The Supply- and Demand-Side Impacts of Credit Market Information

Alain de Janvry*, Craig McIntosh**, and Elisabeth Sadoulet*

September 2009

Abstract

We utilize a unique pair of experiments to isolate the ways in which reductions in asymmetric information alter credit market outcomes. A Guatemalan microfinance lender gradually started using a credit bureau across its branches without letting borrowers know about it. One year later, we ran a large randomized credit information course that described the existence and workings of the bureau to the clients of this lender. This pairing of natural and randomized experiments allows us to separately identify how new information enters on the supply and the demand sides of the market. Our results indicate that the credit bureau generated large efficiency gains for the lender, and that these gains were augmented when borrowers understood the rules of the game. The credit bureau rewarded good borrowers but penalized weaker ones, increasing economic differentiation.

JEL Classifications: O16, O31, G14
Keywords: Credit markets, Credit Bureau, Asymmetric Information, Randomized Experiment

1. INTRODUCTION

Asymmetric information problems bedevil credit transactions, and these problems are particularly severe in poverty-focused credit markets where borrowers lack collateral and credit histories. A broad range of social capital-based lending mechanisms have emerged in recent years to overcome these problems, helping more than a hundred million previously unbanked borrowers enter credit markets over the past decade. Ironically, this surge in lending is increasingly undermining the very mechanisms that made such lending possible, as competing lenders reduce each others' ability to contract without formal collateral. In this environment, credit bureaus can become an attractive means for reducing asymmetric information.

We use a unique confluence of data and identification methods to analyze how lending outcomes have responded to the introduction of a credit bureau in Guatemala's microfinance market. In August of 2001 a major microfinance lender began to install hardware permitting branches to communicate information with the bureau, a process that was completed in ten waves over the course of 18 months. The lender did not inform borrowers of the use of the bureau, however, and we found knowledge of it to be almost non-existent in a survey of borrowers implemented after the rollout of the bureau was complete. We therefore conducted a randomized training campaign in which we informed 5,000 borrowers of the use of the system, how the bureau worked, and the opportunities and risks that it presented for them. We then used institutional data from the lender and from the bureau itself to track how a variety of lending outcomes emerged from this unusual structure in which asymmetric information was reduced on the two sides of the market at two different points in time. We are thus able to disentangle the supply- and demand-side impacts of credit market information.

A new experimental literature has sprung up in recent years working to separate the effects on credit markets of moral hazard and adverse selection as they operate through the price mechanism. Karlan & Zinman (2009) use a two-stage experiment in which borrowers are first offered randomized interest rates in order to measure adverse selection effects, and then among those who come for loans the rate is further randomized downward by surprise in order to isolate moral hazard. This provides an experimental decomposition of the impact of interest rates on hidden information and hidden actions, a relationship originally posited by Stiglitz & Weiss (1981). What we offer here, however, is the estimation of a more primal

1

relationship, because we are able to see how asymmetric information alters outcomes in credit markets *directly,* rather than via the indirect mechanism of interest rates. That is to say, rather than holding the overall amount of asymmetric information constant and using prices to alter who ends up on the credit market, bureaus offer lenders a vehicle through which they can improve selection and alter behavior based on a first-order reduction in the quantity of asymmetric information in the marketplace.[1]

Before the existence of the bureau, borrowers were supposed to disclose information about past defaults and current debts on their loan screening forms. The improved screening ability generated by use of the bureau results in large increases in profitability for the lender, indicating that there was strategic behavior by potential clients over past loan information and that it was indeed creating a substantial adverse selection problem for the lender. The specific sequence through which the bureau was rolled out for use by the lender and information about the bureau randomly provided to borrowers allows us to identify separately the roles of adverse selection, moral hazard, and incentives on group composition. This two-fold experiment – natural and randomized – allows to perform this identification that is, to our knowledge, new to the literature.

To formalize these tests, we first develop a simple theoretical model of the lender's credit scoring problem, and use it to understand the effects of the new information revealed in the bureau. We capture in the model an unusual feature of microfinance credit bureaus, which is that they report on the behavior of *groups* rather than *individuals* when loans are made to jointly liable borrowers. Because the lender can observe and correlate demographic characteristics and outcomes from its own clientele in a scoring model, the *new* information revealed by the bureau is orthogonal to what could previously be predicted. Despite this, we show that the probability of the lender selecting a borrower in (out) as a result of the bureau is increasing (decreasing) in the pre-bureau score, and that the variance of the prediction error on the client quality is an increasing function of the size of the group in which a borrower takes loans. We then specify how borrowers respond to the bureau both through a reduction in moral hazard and through altering the process by which they screen new members in joint liability groups.

---

[1] In this sense our asymmetric information problem is more similar to Navajas et al. (2003), who consider a lender choosing between a costly but effective screening technology and joint liability screening.

We find impacts of informational changes on both sides of the market. In terms of adverse selection, the ejection rate rises by 15% when the lender first uses the bureau, but this impact on the size of the loan portfolio is more than compensated for by new loans made to borrowers to whom the institution had never previously lent. Borrowers selected using the credit bureau are better clients, with better repayment performance and higher growth of future loans. Ongoing clients who are not ejected are able to take larger loans but their performance exhibits a small deterioration. The selection process benefits more women than men. In terms of moral hazard, the repayment performance shows a modest and temporary improvement when borrowers become aware of the bureau. Groups then also exhibit an adverse selection effect, ejecting some worse-performing members, and unexpectedly it is now women who lose access to credit more than men. We use data from the bureau itself to demonstrate that the training induces a 10% jump in the probability that a client will take a loan from an outside lender, and we see sharp differences across borrower types in the ability to handle this surge in total debt. Overall, the bureau permits a substantial expansion of credit among lenders in the system while simultaneously driving down delinquency.

2. REDUCING ASYMMETRIC INFORMATION

Credit bureaus are a formalized approach to information sharing among lenders. In the absence of such an institution, lenders of non-collateralized loans must resort to an array of informal mechanisms ranging from joint liability (Besley & Coate, 1995; Ghatak & Guinnane, 1999; Gine & Karlan 2006), to relationship banking, and to information sharing in the 'credit officer lunch' (McIntosh et al, 2006). Within social networks, North (1990) and Greif (1994) show that the transfer of reputation and multilateral punishment can enable a 'localized honesty equilibrium' in spite of the individual pursuit of self-interest. The broadening of this equilibrium to a 'generalized honesty equilibrium' supporting anonymous exchange in very large groups requires the emergence of institutional innovations to formalize the transfer of reputation and the sharing of information (Platteau, 2000). Microfinance markets provide a particularly interesting environment in which to study the impact of this formalization, both because of the rapid creation of bureaus in recent years and because microfinance provides a continuous range of contracts across which informal and formalized tools are interchanged (Navajas et al., 2003; Morduch, 1999; Morduch and Armendariz de Aghion, 2005).

The decision to form a bureau is one fraught with strategic risks for lenders (Padilla & Pagano, 1997). Lenders are of course happy to see the data of others, and in general do not have problems sharing 'negative' information (on default), but sharing their own 'positive' information (on current lending) creates the possibility that their competitors will try to cherry-pick their own best clients (Gehrig & Stenbacka, 2007). Set against this peril are expected benefits from a decrease in portfolio risk (Campion & Valenzuela, 2001; Jappelli & Pagano, 1999), restraining multiple contracting by borrowers (McIntosh et al, 2006), and the preservation of reputation effects through the formation of long-term credit histories (Vercammen, 1995; Padilla & Pagano, 2000)[2].

Concern over a rising level of default in the Guatemalan market led the country's three major microfinance lenders (Genesis, BanCafé, and Banrural) to agree to the formation of a bureau (called Crediref) in 2001.[3] Strategic fears about the use of the bureau were alleviated through several simple mechanisms. First, only institutions that share information into Crediref are allowed to consult it. Secondly, the system does not allow users to identify the lender who issued the loan, respecting lender information privacy. The bureau has attracted smaller lenders with the passage of time, and now contains data from 14 different institutions.

Our setting does not allow us to observe directly the quantity of asymmetric information in the marketplace. Genesis is a well-respected lender that used the typical labor-intensive practices for screening loans prior to the creation of the bureau, including detailed application forms and visits to the enterprise and home of the applicant. The bureau includes the majority, but not all of, the major microfinance lenders and contains no information on informal loans, utilities payments, or commercial banking activity. Therefore substantial information existed in the market prior to the use of the bureau, and even subsequent to its advent asymmetric information persisted. Nonetheless, we are able to conclusively detect the implications of improved information in this marketplace. When the lender begins using the bureau, average loan sizes increase and the overall lending portfolio increases without an increase in default (interest rates, typical of the microfinance sector, remain fixed over the short run). This pattern, in the absence of any real awareness of the existence of the bureau

---

[2] Note that both of these papers discuss the possibility that sharing *too much* information (too long a history in Vercammen (1995) and information over types in Padilla & Pagano (2000)) can create a disincentive to effort.

[3] BanCafé and Banrural are both national full-service banks which only share microlending information in Crediref, and not information from their commercial banking divisions.

among clients, indicates decreased adverse selection. When we train borrowers mid-loan on the existence of the bureau, we see a modest repayment improvement within a given client base on loans administered prior to the training. This is evidence of pure moral hazard in the marketplace. Finally, the pattern of turnover in groups changes after the training, indicating that a form of adverse selection exists in the group selection process as well, despite the fact that theory typically predicts such groups to be perfectly informed over each others' behavior.

Our study follows Genesis which has 39 branches distributed over most of Guatemala. This institution staggered the entry of its branches into Crediref over the period between August 2001[4] and January 2003 because a team of computer experts had to be sent branch-by-branch to set up the hardware, software, and networks required to perform ongoing credit checks and the monthly data exchange. While this rollout was not randomized, we perform a variety of tests on the sequencing and find that it was not endogenous to outcomes. Due to concerns over client reaction to this new bureau, the rollout was not accompanied by any publicity campaign. Surveys we conducted in the summer of 2003 found virtually no knowledge of the system on the part of clients.[5] Having suggested to the lender that this situation was likely to be inefficient as well as inequitable, we worked together with the training department of Genesis to design materials for an educational program.

The randomized training program was conducted from June to November of 2004. For logistical reasons, we trained only group borrowers. We stressed the expanded importance of behavior with one lender to the contracts now available from other lenders. This point was made both in a negative fashion (meaning that repayment problems with any participating lender will decrease options with other lenders) and in a positive fashion (emphasizing the greater opportunities now available for climbing the 'credit ladder' for those who repay well).[6] The training also discussed the information made available by the credit bureau and the timing and circumstances under which checks of credit bureau records were conducted by member institutions.

---

[4] Between August 2001 and March 2002 the bureau was a precursor to Crediref known as Infornet, whose information was later subsumed into Crediref.

[5] Luoto et al. (2007). Note that if borrowers understood *some* component of the new system, then the rollout of the bureau is accompanied by a muted moral hazard response from borrowers as well as the starker screening effects..

[6] Galindo & Miller, 2001. As a cautionary tale of the unpredictable consequences of training programs, Schreiner (1999) finds that the randomized Unemployment Insurance Self-Employment Demonstration actually *discouraged* the most disadvantaged from entering self-employment.

The way in which borrowers respond to the bureau will depend on the contracts and information in use at the time the bureau is introduced. Genesis extends loans to individuals, to Solidarity Groups (SGs) of 2-5 members, and to large Communal Banking (CB) groups. CB loans are uncollateralized and rely entirely on joint liability and dynamic incentives to ensure repayment. Thus CB borrowers are likely to be relatively credit constrained and to be unable (prior to the bureau) to readily substitute credit from one lender with credit from another. The bureau is only rarely used to check CB clients, with 2.5% of loans checked. SG borrowers, on the other hand, are checked just as frequently in the bureau as individual borrowers (63% of loans), and collateral and cosigning are commonly used. Thus SG clients are likely to be less credit constrained and to be more able to obtain credit from outside lenders, even in the absence of the bureau. Because individual borrowers have no joint liability at all, we expect asymmetric information to be most severe for these borrowers, and hence the bureau should have the sharpest impacts on their selection, behavior, and access to loans.

For group loans, the bureau reports repayment and indebtedness information on *the group as a whole*, rather than on each member of the group separately. If an individual who is taking a loan in a group of 5 is checked in the bureau, the check will reveal the size of the group, the characteristics of the loan (loan size, term), and the repayment performance for the whole group, but not for that specific person. Clearly, this noisy measure of individual quality makes the bureau less informative about people who have taken loans only in large groups. These large groups are supposedly governed by joint-liability lending rules, but often credit officers are too well informed to punish an entire group because of a single bad individual, as the rules dictate they should. We suggest that the blunt reporting of group information into the bureau will in fact strengthen the joint liability mechanism, and hence the incentives for groups to be choosy about their membership, reducing adverse selection at the group level. Good groups should try to expel bad members and good members of bad groups should be trying to leave them.

We now formalize the intuition of a bureau's impact on behavior of the lender and borrowers with a simple theoretical model.[7]

---

[7] In this paper we focus on the details of statistical discrimination when a bureau is used in a group lending environment. For a theoretical environment which decomposes the adverse selection and moral hazard effects of an improvement in lender information for individual borrowers, see McIntosh & Wydick (2007).

3. THEORY

3.1. LENDER'S SELECTION OF BORROWERS AND USE OF THE BUREAU.

In this section, we model the screening process applied by the lender and show how the use of the bureau affects client selection. The screening process is formalized as the use of a scoring model based on observable characteristics collected by the lender and on information available in the credit bureau, when in use. We assume that the pool of borrowers is unaffected by existence of a credit bureau, in accordance with the context of the empirical work that follows in which clients were unaware of the existence of the credit bureau. Each borrower is characterized by a set of observable variables X. Assuming that the lender uses a linear regression to form the 'credit score', we can write the expected profitability of a given borrower conditional on the observable information as $E\left(\pi^L \mid X\right) = X\beta$. The true, *ex post* profitability of lending π^L to an individual with observable characteristics X is given by $\pi^L = X\beta + a$. Here *a,* the unobserved quality, must be mean-zero and orthogonal to X because the score, $X\beta$, exhausts all available information (meaning that any unobservable attributes will project into β to the extent that they are correlated with X).

The credit bureau provides a signal on client quality. Lenders using group-lending credit bureau information, however, face a fundamental problem in interpreting how much information over individual quality can be inferred from this group behavior. Two primary factors drive this: first, the size of the group. Because an individual has an increasingly strong effect over the group outcome as the size of the group decreases, repayment behavior from small groups is inherently more informative than repayment from large groups. The second factor is the extent of assortative matching which exists at the group level. In a population which featured no assortative matching, the lender would recognize that the repayment performance of large groups was completely stochastic and therefore contains no information over individual quality. If, however, borrowing groups feature assortative matching over *unobserved* components of borrower quality, then group repayment performance continues to contain some information about the expected quality of a borrower even as the group size becomes very large.

To capture these features, we parameterize the problem as follows: we assume that groups match assortatively around an intended average group quality $a_g \sim N(0, \sigma_g^2)$. We define assortative matching over the unobserved quality a because, to the extent that matching takes place on observables X, it has already been captured in the pre-bureau credit score $X\beta$. This process of assortative matching is imperfect, however, and the mismatch between an individual's actual quality and the intended quality of the group has two components: a predictable element that would be revealed from a credit bureau information $u \sim N(0, \sigma_u^2)$, and a completely unpredictable component of repayment behavior given by $\varepsilon \sim N(0, \sigma_\varepsilon^2)$.

From here it remains only to observe that $\varepsilon \perp \{a_g, u\}$ because ε is by definition the residual from the observation in the credit bureau, and that $u \perp a_g$ by its definition as the residual from the intended group quality. Therefore we can write $a = a_g + u + \varepsilon$, and because each of these three component terms is mean-zero and independent, $a \sim N(0, \sigma_a^2 = \sigma_g^2 + \sigma_u^2 + \sigma_\varepsilon^2)$. This is equivalent to saying that the average group quality works like a random effect in individual quality. The strongest predictive information over individual repayment therefore consists of the set $\{X, a_g, u\}$.

The new information about an individual revealed by the group-lending credit bureau information is the group average of the predictable components of quality: $\bar{\alpha} = a_g + \bar{u}$, and the size of the group from which the signal comes, N. Given the orthogonality of a_g, u, and ε, the group signal $\bar{\alpha} \sim N\left(0, \sigma_{\bar{\alpha}}^2 = \sigma_g^2 + \sigma_u^2/N\right)$. In addition, $E\left(a | \bar{\alpha}\right) = \bar{\alpha}$, and estimating a linear scoring model for individual behavior, conditional on X and $\bar{\alpha}$ would lead to:

$$E\left(\pi^L | X, \bar{\alpha}\right) = X\beta + \bar{\alpha}.$$

The observed average group quality $\bar{\alpha}$ is a signal on a, and its precision given by $\operatorname{corr}\left(a, \bar{\alpha}\right) = \dfrac{\sigma_{\bar{\alpha}}}{\sigma_a}$ is increasing in the variance of the group signal relative to the total variance, therefore decreasing in the number of members in the group from which the signal comes.

Thus,

- As $N \to \infty$, $\bar{u} \to 0$ and so $\bar{\alpha} \to a_g$; meaning that only the intended average group quality and nothing about individual deviations can be inferred from the bureau. This is still a meaningful signal on an individual's quality.

- As $N \to 1$, $\bar{\alpha} \to a_g + u$, which is the strongest signal over individual quality that can be observed through the bureau (because ε is fundamentally unpredictable).

Based on these expressions, we now analyze which types of clients will be most affected by use of the credit bureau, either positively as differentially screened in with the bureau while they would not have been selected without the bureau, or negatively when the bureau reveals information that leads to their rejection.

Borrowers differentially screened in and out with use of the bureau

Borrowers differentially screened *in* using bureau information are those for whom $0 < -X\beta \le \bar{\alpha}$. Hence the proportion of the applicant pool that is offered a loan as a consequence of the use of the credit bureau is:

$$\Pr\left(X\beta + \bar{\alpha} \ge 0 \mid X\beta < 0\right) = \Phi\left(\frac{X\beta}{\sigma_{\bar{\alpha}}}\right).$$

This expression shows the probability of being screened in to be positively correlated with the pre-bureau score and with the variance in performance, although X and u are orthogonal in the population as a whole.[8]

Borrowers differentially screened *out* using bureau information are those for whom $\bar{\alpha} < -X\beta \le 0$ (note that for current borrowers the vector X contains a history of inside lending information, which it does not for those newly screened in). The probability that a current borrower be ejected as a result of the use of the credit bureau is:

$$\Pr\left(X\beta + \bar{\alpha} < 0 \mid X\beta \ge 0\right) = \Phi\left(\frac{-X\beta}{\sigma_{\bar{\alpha}}}\right).$$

This is a decreasing function of the normalized score prior to the use of the credit bureau, meaning it is the borrowers closest to being unprofitable and those with the largest variance

---

[8] In the empirical section we will allow the variance in performance to vary with X.

that are most likely to be ejected with the use of the credit bureau. This gives the first set of claims that we will take to the data:

Claim 1. When the credit bureau is consulted, and at given pre-bureau score, the decision to eject or select a client is directly related to the quality of the repayment performance revealed by the bureau.

Claim 2. The probability that an applicant would be newly selected in (out) with use of the credit bureau is an increasing (decreasing) function of her pre-bureau score. The probability of being newly selected in *or* out is increasing in the variance in performance among applicants who are observationally similar.

Reduction of the prediction error.

Denote by ω the prediction error on the true profitability. From the expressions above we can readily derive the following:

$$\text{var}(\omega|X) = \sigma_g^2 + \sigma_u^2 + \sigma_\varepsilon^2 > \text{var}(\omega|X, \bar{\alpha}(N \to \infty)) = \sigma_u^2 + \sigma_\varepsilon^2$$
$$> \text{var}(\omega|X, \bar{\alpha}(N=1)) = \sigma_\varepsilon^2.$$

In other words, even in very large groups a bureau is more useful than no bureau because it allows lenders to remove the variance that comes from group-average quality. However, as the group size decreases towards one it also allows lenders to remove the variance that comes from deviation from group average quality, and hence the bureau is increasingly useful as group size decreases.

Reduction of the inclusion error

A client will be mistakenly given a loan if the prediction error ω is so strongly negative that it overwhelms an otherwise positive expected profitability $E\left(\pi^L\right)$. We can thus calculate the probability of mistakenly including a bad client (or mistakenly ejecting a good one) with our different benchmark levels of information:

Mistakenly giving a loan to a bad client: $\Pr\left(X\beta + a < 0 \middle| E\left(\pi^L\right) \geq 0\right) = \Phi\left(-E\left(\pi^L\right)/\sigma^*\right)$,

Not giving a loan to a good client $\Pr\left(X\beta + a \geq 0 \middle| E\left(\pi^L\right) < 0\right) = \Phi\left(E\left(\pi^L\right)/\sigma^*\right)$,

where: $E\left(\pi^L\right) = X\beta$ and $\sigma^* = \sigma_a$ in absence of bureau,

and $\quad E\left(\pi^L\right)= X\beta +\bar{\alpha}$ and $\sigma^* = \sqrt{\dfrac{N-1}{N}\sigma_u^2 + \sigma_\varepsilon^2}$ with group bureau information.

Because the variance of the prediction error is higher when there are more sources of uncertainty, the likelihood of making a mistake without a bureau is higher for any predicted level of profit. Similarly, it is more likely that the lender makes a mistake with a bureau reporting on large groups than one reporting on individuals.

Claim 3: The probability of erroneously selecting a bad client or rejecting a good client decreases with use of the credit bureau. The quality of selection is higher for borrowers from smaller groups.

Without any observations on the quality of rejected applicants, we cannot verify the theoretical results that the credit bureau would reduce the rejection of potentially good clients, but we will verify the reduction in enrollment of bad clients.

Decision to use the bureau

In order to understand the influence of group size on the decision to use the bureau, we must model how lenders form expectations over the group sizes that they will observe in the bureau for a given borrower. Perhaps the simplest way to think about this is that the lender observes the group size of each borrower on *inside* loans (or loan application for candidates), and expects that each borrower takes *outside* loans from groups of the same size. Denoting this quantity by $\hat{N}$, the variance of the prediction error of the signal observed in the bureau will then be $\sigma_u^2 \dfrac{\hat{N}-1}{\hat{N}} + \sigma_\varepsilon^2$. Since the variance of the prediction error without the bureau is $\sigma_g^2 + \sigma_u^2 + \sigma_\varepsilon^2$, the benefit from using the bureau comes from the expected decrease in prediction error. Prior to consulting the bureau, expectation on profitability is based on currently available information $\hat{E} = E\left(\pi^L\middle|X\right)$.

Imagine that checking any individual in the bureau entails a fixed cost F. If the cost of making a loan to a borrower who turns out to be bad is c^b, then a lender will only use the bureau if the expected decrease in inclusion error costs outweighs the fixed cost of the bureau, or if:

11

$$c^b \Delta\Phi > F, \text{ where } \Delta\Phi = \Phi\left(\frac{-E\left(\pi^L \mid X\right)}{\sqrt{\sigma_g^2 + \sigma_u^2 + \sigma_\varepsilon^2}} \right) - \Phi\left(\frac{-E\left(\pi^L \mid X\right)}{\sqrt{\frac{\hat{N}-1}{\hat{N}}\sigma_u^2 + \sigma_\varepsilon^2}} \right),$$

conditional on $E\left(\pi^L \mid X\right) \geq 0$. This is increasingly unlikely as the pre-bureau score increases. As the score includes all observable behavior for current clients, this suggests that the lender is more likely to seek the bureau's information for poorly performing clients.

Similarly, if the costs of mistakenly rejecting a client who would have turned out to be good are the foregone profits c^g, those who appear to be bad borrowers without the bureau will only be checked if $c^g \Delta\Phi > F$, conditional on $E\left(\pi^L \mid X\right) < 0$. This is increasingly unlikely as the pre-bureau score decreases.

Hence use of the bureau is more likely if (i) the pre-bureau score is close to break-even; (ii) costs from mistakes are high, or the fixed cost of using the bureau is low; (iii) there is strong assortative matching over group quality, so that σ_g^2 is large compared to σ_u^2 and σ_ε^2; and (iv) an individual takes loans from a small *inside* group. While we cannot identify variation in the second and third dimensions of heterogeneity, we will confirm with the data the other two results.

Claim 4: Use of the credit bureau is more valuable (and hence the lender is more likely to seek its information) for clients taking loans individually or in small groups, and for current clients with low payment performance.

Claim 5: If the cost of carrying a bad client is high compared to that of rejecting a good client, one expects to see use of the credit bureau to induce large-scale screening out among current clients, despite the direct information the lender already has on its clients.

3.2. CONTRACTS FOR ONGOING BORROWERS

We now turn to use of the bureau for on-going clients, and what it reveals on the remaining asymmetric information between the lender and its clients. Ongoing borrowers are those for whom $E\left(\pi^L \mid X\right) \geq 0$ and $E\left(\pi^L \mid X, \bar{\alpha}\right) \geq 0$. Our theory suggests three different reasons why we may observe changes in the contracts offered to ongoing borrowers:

1. An ongoing borrower who is checked and continues to receive a loan must have been found to be of high quality, or to have little outside debt. The natural response is to increase the loan size to this borrower, which in turn is likely to result in higher default for that borrower than would have obtained if the loan size had remained at the pre-credit check level (McIntosh & Wydick, 2007).

2. A positive economic shock will likely improve repayment to all lenders. Therefore individuals who are checked and found to be of high quality in one period could have received above-average shocks in the period in which they were checked. If business shocks are i.i.d. over time, mean reversion will cause an increase in default among ongoing clients immediately after the use of the bureau.

3. The change in contract offers may shift the demand side of the market as well; changes in collateral requirements or interest rates may alter credit demand from a single lender, or may alter the preference ordering across lenders to which a borrower has access. This represents a general equilibrium effect of the revelation of $\bar{\alpha}$ on borrowers, even though they are unaware of *why* the contract offer has changed.

Claim 6: Borrowers that are kept as clients after consultation of the credit bureau are more likely to experience an increase in loan size and a deterioration in repayment performance than those that would remain clients without consultation of the credit bureau.

3.3. IMPACTS FROM BORROWER TRAINING

With the rollout complete and no revelation of existence of the bureau to borrowers, we have a 'lender only' bureau. The lender-driven adverse selection effects have by then taken place, and borrowers have responded to the alteration in credit contracts, but they have not incorporated the *incentives* created by the existence of the bureau into their decision making. By training clients at this point, we interact only with those borrowers who are deemed profitable with the bureau in use, and so we do not observe as large an effect as we would have had we trained borrowers prior to implementation of the bureau (when 'bad' borrowers would have had a chance to improve behavior before it was observed).

Moral Hazard Effects.

We conducted the trainings in the middle of a group's lending cycle, and so can look for discontinuous impacts of the training on groups of fixed composition. This allows us to cleanly isolate the intensive margin impacts of the training, which should arise purely as a result of decreased moral hazard. Such moral hazard might arise from risky project selection (Ghatak & Guinnane, 1999), shirking in project implementation (Conning, 1999), or from hidden outside debt in the presence of incomplete contracting (McIntosh & Wydick, 2005).

The fact that the bureau is reporting on the group signal $\bar{\alpha}$ indicates that the strength of the incentive to reduce moral hazard effects is a function of *N*, the group size. Group average reporting creates a classical free rider problem in which the benefits of risky behavior are captured by the individual, whereas the costs are distributed among members of the group. We therefore expect that the *restraint* effects of the training, whereby inside default will decrease because of strengthened repayment incentives, will be larger in Solidarity Groups than in Communal Banks. Conversely, the *new opportunities* effects, whereby the existence of the bureau expands the use of outside credit for good borrowers at the cost of higher default, should be largest for those in Communal Banks.

Claim 7: The power of the credit bureau in curtailing moral hazard behavior is stronger for smaller groups. The incentive for borrowers to reduce moral hazard is larger for those more likely to seek loan opportunities outside the lender.

Group Borrower Selection

Although group members are not acquiring any new information on the other members, learning that of the existence of a credit bureau can still lead to changes in group-driven selection. To understand the logic of this selection, we can formalize a group's decision as follows. Group profits can be written as a function of the mean characteristics of group members, $\pi^B(\bar{X}, \bar{a})$.[9] Linearizing this system, we can write this expected group profit as $\pi^B = \bar{X}\beta^0 + \bar{a}\gamma^0$. We assume that borrowers have access to some outside credit source with profit π^*, and that prior to the training borrowers believe that their outside options are unrelated to their group performance in Genesis. Because the training informs borrowers

---

[9] We follow the convention in the joint liability literature in assuming that borrowers within group loans are perfectly informed as to each others' characteristics and actions, and hence the training alters only incentives and not information sets for borrowers.

that the inside lender has already been using the bureau for a year or more, it has no immediate effect on borrower profits from inside loans. The training's sole effect is to make borrowers aware that outside profits have become a function of inside group performance, which we linearize as $\pi^* = \overline{X}\beta^1 + \overline{a}\gamma^1$.

To the extent that $\beta^0 = \beta^1$ and $\gamma^0 = \gamma^1$, this knowledge induces no changes in selection on inside loans because the determinants of profitability are the same. Writing the elements of the vector β as $b_1,...,b_j,...,b_J$, we expect to see any characteristic for which $b_j^1 > b_j^0$ positively selected in Genesis subsequent to the training, and negative selection for attributes with $b_j^1 < b_j^0$. Perhaps the most concrete example of this is that the bureau reports heavily on missed *intermediate* payments[10], whereas access to future loans in Genesis is based almost exclusively on *final* repayment performance. Hence we expect to see that the type of borrowers who miss intermediate but not final payments will be pushed out after the trainings.

Claim 8: Group selection is expected to result in ejection of members whose repayment behavior has relatively more weight in the credit bureau information than with the lender.

4. DATA

The main source of information for most of the analysis is the administrative data from Genesis, reporting all loans extended in its 39 branches from February 1999 to July 2005. This database includes some minimum information on all clients recorded at the time of their first loan (gender, age, education, marital status, and urban/rural residence), information on all loans (amount, term, periodicity of repayments, and the balance if any in July 2005), and on all transactions on each loan account (payments on principal, on interest, and on fines, and balance). For group loans, information is given on the identity of the members and the individual amounts they received, but repayment information is solely reported at the group level. From the detailed accounts on payments, we constructed two indicators of loan performance: the total fines paid as percentage of the principal, and

---

[10] Crediref reports include a line for each loan that categorizes each payment into 'on time' and a variety of delinquency categories. While there was some confusion among credit officers as to how to read this string, we found they had no difficulty in differentiating a loan that had had intermediate problems from one that had not.

whether the final payment was settled more than 2 months after the original due date. Information is also available on the dates at which the credit bureau was consulted for information on each client or applicant, although no outcome is recorded.

Table 1 presents basic summary statistics that give a sense of the differences between borrowers and contracts across the three loans types offered by Genesis during the six months prior to use of the credit bureau. Communal banks are very different from individual clients and solidarity groups. They are larger groups (on average 7.7 members taking loans vs. 4.5 in SGs) composed mainly of women (99.7% vs. 46.7% in SGs), with low education (41.4% uneducated vs. 26.5% in SGs), and with much smaller loans (US$324 per client vs. $662 in SGs).

We also had access to the information recorded in the credit bureau on all Genesis clients, as of October 2005. This includes 263,116 loans. For each loan, the reported information includes whether the loan is an individual loan or a group loan, its total amount (but not the number of members if group loan, which prevents us from inferring a potential level of indebtedness), and a series of codes referring to up to 20 periods indicating whether the repayment status was in good standing, in arrear, or if the payment was the last payment. Crediref collects information from its members on a monthly basis, and only the date of the last entry of information into Crediref is recorded. The earlier date is June 2002, while 40% were still active loans and thus have information recorded in October 2005. From these monthly repayment status, we created an indicator variable for whether the loan had ever been in arrears.

Matching clients from Genesis and Crediref databases was done from a common identification code. Recognizing the Genesis loans in the Crediref database was done from the loan amount and from an approximate correspondence between the last dates recorded for payment in Genesis and payment status in Crediref.

Use of the bureau to screen clients is not universal. The fixed cost of a check in Crediref starts at $1.60 (decreasing to $0.67 per check when over 6,400 checks per month are made), thus making Genesis less likely to investigate small borrowers. Almost no CB clients have their records checked. This is consistent with our theory indicating that the informational gains from the bureau are decreasing in the expected group size, and hence we ignore the CB borrowers throughout the analysis of the impact of lender use of the bureau.

5. Empirics: The Lender Begins Using the Bureau

Table 2 provides some simple tests of exogeneity of the rollout of the bureau. Let f_{lbt} be an outcome, such as loan delinquency, late fees, and group size, for loan l issued by branch b in month t. We first calculate borrower-level changes in outcomes from the previous loan, Δf_{lbt}, for all individual borrowers and Solidarity Groups. We then calculate branch/month averages of these changes in outcomes, separately for individual borrowers and Solidarity Groups, for 1, 2, 3, and 4-6 months prior to entry in the credit bureau in July 2001, $\overline{\Delta f}_{bt}$, and regress them on a set of month dummies δ_t and a variable T_b which gives the numerical month in which the bureau was introduced in branch b:

(1) $\qquad \overline{\Delta f}_{bt} = \delta_t + \beta T_b + u_{bt}$.

This tests whether outcomes were changing at a different rate in branches that received the bureau earlier relative to those that received it later, which is the identifying assumption of an impact regression using branch- or borrower-level fixed effects. Results reported in Panel A of Table 2 show no indication of any differences in rates of change in loan performance (loan delinquency and late fees) or number of members in solidarity groups with the timing of branch entry in the credit bureau. These regressions were also tried without month fixed effects and the results were similar.

We now look more closely at the repayment behavior over the short period prior to entry in the credit bureau. Figure 1 shows the observed branch level performance before and after its entry in the credit bureau. The pattern for individual loans is quite clear with a net discontinuity concomitant with use of the credit bureau. There is no particular pattern in the few months prior to entry, which would create concern with a potential endogenous sequence in the rollout, either in response to repayment problems (an Ashenfelter dip), or following an on going improvement in performance. For solidarity groups, the picture is less powerful. The pattern is dominated by an overall increase in default among solidarity groups that occurred between May 2000 and January 2002, but there is no evidence of a short term pre-entry burst in default, nor of a break in the trend just prior to entry in the credit bureau. These results are confirmed by regression of the average branch outcome $\overline{f}_{bt}$ on a set of month dummies δ_t and variables m_{-n} that denote the month n before entry in the credit bureau, estimated over the period prior to entry in the credit bureau:

$$(2) \qquad \overline{f}_{bt} = \delta_t + \beta_1 m_{-1,bt} + \beta_2 m_{-2,bt} + \beta_3 m_{-3,bt} + \beta_4 m_{-4,bt} + u_{bt}.$$

Results reported in panel B of Table 2 show no specific patterns, positive or negative, in the months preceding entry of a branch in the credit bureau.

We thus proceed to utilize the staggered entry of Genesis branches into the bureau to estimate causal impacts on credit outcomes for individual borrowers and Solidarity Group members. Because we observe whether each loan is issued to a new or to an ongoing borrower, we can disentangle the screening effects of the bureau on the extensive margin from changes in contracts on the intensive margin. We can also track the differences over time between borrowers who entered Genesis before and after the bureau was being used, and so measure the longer-term effects of improved information.

5.1. SELECTION OF CLIENTS USING THE BUREAU

We run a regression at the loan level to explain the difference in differences that results from the staggered treatment dummy R_{lbt} on the loan performance f_{lbt}. R_{lbt} is one for loans to clients that could have been checked by the bureau and zero for loans given before the bureau was introduced in each branch, and hence measures the Intention to Treat Effect of a branch *having* the bureau, not the effect of *using* it.[11] Our basic regression specification is then:

$$(3) \qquad f_{lbt} = \delta_t + \delta_b + \beta R_{lbt} + u_{lbt},$$

where δ_t and δ_b are month- and branch-specific fixed effects, and u_{lbt} is an error term clustered at the branch level.

Table 3 demonstrates the substantial increase in the fraction of clients selected in and out under use of the bureau. The sample includes all loans made by Genesis between February 1999 and May 2004. In the top panel, we look at clients leaving. The dependent variable f is equal to 1 for loans whose taker does not take a loan from Genesis afterwards. The subscript t in equation (3) refers to the month in which the last payment is made. The treatment of being screened is a dummy variable equal to 1 for loans whose last repayment was made after the credit bureau became available in a branch or, to account for typical lapse

---

[11] We measure the ITE because the decision to check an individual is non-random, and so estimating the Treatment-on-the-Treated would require that we try to predict who in the branches without the bureau would have been checked. Any unobservable in this equation would create bias.

between two loans, at most 2 months before that date. In expectation, the *first* use of the bureau for a given individual should reveal more new information than subsequent checks on the same individual. We therefore distinguish two treatment variables. The "first-time screened" treatment applies to the first time a client would have been screened with the bureau, i.e., to the last loan granted before the availability of the bureau and with last payment made no earlier than 2 months before the availability of the credit bureau, while the "subsequent screenings" applies to all subsequent treated loans.[12] In the lower panel, the variable of interest *f* is whether the loan is granted to an entering client. Time *t* refers to the month in which the loan is granted, and the treatment variable is a dummy variable equal to 1 if the credit bureau is available to the branch in this month. We see that the bureau causes a symmetric change in the percentage of individual borrowers who leave and who enter (column 1); both figures increase by roughly 14 percentage points (beyond base rates in which 42% of the clients are new and 51% leave across cycles prior to the bureau). In other words, there is a period of upheaval in the client base triggered by the use of the bureau. This drastic change in the rate of ejection of clients suggests the value of the credit bureau even for clients on whom the institution has rich information from a history of lending, as mentioned in claim 5. Figure 2 shows the large increase in new individual clients that occurs for roughly six months after the bureau is implemented. For solidarity groups, the picture is somewhat more nuanced; individuals within these groups are much more likely to be expelled, by 19 percentage points (column 2), but the groups themselves become more durable as a result of the bureau (column 3). The net effect of decreased enrolment in existing groups, expulsions, and the creation of new smaller groups is the dramatic decrease in average Solidarity Group size, illustrated in Figure 3, falling from 4.3 to 2.7 members. In other words, the bureau causes the lender to rely less on joint liability as a screening tool.

Another illuminating comparison is looking at the Crediref records of ongoing borrowers who were checked and ejected from Genesis versus those who were checked and not ejected. Table 4 draws this comparison. Ongoing borrowers are defined as those that have had a loan with Genesis until at most 2 months prior to the introduction of Crediref in their branch. Clients are sorted according to whether their Crediref record was checked or not in the two months that follow availability of the credit bureau. Ejection is here again

---

[12] This includes borrowers that were clients before the credit bureau that are checked later and clients selected after the credit bureau.

defined as the client not taking subsequent loans with Genesis.[13] As can be seen from the number of observations in each group, 59% of individuals and 63% of SG members are checked, and of these 12% and 9% are ejected, respectively. While the decision to screen is endogenous, and so no causal interpretation can be drawn from this comparison, it is clear that Genesis is more likely to check the records of individual clients with the worst performance. Among these clients, those who leave are about twice as likely (36.7% vs. 20.8%) to have defaulted on external loans than those who remain clients. Although this looks like a substantial difference, it is small compared to the ten-fold difference seen between first time candidates who are rejected and those who are taken as clients.[14] Because past performance is part of the information set for prior clients, we expect the influence of the information in Crediref to be more muted for ongoing than for prospective borrowers. For Solidarity Group members, although the information in Crediref signals those who will be ejected, the internal information also conveys differences. The ejected clients come from groups with worse performance (33.7% of loans with arrears) than the non-ejected (16.8%). Note that for all clients, the performance recorded in Crediref can be from both individual and group loans. The role of Crediref is thus particularly clear in helping discriminate among individual clients with mediocre internal performance. Together these results show an extensive use of the credit bureau, even for clients on whom the lender has direct information (claim 5), particularly for clients with repayment problems (claim 4). The information revealed by the credit bureau informed the lender's decision on whether to eject a client or not (claim 1).

Table 5 examines how the use of the bureau altered the characteristics of clients selected in and out of Genesis, and so gives us a direct test of the theory outlined in Section 3.1 and claim 2. Our theoretical model indicates that what is learned through the bureau should be orthogonal to the pre-bureau score, but that nevertheless selection in (out) will pick borrowers who are better (worse) on observables. It also indicates that the likelihood of selection in and out will be increasing in the variance of repayment performance conditional

---

[13] It thus also includes the clients that decide on their own not to renew their demand for a loan. Note also that some of the non-checked clients may have left on their own or be ejected by Genesis without checking their records. These confounding factors however all play in favor of blurring the contrast that we are making in the table between the internal performance of the checked and unchecked borrowers, as well as the external performance of the ejected and not ejected.

[14] Genesis records give the identity of all applicants and a date at which their Crediref record was checked. However no further information is available on the applicants that are not accepted. We have thus no other observables to compare accepted and rejected applicants apart from their Crediref records.

on the score. In order to investigate these hypotheses, the first column of Table 5 gives coefficients from a pre-bureau scoring model, and the second column reports how the variance of outcomes relates to these same characteristics. We only use individual borrowers for this analysis because we can relate their own performance to their characteristics without confounding the influence of other group members. These are OLS estimations of performance indicators on client attributes and time and branch fixed effects for the 13,550 loans granted before Crediref, from February 1999 to July 2001. The low performance indicator is the ratio of late fines to principal paid over the whole loan cycle. The variance indicator is the square of the residual in the performance indicator estimation. This is meant to mimic the scoring model that Genesis could be using, although it is only based on the subset of characteristics that are recorded in the Genesis information system to which we have access, rather than the complete information that credit agents have from their personal contacts with clients. Results show that lower performance and larger variance in performance are associated with male, some secondary education (for low performance), and lower age. None of the other covariates (marital status, ethnicity, some primary school education as opposed to none, and rural/urban location) were significant in this simple scoring model.[15]

The third and fourth columns of Table 5 report on separate regressions that measure changes in the composition of entering and departing borrowers,

$$(4) \qquad X_{ibt} = \delta_t + \delta_b + \beta R_{ibt} + \gamma X_{ibt}^* + u_{ibt}.$$

X_{ibt}, representing either gender, education, or age of the new entrant or the departing client, is regressed on the R_{ibt} rollout dummy, controlling for the other two characteristics X^*. So the coefficients β should be interpreted as the effect of the bureau on the average characteristic of new and departing clients, conditional on the other two characteristics. Results show that the bureau pushes individual lending to be less male (more dropouts) and less educated (both less new clients and more dropouts), but has no effect on the age composition. These regressions were performed on all new clients until May 2004, but only on departing clients that have been selected before the credit bureau, to avoid confounding

---

potential selection effects. Very similar results however are found when including all departing clients, and without conditioning on the other two characteristics. This confirms the theoretical result that the bureau causes large selection effects in client categories that have low pre-bureau *average* and large *variance* in performance (claim 2).

5.2. EXTENSIVE MARGIN IMPACTS OF THE BUREAU

We now analyze the impact of the use of the bureau on the quality of the selected clients. A good client pays back, but also takes repeated and increasing loans. In fact a major problem of this and many similar microfinance institutions is the large turnover of clients that only take one or possibly two loans. Since the selection of a new client entails large fixed costs, stabilizing the clientele would bring large efficiency gains. By tracking the differences over time between borrowers who entered Genesis before and after the bureau was being used, we measure the longer-term effects of improved information.

In the first three columns of Table 6 we use loan-level data and the specification in (3) to measure the impact of the bureau on client repayment performance on their first loan. The sample includes the first loan of all new clients selected between February 1999 and May 2004. The time index refers to the month the loan was granted, and the treatment variable to whether this selection occurred after the availability of Crediref. For loans given to individual clients, where we would expect the effects of new information to be strongest, we see a 24% drop in the share of loans that were charged important late fees. Loans more than 2 months delinquent, which would be technically under default, are not changed. For group borrowers, on the other hand, the primary effect is an increase of 42% in mean loan size. So the decrease in the lenders' adverse selection causes loan sizes to go up, while repayment on these new loans does not deteriorate. Given that lenders' profits are made on margin, loan size increases of this magnitude with no corresponding increases in repayment problems translate into substantial earnings for the lender.

The last three columns compare performance on subsequent loans for those selected before vs. after the bureau, with an estimation similar to (3) at the individual client or SG level:

$$(5) \qquad f_{ibt} = \delta_t + \delta_b + \beta R_{ibt} + u_{ibt} .$$

The time variable refers to the month the client or SG was selected, and the treatment variable to the availability of the bureau at that selection time. In column (4), the dependent

variable is whether these clients or SGs have taken a second loan with Genesis or not. In columns (5) and (6) we characterize these second loans by the time elapsed since the end of the first loan, and the growth in loan size. The bureau causes significant improvements in the ability to pick good stable clients, but only for people who take individual loans, a result in support of claim 3. Individual borrowers selected with the bureau are 53% more likely than those selected before the bureau to go on to take subsequent loans. These subsequent loans are taken somewhat sooner, and the size of these loans is 12% larger. Group borrowers, on the other hand, display no impacts on subsequent performance. This is consistent with the joint liability mechanism providing a richer information set when group borrowers are screened without a credit bureau.

5.3. IMPACTS OF THE BUREAU ON ONGOING BORROWERS

Having seen the supply-side changes that occur on the extensive margin, Table 7 carries out the analysis on the intensive margin by including only borrowers who took loans both before and after the bureau was being used in their respective branch. Here we can include borrower-level fixed effects, and so the treatment effect measures changes in contracts for ongoing clients. Since we have limited the sample to those for whom $E\left(\pi^L\middle|X\right)>0$ and $E\left(\pi^L\middle|X,\bar{\alpha}\right)>0$, we follow a consistent cohort through the implementation of the bureau. We estimate:

(6) $f_{ilt} = \delta_t + \delta_i + \beta R_{ilt} + u_{ilt}$,

for everyone whose first loan was before the bureau and last loan was after the bureau, continuing to cluster standard errors at the branch level. The dependent variable is a characteristic of the loan and the time index refers to the month in which the loan was granted. The treatment variable indicates whether the loan was granted before or after the availability of the bureau. Note that this estimation measures an intention to treat effect since not all on-going clients were checked in the bureau, as seen in Table 4. This test provides a measure of the intensive margin effects, but it may be biased by mean reversion in the manner suggested in Section 3.1.

Loan sizes for ongoing individual clients and SGs increase by 7.8% and 23.9%, respectively, when the bureau is in place. However, for individual loans, this is accompanied by a sharp worsening of repayment (claim 6). Such a worsening is not surprising if loan sizes

have increased, but evidence that mean-reversion is involved is provided by the extremely low mean default rate among these ongoing clients: 2% versus an institutional average of over 4%. Thus the 'treatment effect' measured in the second column of Table 7 essentially shifts performance among borrowers retained when the bureau comes in back to the unconditional mean in the institution. Given the potential positive bias in the estimated parameter, we cannot conclude on the causal effect of the credit bureau on on-going borrowers, except for saying that if there were any effect in reducing bad repayment performance it does not dominate the selection bias effect.

5.4. AGGREGATE IMPACT OF THE BUREAU AT THE CREDIT OFFICER LEVEL

We conclude this section on the use of the credit bureau by the lender with some statistics on the aggregate effect of using the bureau on the institution's portfolio. The introduction of the credit bureau in branches was accompanied by installation of new hardware and software, and a substantial overhaul of the institution's information system. While this has improved the efficiency of the lender's administration, credit officers did not directly deal with any of these tasks. They continued to operate in the field using regular paper forms to collect information that was digitalized afterwards by administrative personnel. The only significant effect of the installation of the new computer system on the credit officer performance therefore came from availability of the credit bureau information. Credit officers received some training in reading the information provided by the bureau. While this may have raised awareness of the importance of repayment performance, there was no change in guidance for the traditional selection procedure that followed consultation of the credit bureau, nor in the incentives given credit officers to maintain good repayment in their portfolio. It is therefore unlikely that the measured changes in performance of borrowers and efficiency of loan officers came from factors other than introduction of the credit bureau.

We perform the analysis at the credit officer level, which allows us to control for effects due to changes in the number and composition of credit officers. An efficiency effect of the bureau is the ability for a given employee to increase the number of new borrowers whose applications he processes in a given period of time. Results reported in the first two columns of Table 8 use credit officer fixed effects in a regression on data at the credit officer/month level:

(7) $$f_{obt} = \delta_t + \delta_o + \beta R_{bt} + u_{obt},$$

where the dependent variable is a characteristic of the portfolio of credit officer o of branch b in month t, and the treatment variable is whether the credit bureau was available to the credit officer in that month. Results show an increase by 64% in the number of new individual clients and 149% in the number of members of new SGs, suggesting large efficiency gains in the selection of clients. Total number of loans per month increases by 27%, and aggregate performance also improves, with a large 33% (but statistically not significant) reduction in the share of loans with large fines. This translates into a large 52% increase in branch level loans and amount lent without deterioration of the portfolio at the branch level (columns 3 and 4).

A back of the envelope calculation of costs and benefits further indicates the financial return to the lender of using the bureau. Investment costs are $9,250 per branch in hardware, telephone, and furniture, plus $1,280 to the institution in software. This sums up to $9,280 per branch. These assets were expected to need replacement or updating on average every 3 years, which we take as benchmark. Variable costs are estimated at $72 per branch per month. The main benefit stems for the increase in lending by $24,000 per month per branch, caused by use of the credit bureau by the institution (Table 8). The majority of loans are for 6 or 12 months, averaging 10 months. The cost of capital is 17% and the lending rate is 25%. Since this additional lending was undertaken with no increase in the number of credit officers, the benefit associated with it is likely a very large share of the 6.6% margin made on 10 months loans. Assuming 3% benefits would lead to a monthly benefit of $725. Net benefit would thus be $653 per month, paying back the fixed investment in 14 months, much before the anticipated 3 years before replacement.

6. EMPIRICS: BORROWERS LEARN THAT THE LENDER IS USING THE BUREAU

We now turn to the impacts of the randomized credit bureau training course that was implemented 18 months after the staggered rollout was complete. The sample used in this analysis consists of seven branches randomly selected from the 39 Genesis branches to represent the variety of Genesis clients. Within each of these seven branches, we randomly selected a predetermined number of groups for treatment, the others forming the control

groups. [16] Observations are thus weighted in the analysis, in correspondence to the sampling scheme.

Once selected, groups were notified that they were eligible to receive a free information session, and they were requested by their credit officer to appear at a specific time and place in order to receive the information. Attendance was entirely voluntary, and if a group did not show up the first time, two subsequent efforts were made to call it for the session. The percentage of chosen units that were in fact treated varies from 31% to 100% across branches, with an average response rate of 62%. The information sessions took place over a period of four months, from July to November 2004, with the order in which groups were called randomly defined. The timing of the treatment is thus specific to each treated group and we assign the median of the treatment dates within each branch to the control groups. The analysis that follows is based on loans granted from February 2003 on, when all 7 branches had access to the credit bureau, and with due date by June 2005. This includes 2582 loans taken by a total of 1040 groups, 573 selected for training and 467 control groups.

The quality of the randomization can be gauged from Table 9, panel A. Comparing the mean values of group-average characteristics such as age, marital status, education, gender, and ethnicity for the loans extended to the 994 groups that were active during the 17 months preceding the randomization, from February 2003 to June 2004, we find no evidence of significant differences between the selected and control groups. On loan performance, the situation is less ideal. The selected groups perform somewhat better than the control groups, even though, because of the large variation within each category, the difference is statistically not significant.

Because we use branch and month fixed effects in all the regressions, the relevant question in terms of bias to our estimators is whether a significant difference remains when controlling for these fixed effects. By constructing false treatment effects within the pre-treatment time period, we verified in Table 9, panel B, that none of the false treatment effects are significant, suggesting that there are no serious biases when we use a double difference.

---

[16] This selection was done by randomly selecting one branch in each of seven groups of similar branches constituted by credit officers with intimate knowledge of the institution. However, despite the randomization, the average characteristics of the groups from these selected branches do not perfectly match those of the non-selected branches. We therefore limit the analysis to the groups from the selected branches. Within each branch the percentage of groups that was selected for training was also randomly assigned values from 25% to 87% with the objective of looking at spillover effects. We however found no such differential effect and thus ignore this issue in this paper.

Because of the relatively high non-response rate, our analysis focuses on the Intention to Treat Effect. Assuming that the non-experimental implementation of the training program would have a similar non-response rate, the ITE is in fact the quantity of interest for the lender.

The interpretation of the ITE is complicated by the fact that some borrowers must certainly have understood at least some of the implications of the bureau before the training. As a general matter, knowledge of the workings or indeed the existence of Crediref was very low among clients; not one of 184 clients surveyed in the summer of 2003 was aware that information was being shared between microfinance lenders. That said, certainly some clients would have possessed better information, or at least more realistic expectations, over the process of information sharing. Such clients will appear to have a lower impact (and hence a smaller moral hazard response) simply because they learned less from the sessions. A causal impact of the treatment, then, is composite of the amount that was actually learned and how what was learned effects behavior.

6.1. THE INTENSIVE MARGIN: IMPACT WITHIN A LOAN CYCLE

Since group composition is fixed within a loan cycle, the instantaneous impact of the information program on inside repayment isolates the moral hazard effect that arises from the desire to use reputation to leverage credit from other sources. Hence in the short run our experiment raises the value that clients place on good performance to access outside credit. Over time, the repercussions of changes in group membership undertaken due to learning about the bureau begin to have their own effects upon inside repayment, adding adverse selection to moral hazard effects.

The trainings occurred mid-loan for 735 of the groups, and so we have the ability to see whether a given group of people change their behavior once the existence of the bureau is revealed. This analysis is conducted at the loan *payment* level, separately for Solidarity Groups and Communal Banks. The observations are the different intermediate payments made on the loans that were active at the time of the treatment. Because repayment problems tend to come only after a certain time is elapsed, we control for where in the loan cycle the repayment takes place. A complication occurs in that loans are of different lengths and require various numbers of intermediate repayments. To make these repayments comparable, we therefore divide the length of each loan cycle in 10 equal time intervals,

which we refer to as deciles, and we control for the deciles rather than the rank of the repayment. We thus estimate:

$$(8) \qquad f_{plt} = \delta_l + \delta_t + \sum_d \delta_d D^d_{plt} + \beta T_{plt} + u_{plt},$$

where f_{plt} is an indicator of performance for payment p made at time t on loan l that was active at time of treatment. The decile dummy variable D^d_{plt} is equal to 1 if the payment belongs to decile d. The treatment variable, defined at the payment level, T_{plt} is set equal to 1 if the payment p is in loan l taken by a group g that was selected for treatment and $t > \tau_g$, the treatment date for group g. Results are reported in the first column of Table 10.

Performance on the final payment and overall loan performance can only be analyzed by simple difference between treatment and control groups within a branch. This is done with a loan level estimation:

$$(9) \qquad f_{lb} = \delta_b + \beta T_{lb} + u_{lb},$$

where the treatment variable indicates if the group was selected for treatment or not. Results are reported in columns 2 and 3 of Table 10.

We see in column 1 of Table 10 that there was no significant change in performance on intermediate payments as a result of the training. The fraction of loans which were late as of their final payment, however, decreases by 10 percentage point (over a base value of 20 percent) with the training, but only for SGs (column 2). Changes in the fraction that ultimately pays large fines or goes into default are not statistically significant, but show large point estimates of -7 percentage points (over averages of 14 and 19%, respectively) for SGs and -3 percentage points (over averages of 3.5 and 4.5%) for CBs, suggesting important improvement in payments after the information session, particularly for the first (column 3). This is consistent with claim 7 that the power of the credit bureau in curtailing moral hazard is higher for smaller groups.

6.2. IMPACT ACROSS LOAN CYCLES

We have data on repayment behavior from Genesis for one year after the intervention. Over this intermediate time frame, we expect the moral hazard impacts to dominate although, in groups that take one or more loans after having received the information, repayment behavior is also plausibly being affected by the selection response of

group members. These impacts are measured by estimating the repayment performance at the loan level over the long period February 2003-June 2005 (excluding loans that were active at time of treatment). We use group fixed effects estimators:

$$(10) \qquad f_{lgt} = \delta_t + \delta_g + \beta T_{lgt} + u_{lgt},$$

where f_{lgt} is a measure of repayment performance of loan l of group g with last payment at time t. The treatment variable T_{lgt} equals 1 if the group g was selected for treatment and $t > \tau_g$, the treatment date for group g.

Results are reported in column 4 of Table 10. With only a few months of observations after the information courses, the evaluation of the repayment performance is based on very few post treatment date loans (181 SG loans and 23 CB loans), and none of the CB loans, from either treated or non-treated groups, bore a large fine or were in default. We however observe a significant decline of 6 percentage points in delinquency for SG loans for the treated groups relative to the non-treated groups.

6.3. THE EXTENSIVE MARGIN: IMPACT ON GROUP COMPOSITION

We argued in Section 3.2 that changes in group selection induced by the training will arise as a result of differences between β^0, the determinants of profitability on inside loans, and β^1, the determinants on outside loans (claim 8).

We therefore begin our analysis of selection after the training by asking what kinds of borrowers were repaying well before they were aware of the use of the bureau. We estimate the correlations between group repayment performance and the average characteristics of groups in the seven branches that were in the training experiment during the period February 2003 to the treatment date,

$$(11) \qquad f_{lgt} = \delta_t + \bar{X}_{lg}\beta^0 + u_{lgt},$$

where $\bar{X}_{lg}$ is a vector of group-average attributes for each loan. We use only SG borrowers because the group-level characteristics are not consistently entered for CB clients. The purpose of this regression is for us (the econometrician) to observe these correlations as suggestive information on which types of members are likely to be worse performers, because we have no information on individual behavior. In contrast, group members observe each others' individual behaviors and thus act on the basis of perfect information.

Perhaps the most interesting correlate of repayment performance seen in the first two columns of Table 11 concerns the share of female borrowers in the group: women are more likely to have small repayment problems which result in fines (with a t-statistic of 1.6), but they are less likely to have loans more than 2 months delinquent (or substantial late fees above 1% of the principal, not reported). This may be related to the fact that the large majority of these women are primary care-givers in the household, and so as business people they must multitask in a manner that causes more frequent small lapses. We also see that groups with more educated members are worse payers.

With this basic sense of the correlates of repayment performance, we move to an analysis of the training's impact on group composition. We use the average characteristics of those entering and those leaving existing solidarity groups in the aftermath of the training. New clients are defined as members that join a group after the first loan of the group, and dropouts as those that quit the group before the last group loan. Members that do not participate in a particular loan cycle but return to the group for a subsequent loan are counted as continuing borrowers. As there are not enough observations on groups with entry and departure of members to use a group fixed effect estimator, we use branch fixed effects:

(12) $\quad \bar{X}_{lgbt} = \delta_t + \delta_b + \beta T_{lgt} + u_{lgbt}$.

Observations are weighted by the number of people represented by each group average.

Results are reported in the last two columns of Table 11. As we might expect in an environment where borrowers were fully informed about each other's quality prior to the training, there are no sharp compositional impacts of the training. And with a small number of observations, standard errors are large. But again the most intriguing result is on gender: SGs swing toward a more male composition after the training, with women more likely to dropout (17 percentage points increase in women share among dropouts) and less likely (but not significantly) to be accepted. Given that we expect the primary reason for differential selection after the training to be characteristics which are more important to outside lenders than inside lenders, it appears that the greater propensity of women to miss intermediate payments, combined with the detailed reporting of arrears on intermediate payments in Crediref, has decreased female access to SG credit. The difference between the coefficients on leaving and entering clients for women is 0.42, suggesting that the turnover of a member is 42 percentage points less likely to be a woman than it would have been without the training.

6.4. BEHAVIOR WITH OTHER CREDIREF LENDERS

A key assumption for a credit bureau to create incentives on client behavior is that it allows good borrowers to use the reputation they acquired with a particular lender for gaining access to outside borrowing opportunities. We therefore conclude this section on the effect of informing clients on the existence of the credit bureau with some support that this mechanism is at play.

Because we have access to data from the bureau itself for Genesis clients, we can test directly for the ways in which improved understanding of the credit system altered their behavior with *other* lenders. The entry of Genesis client data into the bureau took place along with the staggered rollout of use of the bureau, and hence as of the time of the training all 'inside' information had been observable to other lenders for a year or more. However, to the extent that clients did not know that the bureau existed, they possessed a form of reputational capital of which they were unaware. Treatment effects of the training thus arise from differences in the process by which borrowers *seek* outside loans, not from a systematic shift in the *supply* of credit.

We characterize a client's outside borrowing by the number of loans taken from other lenders (the quantity is difficult to interpret for group loans) and by the repayment performance of each loan as to whether there has been any late payment. The analysis uses only Genesis clients who were members of a group at time of treatment, and their treatment status is that of the group to which they belonged.[17]

The date recorded for each loan in Crediref is that of the last data entry, which corresponds to the closing date of the loan (except for the current loans which have their last transaction recorded in June 2005). In this analysis we consider as pre-treatment all loans completed before the treatment date. Post-treatment loans are those for which payments were recorded after the training date. We estimate the following equation at the client level:

(13) $\Delta_{igb} = \delta_b + \beta S_g + u_{igb}$,

---

[17] When clients belonged to two groups, they were considered treated if at least one of their groups was treated. About 3% of the control SG clients (20% of the control CB) changed group, joining a treated group after the treatment date. We also perform the analysis by attributing them the status of treated starting from the date they joined the treated group. Results are very similar and not reported here.

where Δ_{igb} characterizes the change in outside loans reported in Crediref from the pre-treatment to the post-treatment period of individual i from group g in branch b, δ_b represents the branch level change in outside borrowing for the members of the control groups, and S_g is a dummy variable indicating that the group g was selected for treatment. The parameter β measures the ITE effect of the information sessions. Errors are clustered at the group level. We use as indicator of outside borrowing the change in the number of loans and whether individual i started taking outside loans after the treatment date. Some of the loans classified as treated may have started before the information sessions; we have no way to know. However there is no reason for the number of such loans to differ across treatment and control groups, and hence it is absorbed in the branch fixed effects.

The repayment performance on outside loans is analyzed with the following branch fixed effects regression at the loan level:

(14) $\qquad f_{libt} = \delta_b + \delta_t + \beta T_{lit} + u_{lit}$,

where f_{libt} is a measure of performance for loan l taken by individual i from branch b last recorded in Crediref at time t. The treatment variable T_{lit} is equal to 1 if individual i was member of a group selected at time of treatment and $t \geq \tau_i$, the treatment date.

Results are reported in Table 12. They show treated SG starting to take outside loans 27% more frequently than mean value in the control group, and CB members increasing the number of outside loans and starting to take outside loans by 77% and 33%, respectively. This impact represents a substantial increase in credit received. Note that the change in reported outside borrowing includes a direct reporting effect associated to the expanding coverage of the bureau, which probably accounts for most of the observed change in the control group. And there is weak evidence of a positive effect on performance for SG borrowers with less ever missing a payment.

Who among these members responded to the training by taking on outside loans and how did they fare? Table 12 reports the contrast in ITE for good and bad Genesis clients as well as for more and less experienced clients. Good clients are defined as never having had a late repayment with Genesis (meaning they are members of good groups for their group loans), while those named bad clients have had at least one late repayment. More experienced clients are those who had 4 or more loans with Genesis, of any type, individual

or group, before the information session, and less experienced clients 3 or less loans with Genesis. Results show that good CB clients respond to information about their public reputation by starting to take outside loans or increasing the number of loans taken outside while bad clients do not. Less experienced CB clients respond by starting to take outside loans more than their experienced counterparts. However, they show a sharp and significant deterioration in their outside payments. This negative effect on performance does not happen with SG borrowers, who to the contrary tend to improve their outside performance.

One possible explanation for this response by CB relates to the extensive use of dynamic incentives in microfinance. Loan sizes begin very small and grow slowly, holding clients credit-constrained and thus putting a high value on the future relationship with the lender. While we expect the bureau to reinforce dynamic incentives with the credit system as a whole, the rise in default among CB clients indicates that the bureau may have a perverse effect by expanding credit access to new borrowers too fast.

7. CONCLUSION

We utilized an unusual pair of experiments which allowed us to decompose the impacts of a new credit bureau into two parts: what happens when a lender observes new information about borrowers, and what happens when borrowers become aware that lenders can observe this new information. We find that the new information in the hands of the lender has stark impacts. There is a large increase in the turnover of the client base, particularly in the 6 months after introduction of the bureau to a branch. Large numbers of individual and SG clients are ejected, and the categories of clients ejected are indeed those with lower average repayment performance (males, more educated borrowers) and larger variance in performance (males). There is a dramatic improvement in the repayment performance of new individual clients and in the size of the loans made to new SG borrowers. The new individual borrowers are better future clients in that they are more likely to take subsequent loans and have faster growth in their borrowing. For solidarity group borrowers, the bureau induces a strong swing toward smaller groups and allows the lender to increase loan sizes without causing repayment problems. At the level of the lender, there is a huge increase in employee efficiency, with the average credit officer selecting 63 percent more new individual borrowers over an average of 6 per month, increasing the number of loans by 27%, and this without any deterioration in portfolio performance. The reduction in adverse

selection from improved credit market information therefore presents strong benefits for lenders, generating more clients, larger loan portfolios, and improved repayment.

When group borrowers learn of the bureau and what it implies for them, we also see significant impacts on borrowing. Informed solidarity groups show an immediate improvement in repayment performance, with a sharp decline in delinquent final payments, solely due to reduction in moral hazard. Subsequent loans also exhibit a sharp decline in delinquency due to reduction both in moral hazard and adverse selection in group formation. There are, by contrast, no impacts of information about the credit bureau on the repayment performance of CB. For SG, there are important group compositional shifts, with a decline in female membership. Female SG members have less delinquent loans, but they have more intermediate repayment problems, which are documented by the credit bureau, to their detriment. Both SG and CB clients use their knowledge of the bureau to start getting access to credit from outside lenders. For CB members, outside loans taken by good clients do not lead to repayment problems. This is not the case for the less experienced clients for whom taking more loans leads to repayment difficulties.

Ironically, in several ways the impact of the credit bureau experiment confirms the efficacy of group lending in combating asymmetric information. First, the improvement in the lender's screening ability is larger for individual borrowers than for SG borrowers, indicating that the degree of asymmetric information was lower for group loans to begin with. Solidarity groups become smaller when bureau information is available, suggesting that the bureau acts as a substitute for information and incentives otherwise provided by larger groups. The trainings have no impact on the inside repayment performance of communal bank borrowers, where group incentives should have been the strongest to begin with. And finally, those less experienced CB borrowers who increase their net indebtedness run into repayment problems, indicating that the credit system in the absence of the bureau was providing as much credit as these clients could manage without default. Hence, while our results show that outcomes can be improved with the use of a bureau, they also highlight the value of group lending in enabling lenders to combat asymmetric information.

Genesis is a well-respected lender that used standard best practices for MFI loan screening prior to introduction of the bureau. Despite this, our results demonstrate that substantial asymmetric information remained in this market, because the bureau was highly effective at improving credit market outcomes. Since bureaus are a relatively low-cost

intervention, this implies that they should be made a part of efforts to achieve financial deepening in developing countries. Their use appears to be clearly to the benefit of lenders, and in a competitive market, this should lead to lower interest rates for borrowers over time. The losers from the introduction of a bureau are those borrowers who are screened out as a result of the information, and ongoing borrowers who may lose insurance opportunities as a result of the decline in solidarity group size. We show that group reporting can in fact reinforce the group mechanisms that underlie microfinance lending. The ultimate outcome is efficiency gains for the innovating institution, gains for the more capable borrowers, and increased economic differentiation across agents.

REFERENCES

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

Campion, A. & L. Valenzuela. 2001. "Credit Bureaus: A Necessity for Microfinance?" *Mimeo.*

Conning, J. 1999. "Outreach, Sustainability, and Leverage in Monitored and Peer-monitored Lending." *Journal of Development Economics,* 60: 51-77.

Galindo, A., & M. Miller. 2001. "Can Credit Registries Reduce Credit Constraints? Empirical Evidence on the Role of Credit Registries in Firm Investment Decisions," IADB working paper.

Gehrig, T., & R. Stenbacka. 2007. "Information Sharing and Lending Market Competition with Switching Costs and Poaching." *European Economic Review,* 51: 77-99.

Ghatak, M., & T. Guinnane. 1999. "The Economics of Lending with Joint Liability: Theory and Practice." *Journal of Development Economics,* 60: 195-228.

Gine, X., & D. Karlan. 2006. "Group versus Individual Liability: A Field Experiment from the Philippines." Yale Economic Growth Center working paper No. 940.

Grief, Avner. 1994. "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualistic Societies." *Journal of Political Economy*, 102: 912-50.

Jappelli, T., & Marco Pagano. 1999. "Information Sharing, Lending and Defaults: Cross-Country Evidence." *Journal of Banking and Finance*, 26(10): 2017-45.

Karlan, D., and J. Zinman. 2009. "Observing Unobservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment". *Econometrica,* forthcoming.

Luoto, J., C. McIntosh, & B. Wydick. 2007. "Credit Information Systems in Less-Developed Countries: Recent History and a Test." *Economic Development and Cultural Change*, 55(2): 313-334.

McIntosh, C., A. de Janvry, & E. Sadoulet. 2006. "How Increasing Competition Among Microfinance Institutions Affects Incumbent Lenders." *Economic Journal*, 115: 987-1004.

McIntosh, C., & B. Wydick. 2005. "Competition and Microfinance." *Journal of Development Economics,* 78: 271-298.

McIntosh, C., & B. Wydick. 2007. "Adverse Selection, Moral Hazard, and Credit Information Systems." University of California at San Diego.

Morduch, Jonathan. 1999. "The Microfinance Promise." *Journal of Economic Literature*, 37(4): 1569-1614.

Morduch, Jonathan & Beatriz Armendariz de Aghion. 2005. *The Economics of Microfinance*, Harvard University Press: Cambridge, Mass.

Navajas, S., J. Conning, & C. Gonzalez-Vega. 2003. "Lending Technologies, Competition and Consolidation in the Market for Microfinance in Bolivia." *Journal of International Development*, 15(6): 747-770.

North, Douglas. 1990. *Institutions, Institutional Change, and Economic Performance*. Cambridge University Press, Cambridge.

Padilla, J., & M. Pagano. 2000. "Sharing Default Information as a Borrower Discipline Device." *European Economic Review,* 44: 1951–1980.

Padilla, A.J., & M. Pagano. 1997. "Endogenous Communciation among Lenders and Entrepreneurial Incentives." *Review of Financial Studies,* 10 (1): 205-236.

Platteau, Jean-Philippe. 2000. *Institutions, Social Norms, and Economic Development*, Harwood Academic Publishers, Amsterdam.

Schreiner, M. 1999. "Lessons for Microenterprise Programs from a Fresh Look at the Unemployment Insurance Self-Employment Demonstration." *Evaluation Review,* 23 (5): 504-526.

Stiglitz, J., & A. Weiss. 1981. "Credit Rationing in Markets with Imperfect Information." *American Economic Review* 71(3): 393-410

Vercammen, J. 1995. "Credit Bureaus and Sustainable Reputation Effects in Credit Markets." *Economica,* 62 (248): 461-478.

Table 1. Summary statistics on the Genesis loan portfolio six months prior to any entry in the credit bureau

	Individual borrowers	Solidarity Groups	Communal Banks
Number of clients	6543	8055	10449
Number of groups		1722	1356
Characteristics of clients			
Percent female	48.6	46.7	99.7
Percent married	78.7	77.7	84.6
Percent with no education	11.1	26.5	41.4
Percent with some primary education	66.7	68.6	55.8
Percent with more than primary	21.0	4.0	2.4
Average age	43.2	41.2	39.7
Percent ethnic	27.9	53.7	-[1]
Percent rural	21.7	26.5	-[1]
Average number of loans	3.6	6.8	4.3
New clients per month	544	667	902
Dropout clients per month	490	754	830
Characteristics of groups			
New groups per month		106	91
Dropout groups per month		118	81
Number of members		4.5	7.7
Characteristics of loans			
Total current loans	8111	3094	1977
New loans per month	518	516	327
Loan size per client (US$)	1137	662	324

[1] Missing information

Data source: Genesis administrative data. Genesis portfolio for the period February-July 2001.

Table 2. Tests of exogeneity of the credit bureau rollout

| | Branch level monthly average performance | | | | |
| | Loan more than 2 months delinquent | | Late fees as share of loan size | | Number of members |
	SG	Individual	SG	Individual	SG
Panel A: Monthly average of loan-on-loan changes					
Month Crediref began	0.0013	-0.0009	0.0001	0.0003	0.0022
	(1.10)	(0.68)	(1.39)	(0.95)	(0.81)
Observations	983	1079	983	1079	983
R-squared	0.05	0.06	0.04	0.05	0.09
Panel B: Monthly average performance					
Month prior to Crediref	0.0330	0.0242	0.0009	0.0043	-0.2143
	(0.39)	(0.59)	(0.20)	(0.94)	(0.87)
Month 2 prior to Crediref	0.0968	0.0149	0.0042	0.0050	0.1960
	(1.21)	(0.44)	(1.06)	(1.34)	(0.68)
Month 3 prior to Crediref	0.0244	0.0195	0.0017	0.0043	-0.1558
	(0.45)	(0.58)	(0.55)	(1.25)	(0.77)
Months 4-6 prior to Crediref	0.0339	0.0202	0.0007	0.0037	-0.0616
	(0.91)	(0.84)	(0.27)	(1.52)	(0.40)
Observations	360	413	360	413	360
Number of branches	34	35	34	35	34

Absolute value of t-statistics in parentheses, robust standard errors clustered at the branch level. * significant at 5%; ** significant at 1%.

Panel A: Branch/month level weighted regression with month fixed effects, for pre-treatment period, January 1998 to July 2001. "Month Crediref began" gives the number of months between January 2001 and the introduction of crediref in each branch.

Panel B: Branch/month level regression with branch and month fixed effects, for pre-treatment period, during the 12 months prior to each branch entry into Crediref.

Table 3. Impact of the credit bureau on screening borrowers in and out

	Individual borrowers	Borrowers within ongoing Solidarity Groups	Entire Solidarity Groups
Fraction leaving			
First time screened[1]	0.133	0.194	-0.429
	(7.86)**	(9.39)**	(7.71)**
Subsequent screenings	-0.020	0.001	-0.275
	(0.75)	(0.08)	(4.62)**
Observations	39,958	51,324	21,048
Number of branches	39	37	37
Fraction entering			
Treatment effect	0.149	-0.012	0.268
	(4.22)**	(1.66)	(5.03)**
Observations	39,958	48,362	21,048
Number of branches	39	35	37

Absolute value of t-statistics in parentheses, robust standard errors clustered at branch level.

* significant at 5%; ** significant at 1%

Loans extended February 1999 to May 2004. All regressions include branch and month fixed effects.

[1]Last loan taken before the availability of Crediref and paid off at most two months before this date.

Table 4. Loan performance of checked and ejected borrowers

| | Not checked | Checked in Crediref | | t-stat for test of difference | |
| | | Not ejected | Ejected | | |
	(1)	(2)	(3)	(1) vs (2)	(2) vs (3)
Individuals					
Internal performance prior to Crediref					
Late fees > 1% of principal (%)	8.8	13.4	12.5	(4.7)**	(0.4)
Loan more than 2 months delinquent (%)	1.4	1.8	1.3	(0.9)	(0.5)
External performance on loans seen in bureau					
Percentage of loans with arrears		20.8	36.7		(6.2)**
Average size of loans with arrears (US$)		48.9	77.4		(1.7)
Number of observations	1663	2152	291		
Members of Solidarity Groups					
Internal performance prior to Crediref					
Late fees > 1% of principal (%)	6.9	7.2	15.3	(0.4)	(4.5)**
Loan more than 2 months delinquent (%)	3.0	2.4	7.4	(1.2)	(4.7)**
External performance on loans seen in bureau					
Percentage of loans with arrears		16.8	33.7		(6.6)**
Average size of loans with arrears (US$)		32.9	105.1		(4.0)**
Number of observations	1662	2562	244		

* significant at 5%; ** significant at 1%

Performance on loans taken by current clients from June 2000 to the entry into Crediref. Current clients are defined as having had a loan until at most 2 months prior to the entry into Crediref. Clients are designated as checked if their Crediref record was checked in the first two months of availability of the bureau, and among them, as ejected if they did not take subsequent loans. Not checked refers to continuing clients whose records have not been checked in Crediref.

Table 5. Impact of screening on individual client composition

| | Correlates of loan low performance prior to use of bureau | | Impact on borrower characteristics[3] | |
	Low Performance[1]	Variance in performance[2]	New clients ITE	Dropouts ITE
Male	0.0034	0.0010	0.028	0.077
	(4.85)**	(2.31)*	(1.27)	(6.10)**
Some secondary education	0.0039	0.0006	-0.111	0.039
	(2.62)**	(0.69)	(5.14)**	(2.98)**
Age	-0.0002	-0.00004	-0.309	0.034
	(4.94)**	(3.03)**	(0.71)	(0.15)
Observations	13,550	13,550	16,063	7,352
R-squared	0.09	0.04		

Absolute value of t-statistics in parentheses, robust standard errors clustered at the branch level.

* significant at 5%; ** significant at 1%

[1] Low performance is measured by the ratio of late payment fines to the principal. Regressions include branch and month fixed effects. Observations are all individual loans extended February 1999 to July 2001.

[2] Regression of square residuals from the performance regression, including branch and month fixed effects.

[3] Each coefficient comes from a separate regression of the row title characteristic on the treatment and the other two characteristics, with branch and month fixed effects. New clients for the period February 1999 to May 2004. Dropout clients from February 1999 to the first screening after the availability of the credit bureau.

Table 6. Impact of screening on the extensive margin: Performance of new clients

| | Performance of first loan | | | Future behavior | | |
	Late fees > 1% of principal	Loan more than 2 months delinquent	Average loan size per borrower (US $)	Probability of taking subsequent loan	Months until subsequent loan taken	Growth of subsequent loan per borrower
Individual loans						
Treatment effect	-0.056	0.008	79	0.240	-0.127	0.122
	(4.29)**	(0.54)	(1.10)	(8.40)**	(1.97)*	(2.35)*
Observations	13040	13040	13040	13040	5815	5814
R-squared	0.02	0.01	0.01	0.11	0.04	0.05
Mean of dependent variable	0.23	0.08	882	0.45	1.98	1.37
Solidarity Group loans						
Treatment effect	0.005	-0.008	209	0.069	0.029	0.039
	(0.24)	(0.35)	(6.85)**	(1.14)	(0.18)	(0.54)
Observations	5443	5443	5443	5443	2881	2881
R-squared	0.03	0.02	0.07	0.14	0.04	0.07
Mean of dependent variable	0.14	0.11	495	0.53	1.6	1.35

Absolute value of t-statistics in parentheses, robust standard errors clustered at branch level. * significant at 5%; ** significant at 1%

New individual clients and new solidarity groups, from February 1999 to May 2004. All regressions include month and branch fixed effects.

Table 7. Impact of screening on the intensive margin: Performance of ongoing borrowers

	Late fees > 1% of principal	Loan more than 2 months delinquent	Average loan size per borrower (US$)
Individual loans			
Treatment effect	0.036	0.023	82
	(2.29)*	(3.74)**	(2.15)*
Observations	11203	11203	11203
R-squared	0.04	0.02	0.19
Number of borrowers	3256	3256	3256
Mean of dep. variable	0.120	0.020	1058
Solidarity Group loans			
Treatment effect	-0.009	0.009	208
	(0.67)	(1.39)	(6.60)**
Observations	8796	8796	8796
R-squared	0.03	0.02	0.31
Number of borrowers	1149	1149	1149
Mean of dep. variable	0.040	0.020	872

Absolute value of t-statistics in parentheses, robust standard errors clustered at branch level.

* significant at 5%; ** significant at 1%

Sample of borrowers who have taken loans before and after the credit bureau. All loans extended after February 1999 and closing by June 2004. All regressions include individual and month fixed effects, and dummies for loan cycle.

Table 8. Aggregate effect of use of the bureau for screening on credit agent performance

	Credit officer portfolio		Branch level portfolio	
	Average value	Treatment effect[1]	Average value	Treatment effect[2]
New clients per month				
Individual clients	1.56	0.991	6.39	5.85
		(5.08)**		(5.12)**
Members of new Solidarity Groups	1.83	2.73	7.49	15.62
		(7.37)**		(5.41)**
Members of new Communal Banks	2.41	0.27	9.89	-0.18
		(0.47)		(0.07)
Loans per month				
Number	7.17	1.927	29.38	15.319
		(3.31)**		(5.99)**
Total loan amount (US $)	11,340	2,343	46,461	24,161
		(1.80)		(3.08)**
Loan performance				
Loans with late fees >1% of principal (%)	0.06	-0.02	0.17	-0.003
		(1.20)		(0.20)
Loans more than 2 months delinquent (%)	0.04	0.007	0.08	0.013
		(0.51)		(0.75)

Absolute value of t-statistics in parentheses, robust standard errors clustered at the branch level.

* significant at 5%; ** significant at 1%

[1]Each coefficient comes from a separate regression on the treatment, with month and credit officers fixed effects. Monthly indicators for 425 credit agents on loans extended after February 1999 and closing by June 2004.

[2]Each coefficient comes from a separate regression on the treatment, with month and branch fixed effects.

Table 9. Randomization tests for the information experiment

	Solidarity Groups			Communal Banks		
Panel A: Comparison of pre-treatment covariates and outcomes						
	Control groups mean	Selected groups mean	t-stat of test of difference	Control groups mean	Selected groups mean	t-stat of test of difference
Group characteristic:						
Number of members	2.28	2.27	(0.15)	9.12	9.28	(0.94)
Loan amount per capita (in US$)	886	854	(0.47)	297	296	(0.02)
Percent divorced	1.61	1.46	(0.20)	1.00	0.48	(1.04)
Percent widowed	4.10	4.51	(0.24)	3.69	4.15	(0.96)
Percent with no education	3.34	4.38	(0.99)	34.0	35.9	(0.81)
Percent with some primary educ.	89.5	90.4	(0.57)	61.9	60.5	(0.76)
Percent with more than primary educ.	6.80	5.02	(1.03)	4.11	3.44	(0.67)
Percent female	55.3	50.5	(1.01)	100	99.9	(0.97)
Average age	38.8	37.7	(1.26)	36.0	35.6	(1.29)
Loan performance						
Late fees > 1% of principal (%)	11.2	6.80	(1.66)	0.62	0.00	(0.92)
Loan more than 2 months	5.88	4.17	(0.53)	0.96	1.15	(0.65)
Number of observations	354	680		299	285	
Panel B: False treatment effect on repayment outcomes						
Borrowers pays late fees > 1% of principal		-5.21			-0.41	
		(1.45)			(0.08)	
Loan more than 2 months delinquent		-0.67			-3.23	
		(0.34)			(1.48)	
Observations		819			368	

Loans taken from February 2003 to June 2004 by the 581 groups participating to the training experiment. Observations weighted according to the sampling design.

Panel B: Regressions of loan level performance on a false treatment variable, with group and time fixed effects. The pre-treatment period is divided in two equal periods with a false treatment set in September 2003. All coefficients multiplied by 100.

Table 10. Impact of information on performance

Outcome:	Intermediate payments[1] ITE	Within loan treatment Final payment Difference[2] ITE	Loan performance Difference[2] ITE	Across loans treatment Loan performance[3] ITE
Solidarity Group loans				
Delinquent payment[4]	-5.51	-10.22		
	(1.76)	(2.30)**		
Amount of late fees (US$)	0.42	-0.35		
	(1.46)	(0.31)		
Late fees > 1% of principal[4]			-6.76	0.98
			(0.86)	(0.24)
Loan more than 2 months delinquent[4]			-6.87	-6.46
			(0.79)	(3.39)**
Number of observations	4015	396	396	1237
Communal Bank loans				
Delinquent payment[4]	0.88	-1.65		
	(1.76)	(0.44)		
Amount of late fees (US$)	-0.15	-0.48		
	(1.25)	(0.50)		
Late fees > 1% of principal[4]			-3.30	–
			(1.20)	
Loan more than 2 months delinquent[4]			-3.09	–
			(1.30)	
Number of observations	3420	339	339	566

Sample in columns 1-3 includes the loan that was current at the time of the treatment for each of the 735 groups. Sample in column 4 includes all loans granted from February 2003 and closed by June 2005, except the loan current at treatment time. Observations weighted according to the sampling design. Absolute value of t-statistics in parentheses, robust standard errors clustered at branch level. * significant at 5%; ** significant at 1%.
[1]Each coefficient corresponds to a separate regression of an outcome on the treatment variable, loan and time fixed effects, and dummy variables for payments deciles.
[2]Each coefficient corresponds to a simple difference between treated and control groups.
[3]Each coefficient corresponds to a separate regression that includes group and time fixed effects. None of the 23 loans taken after the training have late fees higher than 1% of principal or are 2 months delinquent.
[4]Coefficients multiplied by 100.

Table 11. Impact of information on Solidarity Group size and composition

	Correlates of repayment performance[1]		Impact on the average characteristics of clients[2]	
	Late fees amount (US$)	Loan > 2 months delinquent	New clients ITE	Dropouts ITE
			New/dropout members	
Number of members	-0.56	0.76	-0.029	-0.023
	(0.33)	(0.92)	(0.91)	(0.90)
Average characteristics by group				
Female ratio	5.68	-5.51	-0.252	0.167
	(1.60)	(2.42)*	(1.45)	(2.05)*
Primary education ratio	28.48	13.01	-0.155	-0.105
	(3.77)**	(7.71)**	(2.02)*	(1.45)
Some secondary education ratio	42.60	9.43	-0.006	0.023
	(4.53)**	(4.48)**	(1.02)	(1.73)
Average age	-0.24	0.01	-7.095	-1.492
	(1.34)	(0.05)	(1.88)	(0.85)
Number of observations	921	921	143	216

Absolute value of t-statistics in parentheses, standard errors clustered at the branch level.

* significant at 5%; ** significant at 1%

[1]Loan level regression of the repayment performance on group characteristics, with branch and time fixed effects, for loans granted from February 2003 on and due before the training date. Coefficients for delinquent loans are multiplied by 100.

[2]Each coefficient corresponds to a separate regression of the characteristic of the group new/dropout clients on the treatment effect, with time and branch fixed effects. Observations are loans with new/dropout clients for the characteristics, but all 1738 loans for the number of new/dropout members.

Table 12. Impact of information on outside borrowing

	Number of clients	Change in number of loans[1]	Start taking outside loan[1]	Ever missed a payment[2]
Solidarity Group members				
Treatment effect - ITE	1247	-0.010	0.100	-0.062
		(0.13)	(2.47)*	(1.91)
Mean value in control group		0.370	0.370	0.13
Heterogeneity among Solidarity Group members[3]				
Less experienced clients - ITE	698	0.012	0.145	-0.054
		(0.14)	(2.99)**	(1.45)
More experienced clients - ITE	549	-0.075	0.034	-0.102
		(0.61)	(0.55)	(2.15)*
Good client - ITE	874	-0.032	0.072	-0.011
		(0.35)	(1.45)	(0.42)
Bad client - ITE	373	0.031	0.120	-0.237
		(0.26)	(1.88)	(2.91)**
Communal Bank members				
Treatment effect - ITE	4172	0.178	0.117	0.023
		(2.75)**	(3.59)**	(1.03)
Mean value in control group		0.230	0.360	0.08
Heterogeneity among Communal Bank members[3]				
Less experienced clients - ITE	2717	0.144	0.147	0.047
		(2.50)*	(4.00)**	(2.23)*
More experienced clients - ITE	1455	0.243	0.041	-0.014
		(2.07)*	(0.87)	(0.37)
Good client - ITE	3572	0.206	0.122	0.025
		(3.04)**	(3.48)**	(1.91)
Bad client - ITE	600	0.010	0.099	0.062
		(0.07)	(1.31)	(0.63)

Absolute value of t-statistics in parentheses. * significant at 5%; ** significant at 1%

[1]Analysis at the client level, with branch fixed effects, and standard errors clustered at the group level.

[2]Analysis at the loan level, with time and branch fixed effects, and standard errors clustered at the group level.

[3]Experienced (less experienced) clients are clients having had at least 4 (less than 4) loans with Genesis. Good (bad) clients had no (at least one) delinquent repayment before.

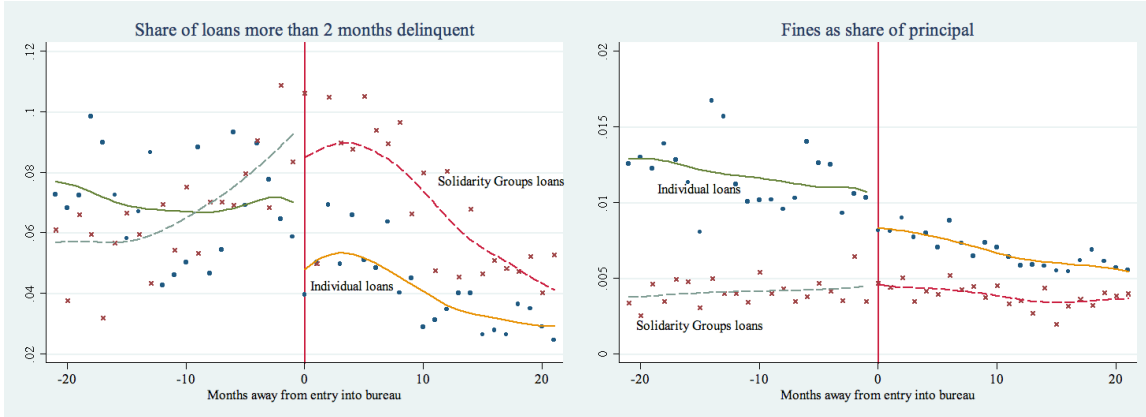**Figure 1. Improvement in performance when the bureau is used**

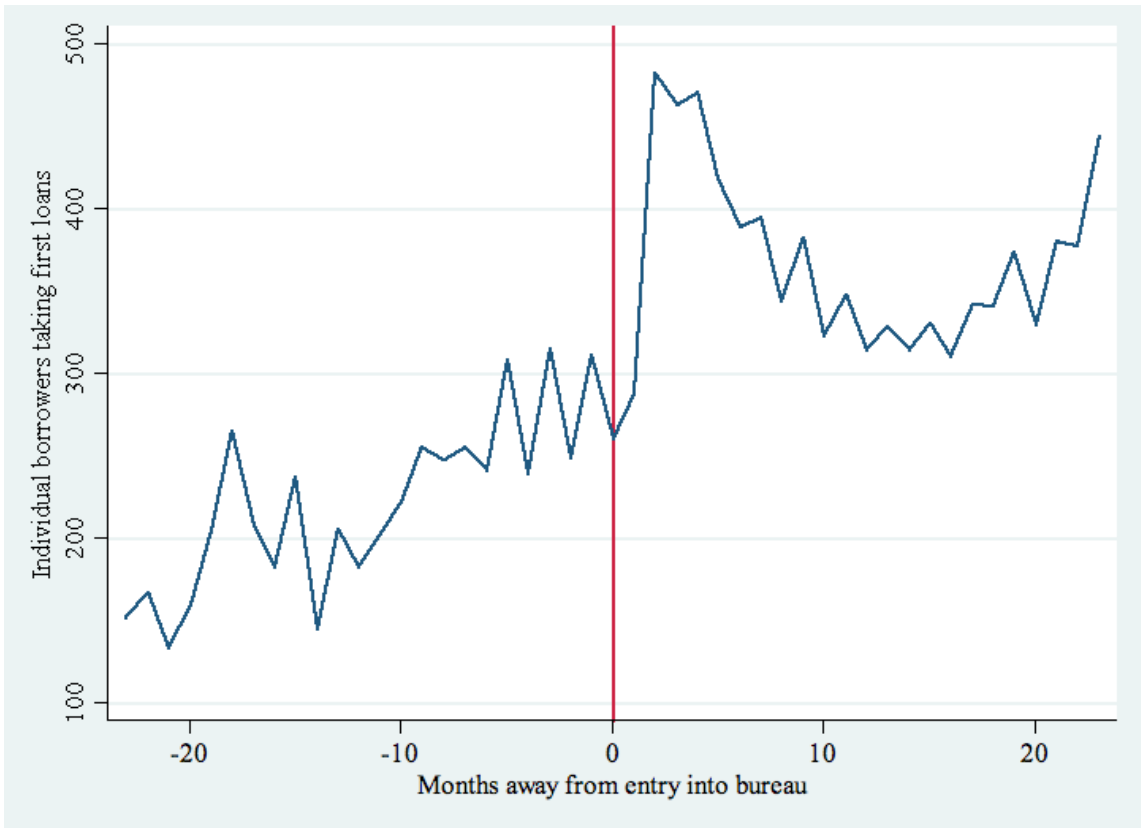**Figure 2. Increase in new individual loans when the bureau is used**

Figure 3. Decrease in the size of solidarity groups when the bureau is used.