol 70(1).

Ghatak, Maitreesh, 1999. Group Lending, Local Information and Peer Selection. *Journal of Development Economics* 60:27-50.

Ghatak, Maitreesh, 2000. “Screening by the Company You Keep: Joint Liability Lending and the Peer Selection Effect.” *The Economic Journal*. 110:601-631.

Ghatak, Maitreesh and Guinnane, Timothy, 1999. “The Economics of Lending with Joint Liability: A Review of Theory and Practice,” *Journal of Development Economics* 60, (October 1999): 195-228.

Glaeser, E., Laibson, D., Scheinkman, J. and Soutter, C., 2000, “Measuring Trust,” *Quartery Journal of Economics,* August 2000.

Guiso, Luigi; Sapienza, Paola and Zingales, Luigi, 2004, “The Role of Social Capital in Financial Development,” *American Economic Review* 94: 526-556.

Kahily, M.A. and Meyer, R. 1993. “The Political Economy of Rural Loan Recovery: Evidence from Bangladesh.” *Savings and Development* 17(1): 23-38.

Karlan, Dean, 2004. “Using Experimental Economics to Measure Social Capital and Predict Real Financial Decisions,” *American Economic Review,* forthcoming.

Knack, Stephen, 1999. “Social Capital, Growth and Poverty: A Survey of Cross-Country Evidence.” The World Bank Social Capital Initiative Working Paper No. 7.

Knack, Stephen and Keefer, Philip, 1997. “Does Social Capital Have an Economic Payoff? A Cross-Country Investigation.” *Quarterly Journal of Economics,* Novermber 1997: 1251-1288.

Knack, Stephen and Zack, P. 1999 "Trust and Growth ", Working Paper.

Krishna, Anirudh and Shrader, Elizabeth, 1999. “Social Capital Assessment Tool,” unpublished manuscript, prepared for the World Bank Conference on Social Capital and Poverty Reduction.

La Ferrara, Eliana, 2003. “Kin Groups and Reciprocity: A Model of Credit Transactions in Ghana" *American Economic Review*, December 2003, 93(5): 1730-1751.

Morduch, Jonathan, 1999a, “The Role of Subsidies in Microfinance: Evidence from the Grameen Bank,” *Journal of Development Economics*.

Morduch, Jonathan, 1999b, “The Microfinance Promise,” *Journal of Economic Literature* 37 (4).

Pollack, Robert A., 1985, “A Transaction Cost Approach to Families and Households,” *Journal of Economic Literature*, XXIII, pp. 581-608.

Putnam, R. (with Robert Leonardi and Raffaella Y. Nanetti), 1993. *Making Democracy Work*. Princeton, NJ: Princeton University Press.

Rai, Ashok and Sjöström, Tomas, 2004, “Is Grameen Lending Efficient? Repayment Incentives and Insurance in Village Economies.” *Review of Economic Studies*, forthcoming.

Sadoulet, L and Carpenter, S., 1999. “Risk-Matching in Credit Groups: Evidence from Guatemala,” ECARE Working Paper.

Sharma, M. and Zeller, M., 1997. “Repayment Performance in Group-Based Credit Programs in Bangladesh: An Empirical Analysis.” *World Development* 25(10): 1731-1742.

Stiglitz, J., 1990. “Peer Monitoring and Credit Markets.” *World Bank Economic Review* 4(3): 351-366.

Uphoff, N. and Wijayaratna, C.M., 2000. “Demonstrated Benefits from Social Capital: The Productivity of Farmer Organizations in Gal Oya, Sri Lanka.” *World Development* 28:11.

van Bastelaer, T. 1999. “Does Social Capital Facilitate the Poor's Access to Credit?” World Bank Social Capital Initiative Working Paper.

Varian, H., 1990. “Monitoring Agents With Other Agents,” *Journal of Institutional and Theoretical Economics* 146(1): 153-74.

Varian, H., 2001. "Economic Scene." *New York Times*. November 21, 2001. Section C, Page 2, Column 1.

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

Yaron, J., 1994. "What makes rural financial markets successful?" *World Bank Research Observer* 9(1): 49-70.

Zeller, Manfred, 1998, "Determinants of Repayment Performance in Credit Groups: The Role of Program Design, Intragroup Risk Pooling, and Social Cohesion," *Economic Development and Cultural Change*, 46(3): 599-620.

Table 1: Individual Summary Statistics
Means

	Method of Arrival to Group	
	Uninvited	Invited
	(1)	(2)
LOAN DATA		
Proportion of loans with default	0.203	0.160
	(0.016)	(0.011)
Default (cond. on default > 0), US$	69.157	62.867
	(8.407)	(4.038)
Default as proportion of approved FINCA loan (cond. on default > 0)	2.797	2.229
	(0.222)	(0.123)
Initial savings, US$	37.800	39.098
	(2.601)	(1.342)
New savings deposits (both required & voluntary), US$	59.121	58.690
	(3.059)	(1.941)
Dropout after first loan cycle, proportion	0.240	0.243
	(0.017)	(0.129)
DEMOGRAPHIC DATA		
Female	0.989	0.996
	(0.004)	(0.002)
Age	34.101	32.102
	(0.495)	(0.389)
Spouse	0.534	0.565
	(0.020)	(0.149)
Completed high school	0.317	0.287
	(0.188)	(0.136)
Individuals	616	1,103
Average number of loan cycles per individual	3.13	2.67

Standard errors of estimated means reported in parentheses.

Table 2: Individual Geographic and Cultural Measures
Means and Standard Deviations

	Uninvited to Group		Invited to Group	
	Mean & Std Error	Std Dev & # of Obs	Mean & Std Error	Std Dev & # of Obs
	(1)		(2)	
DISTANCE DATA (units in minutes walking)				
*Distance from current member to original members of group	13.501	9.928	13.106	7.569
	(0.400)	n=616	(0.231)	n=1075
Distance from current member to members of other groups	13.704	9.632	14.109	7.366
	(0.388)	n=616	(0.225)	n=1075
*Prob(Member from original group lives within 10-minute walk of home)	0.224	0.236	0.205	0.203
	(0.009)	n=616	(0.006)	n=1075
Prob(Person from other group lives within 10-minute walk of home)	0.211	0.214	0.172	0.170
	(0.009)	n=616	(0.005)	n=1075
Distance to FINCA office (town center)	9.565	10.905	9.858	8.957
	(0.439)	n=616	(0.273)	n=1075
CULTURE DATA				
Culture score (0=Western, 8=Indigenous)	2.537	2.224	2.610	2.172
	(0.090)	n=616	(0.066)	n=1075
*Prob(Member from original group is of same culture as individual)	0.201	0.156	0.190	0.140
	(0.006)	n=616	(0.004)	n=1075
Prob(Person from other group is of same culture as individual)	0.185	0.118	0.172	0.098
	(0.005)	n=616	(0.003)	n=1075
SEATING ARRANGEMENTS				
Distance to current members of group	15.795	11.588	15.081	8.984
	(0.612)	n=358	(0.357)	n=632
Distance to persons seated next to each other in meeting	14.953	13.395	13.086	10.503
	(0.708)	n=358	(0.418)	n=632
Prob(Person from same group is of same culture)	0.231	0.141	0.239	0.134
	(0.007)	n=358	(0.005)	n=632
Prob(Person in next seat in meeting is of same culture)	(0.262)	0.3016	(0.261)	0.302
	(0.016)	n=358	(0.012)	n=632

*Variables with asterisks are the key independent variables used in the specifications in Tables 4, 5, and 6. For each variable, the table reports the mean, the standard error of the estimate of the mean, the standard deviation, and the number of observations. Units for distance measures are in minutes walking distance.

Table 3: Group Summary Statistics
Means and Standard Deviations

	Method of Arrival to Group			
	Uninvited to Group		Invited to Group	
	Mean & Std Error	Std Dev & # of Obs	Mean & Std Error	Std Dev & # of Obs
GEOGRAPHIC CONCENTRATION				
*Average distance to original members from current members (minutes)	12.422	5.342	12.413	4.751
	(0.824)	n=42	(0.733)	n=42
*Average percent of original members who live within 10 minutes of current member	0.239	0.200	0.243	0.196
	(0.031)	n=42	(0.030)	n=42
GD: Geographic concentration	0.147	0.104	0.252	0.174
	(0.016)	n=42	(0.027)	n=42
E(GD): Expected geographic concentration	0.127	0.090	0.203	0.168
	(0.014)	n=42	(0.026)	n=42
CULTURAL CONCENTRATION				
*Average percent of original members of same culture as current member	0.197	0.098	0.212	0.111
	(0.015)	n=42	(0.017)	n=42
CD: Cultural concentration	0.119	0.136	0.184	0.155
	(0.021)	n=42	(0.024)	n=42
E(CD): Expected cultural concentration	0.106	0.078	0.167	0.136
	(0.012)	n=42	(0.021)	n=42

*Variables with asterisks are the key independent variables used in the specifications in Table 7.

All results calculated on original group members only.

$$GD_{group} = \sum_{neighborhoods} (s_i - x_i)^2,$$

where s_i is the share of the group from neighborhood i and x_i is the share of the general population from neighborhood i.

$$E(GD)_{group} = [1 - \sum_{neighborhoods} (x_i)^2] / n$$

CD and E(CD) are constructed identically to GD and E(GD), except by cultural group rather than neighborhood.

The Alesina index for cultural concentration is equal to the sum of squared shares of each cultural group.

Table 4: Individual Default
OLS, Tobit, and Probit

	Dependent variable: Percent of loan in default at end of cycle					
	1st Loan Only			All Loans		
	OLS	Tobit	Probit	OLS	Tobit	Probit
	(1)	(2)	(3)	(4)	(5)	(6)
Distance from individual's home to original members of group	0.019	0.343	0.019	0.049	0.297	0.040
	(0.077)	(0.342)	(0.019)	(0.068)	(0.024)	(0.027)
	n=616	n=616	n=616	n=1801	n=1801	n=1801
Percent of original members within 10-minute walk of individual's home	-1.536 ***	-6.077 ***	-0.284 ***	-1.556 ***	-3.754 ***	-0.367 ***
	(0.391)	(1.795)	(0.079)	(0.370)	(1.078)	(0.134)
	n=616	n=616	n=616	n=1801	n=1801	n=1801
Percent of original members with same culture as individual	-0.534 *	-4.230 **	-0.200 ***	-0.396	-1.458	-0.177
	(0.301)	(1.791)	(0.069)	(0.308)	(1.116)	(0.111)
	n=616	n=616	n=616	n=1801	n=1801	n=1801

*** 99% significance; ** 95% significance; * 90% significance
Each cell is a separate specification.
Standard errors corrected for clustering at the group level in all specifications.
Individuals weighted evenly "all loans" specifications.
Individual level specifications include the following control variables (See Appendix Table 2 for results on control variables):
Distance to FINCA (town center), town dummy, neighborhood dummies, age, education, marital status, siblings, children, # in household, year, and age of group when individual joined.
Loan size estimated using approved loan amount, which is savings balance at end of prior cycle.

Table 5: Individual Savings
OLS

	Total Savings Deposits	Mandatory Savings Deposits	Voluntary Savings Deposits	Individual Savings Deposits
	1st Loan Only	1st Loan Only	1st Loan Only	All Loans
	(1)	(2)	(3)	(4)
Distance from individual's home to original members of group	-9.681 *	-3.259 *	-6.428 *	-7.192 **
	(4.995)	(1.749)	(3.702)	(3.294)
	n=616	n=616	n=616	n=1801
Percent of original members within 10-minute walk of individual's home	15.623	21.154 ***	-5.531	27.001
	(23.565)	(5.762)	(22.584)	(20.304)
	n=616	n=616	n=616	n=1801
Percent of original members with same culture as individual	-11.573	8.335	-19.908	-13.590
	(29.700)	(7.392)	(26.783)	(23.721)
	n=616	n=616	n=616	n=1801

*** 99% significance; ** 95% significance; * 90% significance
Each cell is a separate specification.
Standard errors corrected for clustering at the group level in all specifications.
Individuals weighted evenly "all loans" specifications.
Individual-level specifications include the following control variables:
Distance to FINCA (town center), town dummy, neighborhood dummies, age, education, marital status, siblings, children,
in household, year, and age of group when individual joined.

Table 6: Dropout
Probit

	Dependent Variable = 1 if Member Dropped Out After 1st Loan					
	(1)	(2)	(3)	(4)	(5)	(6)
Default	0.115 ***	0.112 ***	0.113 ***	-0.023	0.153 ***	0.197 ***
	(0.037)	(0.037)	(0.036)	(0.056)	(0.043)	(0.041)
Total accumulated savings	-0.013	-0.016	-0.014	-0.014	-0.015	-0.013
	(0.014)	(0.014)	(0.014)	(0.015)	(0.014)	(0.013)
Distance from individual's home to original members of group	0.037			0.023		
	(0.031)			(0.032)		
Distance interacted with default				0.074 ***		
				(0.027)		
Percent of original members within 10-minute walk of individual's home		-0.007			-0.004	
		(0.006)			0.006	
Percent within 10-minute walk Interacted with default					-0.132 ***	
					(0.047)	
Percent of original members with same culture as individual			-0.192			0.000
			(0.157)			(0.144)
Culture interacted with default						-0.332 ***
						(0.094)
Observations	616	616	616	616	616	616
# of dropouts	148	148	148	148	148	148
Log-likelihood	-173.47	-173.76	-173.28	-167.39	-171.58	-166.78
Groups	42	42	42	42	42	42

*** 99% significance; ** 95% significance; * 90% significance

Marginal effects of probit reported.

Standard errors corrected for clustering at the group level.

Individual-level specifications control variables for distance to FINCA (town center), town dummy, neighborhood dummies, age, education, marital status, siblings, children, # in household, year, and age of group.

Table 7: Group Outcomes
Default, Savings and Dropout
OLS

	Average % Default in Group	Total Savings Deposits	Mandatory Savings Deposits	Voluntary Savings Deposits	Percent Return on Savings	% Dropout from Program
	(1)	(2)	(3)	(4)	(5)	(6)
Average distance of original members to current, uninvited members	0.156 **	0.162	2.688	-2.526	0.000	0.055 *
	(0.066)	(3.110)	(2.075)	(2.170)	(0.003)	(0.030)
	n=245	n=245	n=245	n=245	n=245	n=245
Average percent of original members within 10-minute walk of current, uninvited members	-1.290 ***	57.738 **	18.094	39.644 **	0.040 **	-0.441 *
	(0.426)	(21.905)	(11.620)	(15.559)	(0.019)	(0.265)
	n=245	n=245	n=245	n=245	n=245	n=245
Average probability that original member is of same culture as current, uninvited member	-0.348	-25.835	-3.722	-22.113	-0.017	-0.348 *
	(0.562)	(31.488)	(13.747)	(21.476)	(0.028)	(0.198)
	n=245	n=245	n=245	n=245	n=245	n=245

*** 99% significance; ** 95% significance; * 90% significance
Each cell is a separate specification.
Standard errors corrected for clustering at the group level in all specifications.
Groups weighted evenly.
Specifications include the following control variables (See Appendix Table 2 for results on control variables):
Average distance to FINCA/town center (for geographic proximity), % who live within 10-minutes of town center (for geographic proximity), % indigenous (for cultural similarity), % Western (for cultural similarity), town dummy, average age, average education, average # in household, average siblings, average # of children, year, and age of group
Loan size estimated using approved loan amount, which is savings balance at end of prior cycle.

Table 8: Survey of Current Members about Dropouts/Defaults
Summary Tabulations

PANEL A

	All Freq	All %	With Default Freq	With Default %	Without Default Freq	Without Default %
Observations						
Number of lending groups	28					
Number of current members	459					
Number of dropouts/defaults (all dropouts/defaults from prior two loan cycles)	575		119	20.7%	456	79.3%
Number of dropouts	550		117	21.3%	433	78.7%
Number of pairwise relationships between current members and dropouts/defaulters	9,337		1900	20.3%	7439	79.7%
Number of defaulters who remained in the program	44					
Basic information on relationships						
Number of pairwise relationships	9,337	100.0%	1900	20.3%	7439	79.7%
Instances of recognizing name of dropout (hence interview continued)	4,073	43.6%	778	19.1%	3295	80.9%
Lived within 10-minute walk of the dropout	587	6.3%	86	14.7%	501	85.3%
Family member	121	1.3%	14	11.6%	107	88.4%
Knew the dropout before joining the lending group	431	4.6%	59	13.7%	372	86.3%
Has visited the business of the dropout	107	1.1%	7	6.5%	100	93.5%
Culture score within 1 point	1,344	14.4%	225	16.7%	1119	83.3%
Accuracy of information on default						
Current member thought that the dropout left with default	475	5.1%	305	64.2%	170	35.8%
Current member thought that the dropout left without default (& remembered individual)	3,598	88.3%	473	13.1%	3125	86.9%

PANEL B

Change in Relationship: Current members reporting about dropouts	Obs	Worse Freq	Worse %	Same Freq	Same %	Better Freq	Better %
Friendship, if respondent reported that dropout left without default	8,862	245	2.8%	8,598	97.0%	19	0.2%
Friendship, if respondent reported that dropout left with default	475	57	12.0%	418	88.0%	-	0.0%
Trust, if respondent reported that dropout left without default	8,862	87	1.0%	8,765	98.9%	10	0.1%
Trust, if respondent reported that dropout left with default	475	28	5.9%	447	94.1%	-	0.0%
Buying/selling, if respondent reported that dropout left without default	8,862	65	0.7%	8,791	99.2%	6	0.1%
Buying/selling, if respondent reported that dropout left with default	475	10	2.1%	464	97.7%	1	0.2%
Speaking outside of meeting, if respondent reported that dropout left without default	8,862	554	6.3%	8,287	93.5%	21	0.2%
Speaking outside of meeting, if respondent reported that dropout left with default	475	88	18.5%	386	81.3%	1	0.2%

Table 9: Loan Monitoring, Direct Evidence
Results from Survey of Current Members about Recent Defaulters and Dropouts
Probit

Binary Dependent Variables:	Respondent Correctly Reported Whether Dropout Had Default		Respondent Reported Knowing Why Dropout Had Default	
	(1)	(2)	(3)	(4)
Culture score within 1 point of dropout/defaulter	0.049*		0.028	
	(0.027)		(0.045)	
Absolute difference between culture scores		-0.014**		-0.018**
		(0.006)		(0.009)
Lives within 10 minutes of other dropout/defaulter	0.008	0.007	-0.022	-0.024
	(0.031)	(0.031)	(0.025)	(0.024)
Respondent is Indigenous (relative to "mixed")	-0.034***	-0.033***	-0.028**	-0.027**
	(0.012)	(0.012)	(0.014)	(0.014)
Respondent is Western (relative to "mixed")	-0.105***	-0.100***	-0.052***	-0.049***
	(0.017)	(0.017)	(0.016)	(0.016)
Dropout/defaulter is Indigenous	0.009	0.016	-0.003	-0.008
	(0.044)	(0.044)	(0.050)	(0.046)
Dropout/defaulter is Western	0.078	0.098	0.045	0.105
	(0.059)	(0.061)	(0.082)	(0.096)
Extended family member	0.087	0.086	0.002	0.001
	(0.095)	(0.095)	(0.076)	(0.073)
Knew the dropout/defaulter before being a member	0.491***	0.491***	0.173**	0.169**
	(0.028)	(0.028)	(0.072)	(0.072)
Has visited the business of the dropout/defaulter	0.594***	0.594***	0.260**	0.267**
	(0.029)	(0.029)	(0.133)	(0.131)
Mean of Binary Dependent Variable	0.3674	0.3674	0.0757	0.0757
Number of Observations	9337	9337	1900	1900
Pseudo R-Squared	0.0563	0.0566	0.0332	0.0383

In all specifications, standard errors corrected for clustering across observations regarding the same dropout/default individual.
"Similar" culture score defined as within 1 point of each other after scoring each on a scale of 0 to 8, western to indigenous.

Indicator variable included to capture any missing culture data.
Marginal effects reported for coefficients in probit model.
* significant at 10%; ** significant at 5%; *** significant at 1%

Data Appendix Table 1:
Correlations between Geographic and Cultural Concentrations and Direct Social Capital Measures
Tobit

	Homes known of members when joined (1)	Number of members with whom client has bought or sold goods (2)	Instances of direct borrowing or lending between members (3)
Average distance of original members of group	-0.005 ***	-0.014 ***	-0.008 **
	(0.001)	(0.004)	(0.003)
Percent of original members within 10-minute walk	1.544 *	1.656	2.414
	(0.878)	(2.037)	(3.326)
Percent of original members with same culture	1.857 **	-1.091	2.186
	(0.729)	(2.221)	(2.059)
# of observations censored at zero	227	300	538
Observations	948	948	946

*** 99% significance; ** 95% significance; * 90% significance

Each column represents a separate tobit specification with the social interaction measure as the dependent variable.

Standard errors corrected for clustering at the group level.

Includes controls for neighborhood, distance to FINCA, and culture score.

Data Appendix Table 2:
Control Variables Results from Default, Savings, and Dropout Tables
Tobit, OLS, and Probit

	Default		Total Savings		Dropout	
	Typical Results (1)	Probit (2)	Typical Results (3)	OLS (4)	Typical Results (5)	Probit (6)
Indigenous	mixed	0.031	neg	-6.940	pos	0.071
Western	pos *	0.015	pos	2.750	pos	0.009
Distance to town center	mixed	0.007	pos **	0.014 *	pos	0.010
Ayacucho	neg *	-0.337 *	pos	24.992	neg ***	0.490 ***
# of children = 0	pos	0.031	neg	-4.740	pos	0.068
# of children	pos	0.007	neg	-0.529	neg	-0.011
Age	neg	-0.006	pos	0.374	neg	-0.004
Age-squared	pos	0.000	neg	-0.001	pos	0.000
Spouse	pos **	0.057 **	neg	-4.766	pos ***	0.092 ***
Finished high school	neg	-0.022	pos	6.724	neg	-0.019
# of siblings	neg	0.000	pos *	2.129	neg	-0.003
# of women in household	mixed	0.006	neg	-2.650	neg **	-0.034 *
# of men in household	mixed	-0.005	pos	0.555	neg **	-0.043 **

*** 99% significance; ** 95% significance; * 90% significance

"Typical Results" summarizes the typical result across the various permutations of specifications, which depend on which measure of social capital is included and, in the case of default, whether a tobit, probit or OLS, is employed.

The representative examples in columns 2, 4, and 6 use the second geographic dispersion measure, % who live within a 5-minute walk.
Column 2 corresponds to Table 4, Column 3, Row 2.
Column 4 corresponds to Table 5, Column 1, Row 2.
Column 6 corresponds to Table 6, Column 5.

Appendix Table 3: Qualitative Responses on Monitoring of Default

Why did X not repay her loan?

Do not know	260	54.6%
Business was not going well	51	10.7%
Health	50	10.5%
Family problems	46	9.7%
Travel	31	6.5%
Robbery	13	2.7%
She lent it to someone else	8	1.7%
Legal problems	4	0.8%
Death in family	3	0.6%
Did not want to pay	3	0.6%
Studies	3	0.6%
Had other debt	1	0.2%
She was a con artist	1	0.2%
Stopped working	1	0.2%
Work	1	0.2%
	476	100.0%

Why was X allowed to remain in the group even after she had default?

	Evidence of Monitoring?		
Do not know	no	174	52.3%
Bank needed members	no	44	13.2%
Talked to the members	yes	25	7.5%
Talked to director	yes	18	5.4%
Had family problems (sickness, accident)	yes	16	4.8%
She was responsible/punctual (in paying)	yes	14	4.2%
She lent it to someone else	yes	11	3.3%
She got sick	yes	7	2.1%
She wanted to stay	no	4	1.2%
She was traveling	yes	4	1.2%
Trust	yes	4	1.2%
Said they would improve/be more responsible	yes	3	0.9%
Talked to credit officer	yes	3	0.9%
Robbery	yes	2	0.6%
Business was not going well	yes	1	0.3%
Son left for schooling	yes	1	0.3%
Had an accident	yes	1	0.3%
Car broke down	yes	1	0.3%
		333	100.0%

Data come from 2001 survey of current members. Each member was asked privately about the default of all members who had default in the prior two loans.

Left table represents 206 different individuals. Of those 206 individuals, 100 had at least one person report why she did not repay her loan.

Right table represents 44 different individuals. Of those 44 individuals, 38 had at least one person report an "monitoring" explanation for her remaining in the program.