

The Economics of Microfinance

**Beatriz Armendáriz de Aghion
Jonathan Morduch**

April 2004

Forthcoming from the MIT Press, 2005

© Beatriz Armendáriz de Aghion and Jonathan Morduch

Chapter 4

Group Lending

Introduction

4.1 The Group Lending Methodology

4.2 Mitigating Adverse Selection

4.2.1 Numerical Example

4.2.2 Group Lending Beyond Villages

4.3 Overcoming Moral Hazard

4.3.1 *Ex ante* Moral Hazard and the Role of Joint Responsibility

4.3.2 *Ex post* Moral Hazard and the Role of Peer Monitoring

4.4 Evidence on Groups and Contracts

4.5 Limits to Group Lending: Hidden Costs, Collusion, and Emerging Tensions

4.6 Summary and Conclusions

Exercises

Introduction

Once every week in villages throughout Bangladesh, groups of forty villagers meet together for half an hour or so, joined by a loan officer from a microfinance organization. The loan officer sits in the front of the group (the “center”) and begins his business.¹ The large group of villagers is sub-divided into eight five-person groups, each with its own chairperson, and the eight chairs, in turn, hand over their group’s passbooks to the chairperson of the center, who then passes the books to the loan officer. The loan officer duly records the individual transactions in his ledger, noting weekly installments on loans outstanding, savings deposits, and fees. Quick arithmetic on a calculator ensures that the totals add up correctly, and, if they do not, the loan officer sorts out discrepancies. Before leaving, he may dispense advice and make arrangements for customers to obtain new loans at the branch office. All of this is done in public, making the process more

¹ The loan officer is typically a man and the villagers are typically women, but there are exceptions.

transparent and letting the villagers know who among them is moving forward and who may be running into difficulties.²

This scene is repeated over 70,000 times each week in Bangladesh by members and staff of the Grameen Bank, and versions have been adapted around the world by Grameen-style replicators.³ Other institutions instead base their methods on the “solidarity group” approach of Bolivia’s BancoSol or the “village bank” approach operated by microlenders in 70 countries throughout Africa, Latin America, and Asia (including affiliates of FINCA, *ProMujer*, and Freedom from Hunger).⁴ For many, this kind of “group lending” has become synonymous with microfinance.⁵

“Group lending” refers specifically to arrangements by individuals without collateral who get together and form groups with the aim of obtaining loans from a lender. The special feature is that the loans are made individually to group members, but all in the group face consequences if any member runs into serious repayment difficulties. In the Grameen Bank case, the groups are made up of five people. In the BancoSol case, groups can be as small as three people, and in the village banking system pioneered by FINCA groups can range from 10 to 50 women.⁶ The fundamental idea of “group responsibility” (sometimes called “joint liability”) coupled with regular group meetings is common across approaches.

Grameen Bank’s weekly group meetings have some obvious and simple advantages for the lender and customers. Most immediately, they offer convenience to the villagers; the bank comes to them, and any problems (a missing document, being a few taka short) can be resolved on the spot. The bank thus offers the same convenience as a local ROSCA or moneylender. Meanwhile, transactions costs are greatly reduced for

² Helen Todd (1996) provides a detailed and unvarnished study of group lending in Bangladesh. David Borsenstein (1997) provides a journalist’s account of group meetings and the Grameen Bank story. See also Fugelsang et al. (1993).

³ By December 2002, the Grameen Bank had 2,483,006 members organized into 70,928 centers and 513,141 groups. So, on average, there were 35 individuals per center and 7.25 groups per center. Data are from *Grameen Bank Monthly Update, December 2002* (#276, January 25, 2003). See www.grameen-info.org.

⁴ FINCA is the Foundation for International Community Assistance. See www.villagebanking.org.

⁵ In the next chapter, though, we argue that there is much more afoot in microfinance than group lending, although it has played a historically important role.

⁶ Both the Grameen Bank and BancoSol now also make many loans on a strictly bilateral basis, without the “group responsibility” contract. The “individual” contract (as opposed to the “group” contract) is viewed as being more appealing to better-off, more mature members.

the loan officer since the multiple savings and loan transactions of 40 people can take place in a short block of time.

Transacting through groups also has more subtle advantages (and some limitations). In particular, the “group responsibility” clause of contracts can mitigate the moral hazard, adverse selection, and enforcement problems that crippled previous attempts at lending to the poor by outside financial institutions. In chapter 2 we described how these problems are caused by information asymmetries, and one implication is that if the bank gets more information, they can always do better. A solution to the resulting inefficiency is thus to create contracts that generate better information.⁷

But the contracts described below all improve matters *without* the bank necessarily learning anything new. Instead, the contracts take advantage of the fact that group members themselves may have good information about fellow members--and the contract gives the members incentives to use their information to the bank’s advantage. This can occur in subtle ways, and we present different scenarios in turn.

While the advantages of group lending will be spelled out, there is another side to the coin. Might groups collude against the microlender by collectively deciding not to repay? If the group of borrowers is not willing to impose social sanctions upon themselves, can the group nonetheless provide advantages? Another set of questions relates to peer monitoring. What will happen if the population of potential borrowers is dispersed and local information is thus weak and costly to obtain? If group lending takes place in urban areas, where labor mobility is high and individuals also may not have much information about their potential partners, are there still any advantages for groups? And if borrowers cannot observe each other’s effort levels (or are otherwise reluctant to punish shirkers) then group-lending can undermine incentives by encouraging “free riding”. Borrowers will ask themselves: Why should I work hard if I am liable for a penalty when my partner shirks—even when I cannot control their actions? Sections 4.4 and 4.5 investigate ways that group-lending has enabled outside lenders to expand credit

⁷ Such “information revelation mechanisms” are described by Rai and Sjöström (2004). They provide an interesting example of a hypothetical mechanism that reveals information by inducing villagers to “cross-report” on each other, and they show conditions under which cross-reporting can dominate the Grameen-style contract described here. We return to their proposal in the next chapter.

access in low-income communities, but we also point to tensions and imperfections in the approach—which suggest turning as well to some of the alternative mechanisms described in chapter 5.

4.1 The Group Lending Methodology

Access to finance via groups is not new. The example of ROSCAs in the previous chapter shows how groups function to give participants access to a “pot” of communal money, and credit cooperatives similarly function to allow members to obtain loans from their peers. The place of groups in microfinance, however, strengthens and extends earlier uses of groups (although not without some added costs).

To see this, we describe “Grameen-style” group lending. The model has been adapted in different contexts, but replicators have tried to stay true to the main features described below. The Grameen Bank itself has undergone changes in the twenty-five years since it started (most recently with a major overhaul dubbed “Grameen Bank II”), and we will describe elements of what is now called the Grameen “classic” system (Yunus, 2002). This is the model that has figured most prominently in economic research.⁸ We will return to Grameen II in section 4.5 below.

When the Grameen Bank first got started as an experimental bank in the village of Jobra, near Chittagong University, the first loans were made to individuals without a group responsibility clause (Muhammad Yunus, 1994). Instead, economies of scale motivated the first use of groups. But Yunus and his associates soon realized that requesting potential borrowers to organize themselves into groups had another advantage: the costs of screening and monitoring loans and the costs of enforcing debt repayments could be substantially reduced.⁹

To institute this systematically, the bank developed a system in which two members of each five-person group receive their loans first.¹⁰ If all installment are paid

⁸ An excellent overview of the theory of group lending is provided by Maitreesh Ghatak and Timothy Guinnane (1999).

⁹ Grameen restricts membership to people that do not possess more than half an acre of land, although the rule is followed more in spirit than in letter. This definition obviously does not apply to other countries where the Grameen methodology has been replicated.

¹⁰ The maturity period varies across borrowers and countries. But most replicators are advised to extend one-year loans which are to be repaid weekly, that is, in 52 installments. As of 2002, Grameen is allowing for variable loan terms but keeping weekly repayment plans.

on time, the initial loans are followed four to six weeks later by loans to two other members, and then, after another four to six weeks, by a loan to the group chairperson. (This pattern is known as 2:2:1 staggering.) At first, the groups were seen just as sources of solidarity, offering mutual assistance in times of need. For example, if a member of a group fails to attend a meeting, the group leader repays on her behalf, and thus the credit record of the absentee borrower remains clean, and so does the group's. The original premise was that perhaps someone might experience a delay in getting a loan if there was a problem within their group, but there would not be further sanctions.¹¹

Over time, though, formal sanctions became more common. In principle, if serious repayment problems emerge, all group members will be cut-off from future borrowing. The original idea was not that group members would be forced to repay for others, rather it was that they would lose the privilege of borrowing. In practice, of course, a borrower that does not want to lose access to microcredit loans accepts the possibility of having to bail out their fellow group members in times of need. It is not unheard of that a loan officer will stay in a village until group members (or members of the 40-person center) are able to make good on all installments due that week (although the practice is not in keeping with the early vision of top Grameen managers).¹²

In a typical situation, when all goes well with repayments, borrowers are offered a larger loan repayable in the next "loan cycle" (loan cycles—from initial disbursement to repayment of the final installment—were typically a year in the "classic" Grameen system). Thus, if the relationship between Grameen and the borrowers continues, loan sizes grow over the years and credit histories are built up. Eventually loans may be large enough to build or repair a house or to make lumpy investments like purchasing a rickshaw or, in a recent loan innovation, sending a child to university.

¹¹ Jonathan Morduch interview with Muhammad Yunus, December 15, 2002, Dhaka. One advantage of the 2:2:1 staggering, pointed out to us by Imran Matin, is that it increases the chance that a group member is awaiting a new loan when another group member runs into repayment trouble.

¹² Gonzalez-Vega, et al. (1997, p. 88) report that in BancoSol's version of group-lending in Bolivia, loan officers refuse to accept partial loan repayments from a group. So if one member cannot come up with the required money in a given week, the loan officer will not accept any group member's individual contribution for that week—and all members are seen to be in arrears. Funds are only accepted when everyone has 100 percent of their contributions ready to submit. Like the Grameen Bank rules, this creates strong incentives (if enforced) to encourage group members to work hard, manage funds wisely, and help their peers.

4.2 Mitigating Adverse Selection

The adverse selection problem occurs when lenders cannot distinguish inherently risky borrowers from safer borrowers. If lenders could distinguish by risk type, they could charge different interest rates to different types of borrowers. But with poor information, options are limited. As we saw in section 2.2, adverse selection may lead to credit rationing as it induces lenders to charge everyone high interest rates to compensate for the possibility of having very risky borrowers in the customer population. The trouble (and source of inefficiency) arises when safe borrowers are thus deterred from applying for loans. In principle, group lending with joint responsibility can mitigate this inefficiency.¹³

Consider a microfinance institution or a bank committed to covering its costs so that it just breaks even.¹⁴ Assume that the bank introduces the group lending methodology described above, and that it has no idea about the borrowers' characteristics. Borrowers, on the other hand, know each other's types, and, as in section 2.2, borrowers are either "risky" or "safe". As above, the problem is that the bank wants to charge lower interest rates to safe borrowers and higher rates to risky borrowers, but, since the bank cannot easily tell who is who, everyone has to pay the same rate. In practice, then, the safer borrowers--when they actually decide to apply for a loan at the prevailing interest rates-- implicitly subsidize the risky borrowers (who are more costly for the bank to serve). The inefficiency arises when this implicit subsidy is so large that safe borrowers leave the market rather than shouldering the burden--i.e., when the presence of risky borrowers raises the interest rate to levels that are simply unaffordable for safer borrowers. The question here is whether group lending can make it possible to *implicitly* charge safe borrowers lower interest rates and thus keep them in the market.

The fact that groups are encouraged to form on their own is the key to the solution; potential borrowers can then use their information to find the best partners. How they sort themselves depends on the nature of the loan contract. Faced with the

¹³ The exposition here follows treatments by Maitreesh Ghatak (1999) and Beatriz Armendáriz de Aghion and Christian Gollier (2000); also see Ghatak (2000). Varian (1990) provides an earlier paper on group lending and adverse selection, while Laffont and N'Guessan (2000) provide a more recent treatment.

¹⁴ Henceforth we will use the word "bank" bearing in mind that the institution is special in that it is committed to just breaking even, or that it is in a perfectly competitive market so that it cannot charge more than its costs.

prospect of joint responsibility for loans, it is clearly better to be grouped with safe types than with risky types. So, given the choice, the safe types stick together. The risky borrowers thus have no alternative but to form groups with other risky types, leading to a segregated outcome often referred to in the labor economics literature as “assortative matching.”¹⁵

How does this help the bank charge lower prices to safe types? Because investment projects undertaken by risky borrowers fail more often than those of safe borrowers, risky borrowers have to repay for their defaulting peers more often under group lending with joint responsibility; otherwise, they will be denied future access to credit. Safe borrowers no longer have to shoulder the burden of default by the risky types. What this boils down to is a transfer of risk from the bank onto the risky borrowers themselves. It also means that, effectively, the safe types pay lower interest rates than the risky types – because they no longer have to cross-subsidize risky borrowers. Strikingly, the result is that the group lending methodology does the trick even though (1) the bank remains as ignorant as ever about who is safe and who is risky, and (2) all customers are offered exactly the same contract. All of the action occurs through the joint responsibility condition combined with the sorting mechanism.

Moreover, because banks are now better insured against defaults, average interest rates for both risky and safe types can be reduced while banks still make profits. The lower interest rates in turn bring a secondary positive effect. In the adverse selection problem analyzed in section 2.2, “safe” borrowers were inefficiently pushed out of the market by high interest rates; here, the reduction in interest charges faced by safe types further encourages them to re-enter the market, mitigating the market failure.

To see this formally, suppose that the bank requests that borrowers form two-person groups and that each individual in the pair holds herself responsible for her peer.¹⁶ As in section 2.2.2, the analysis is simplified by assuming that individuals try to

¹⁵ The question arises as to why risky types (who earn higher profits than safe types in good periods), cannot simply pay safe types to join with them. Ghatak (1999) provides a proof of why risky types cannot adequately compensate safe types to induce the safe types into mixed safe-risky groups. The numerical example shows this too.

¹⁶ Analyzing 5-person groups is straightforward but adds complications with little extra insight. Similarly, considering risk aversion alters the main results only slightly.

maximize their expected income without concern for risk. As before, we first present the analysis using algebra and then provide a simple numerical example.

Again, each individual has a one-period project requiring \$1 of investment. The fraction of the population that is safe is $q < 1$, and the fraction of the population that is risky is $(1 - q)$. A dollar invested by safe borrowers yields a gross return \underline{y} with certainty.¹⁷ A risky borrower who invests \$1, on the other hand, obtains a gross return $\bar{y} > \underline{y}$ if successful, and this occurs with probability $p < 1$. If not successful, they earn 0, which happens with probability $(1 - p)$. Again, to simplify things we assume that both types have identical *expected* returns, so that $p\bar{y} = \underline{y}$. How do the types sort themselves into groups? Since borrowers know each others' types, safe borrowers pair with other safe types, and risky borrowers pair with other risky types (i.e., there will be assortative matching in equilibrium). Now consider more closely situations where both types of borrowers participate in the credit market. Since the fraction of the population that is safe is $q < 1$, this will also be the fraction of groups made up of (safe, safe) types. If, say, a quarter of the population is "safe", then a quarter of the two-person groups will be made up of "safe" couples.

What is the gross interest rate R_b (principal plus interest) that the bank should charge in order to break even? To make the problem interesting, assume that $\bar{y} > 2R_b$ so that, when lucky, a risky borrower can always repay for her peer. Then the expected revenue of the bank if it sets its break-even interest rate at R_b is straightforward to compute: with probability q the bank faces a (safe, safe) pair of borrowers and therefore gets repaid for sure; with probability $(1 - q)$, the bank faces a (risky, risky) pair, in which case it is always repaid unless both borrowers in that pair have a bad draw; we denote the probability that the bank is repaid in this case is g . Since the chance that *both* are simultaneously unlucky is $(1 - p) \cdot (1 - p)$, the chance that one or both are lucky is $g = 1 - (1 - p)^2$. The expected payment to the bank is thus:

$$[q + (1 - q)g] R_b.$$

¹⁷ By working with gross returns and gross interest rates, we define returns as not being net of the cost of borrowing. The safe types' net returns are $(\underline{y} - R_b)$, for example.

The equation reflects that a fraction q of groups return R_b always (i.e., the safe groups) and a fraction $(1-q)$ of groups return R_b just g proportion of the time. This expected payment must be equal to the bank's cost of funds k in order for the bank to break even in expectation. Solving for R_b gives:

$$R_b = k / [q + (1-q)g],$$

which is smaller than the interest rate in the absence of group lending found in chapter 2 (there, without group-lending, we found that $R_b = k / [q + (1-q)p]$). The fact that the interest rate is smaller here arises because $g > p$; that is, the process of matching means that risky borrowers can pay back their loans more often (thanks to joint liability) than they could if just dealing with the bank as individuals. The risk is thus passed on from the bank to the risky borrowers. The bank can thus reduce the interest rate and lure deserving safe types back into the market.

The beauty of the arrangement is that all borrowers face the *same contract*, but, thanks to assortative matching, the risky types pay more on average. The bank thus effectively price discriminates—without needing to know who is safe and who is risky.

4.2.1 Numerical Example

To see how this works with numbers, return to the numerical example in section 2.2.3. There we showed a situation in which asymmetric information led to inefficiency. Here, we show a group-based contract that solves the problem.

The basic set-up is exactly as before. From the lender's viewpoint, half of the population is safe (they're always successful) and half is risky (they fail 25 percent of the time). Both safe and risky types are risk neutral and need \$100 to undertake a month-long project. Their alternative is to work for a wage for \$45. If the bank lends money, they need to recover costs equal to \$40 per month per loan. The gross revenues of safe types is \$200, and the gross revenues of risky types is \$267. The basic data are shown in Table 4.1.

Table 4.1

Group-Lending Numerical Example: Base Data

The economic environment			
Lender's cost of capital		\$40 per month per \$100 loan	
Borrowers' opportunity cost (wage)		\$45 per month	
Fraction of safe borrowers in the population		50%	
	Gross revenues if successful	Probability of success	Expected gross revenues
Safe type	\$200	100%	\$200
Risky type	\$267	75%	\$200
Group lending contract:			
	Gross interest due if borrower is successful:	\$150	
	Payment due if borrower fails:	\$0	
	Additional payment due if borrower is successful but partner fails:	\$117	
Borrower's expected net returns under the contract:			
		Partner Type	
		<u>Safe</u>	<u>Risky</u>
Borrower Type	Safe	\$50	\$21
	Risky	\$88	\$66

Given this situation, we saw in section 2.2.3 that there was no interest rate at which the bank could cover its costs and still entice everyone to borrow—if it used a standard individual lending contract. Here we show how a contract with joint responsibility can help the bank do better. Consider a contract offered to two-person groups in which the interest rate per borrower is 50 percent, payable only if the borrower's project is successful (i.e., their total payment to the bank is \$150, including principal). The contract also specifies that if a borrower succeeds but their partner fails, the borrower is liable for another \$117 (which is as much as the bank can extract, given risky types' gross revenues of \$267).¹⁸

Now what happens? Borrowers are asked to choose their partners. Does assortative matching occur? Yes: groups will never be mixed by type. To see why,

¹⁸ This is not the optimal contract that the bank could use, but it is sufficient to show how group lending can restore efficiency in the face of adverse selection. Note that joint liability/group responsibility contracts cut off all group members if any one of them defaults. Implicitly this means that they must find a way to make good on the defaulter's debts in order to escape sanctions. We assume that the debts are simply paid by the partners, but an informal loan might be used rather than a grant to the defaulter.

consider the expected net returns under the contract. The four possible scenarios are shown in Table 4.1. If a safe type matches with a safe type, both borrowers know that they will owe \$150 at the end of the month, leaving a \$50 net profit. If a risky type matches with a risky type, they know that they will be successful 75 percent of the time. And $0.25 \cdot 0.75$ of the time, they will owe the “joint liability” payment of \$117. Their expected payment is thus $0.75 \cdot (\$150 + 0.25 \cdot \$117) = \$134$, leaving a \$66 net profit. Can mixed pairs do better? Risky types clearly prefer to group with safe types (net profit = \$88 versus \$66), but can risky types afford to compensate safe types enough to induce them into partnerships? No, since safe types would demand an extra “side payment” of at least \$29 ($= \$50 - \21) to compensate for teaming with risky types. But the risky types’ net gain from teaming with safe types is only be \$22 ($= \$88 - \66). So, like matches with like.

The implication is that safe types now earn enough to make borrowing worthwhile. So everybody wants to borrow, and efficiency is restored. Quick calculations will confirm that the bank wants to lend under this contract too, since it can make a profit of \$4 after covering costs, just breaking even.

4.2.2 Group Lending Beyond Villages

Not all microfinance programs start with close-knit borrowers with rich information on each other. Dean Karlan (2003), for example, describes village banks in the Andes town of Ayacucho (with a population of 150,000). The FINCA affiliate spreads the word about the village banks, and interested borrowers are invited to come to FINCA’s office to put their names on a list; once the list reaches 30 names (typically in less than two weeks), a group is formed. The process is easy and efficient, but a consequence is that few of the group members know each other before joining the village bank.

Section 4.2, in contrast, showed how banks can circumvent credit rationing due to adverse selection through group lending when borrowers are perfectly informed about each others’ types. The village banks of Ayachucho represent a different context, one more typical of urban areas such as Mexico City and Bogotá, where populations are highly mobile and often have little information about each other. Can group lending still

help to overcome adverse selection? Can group lending carry benefits even if the “getting to know each other” process is slow or imperfect?

Consider the extreme scenario where potential borrowers remain completely anonymous; that is, they do not have any information about the characteristics of their peers. Group lending can no longer lead to assortative matching; instead, it will typically involve mixed pairs of safe and risky borrowers. Is this enough to discourage safe borrowers from applying for a loan? Can an appropriately-structured group lending contract improve on standard “individual-lending” contracts?

As in the previous section, risky borrowers will gain from the possibility of matching with a safe borrower who can always repay for them. But can safe borrowers gain too? Yes, if the contract takes advantage of the possibility that when risky borrowers are lucky, they get higher returns than safe borrowers. The optimal group lending contract will in practice extract more from risky borrowers when they are lucky but paired with an unlucky risky borrower, while the contract will not extract as much from a safe borrower who is paired with an unlucky risky borrower. The reason is “limited liability” as described above. Group lending here makes risky borrowers indirectly cross-subsidize safe borrowers, allowing the latter to access loans at a lower interest rate than without group lending. Once again, lower interest rates mitigate the credit rationing problem by increasing the participation of safe borrowers in the credit market.

We show the potential for the welfare-improving use of group lending here, using a stylized example based on the analysis of Armendáriz de Aghion and Christian Gollier (2000). The example follows the spirit of the analysis at the start of section 4.2, and, as above, our goal is to show the potential for gains, rather than to claim that there will *always* be gains. More formally, again let R_b denote the gross interest rate set by the bank (set so that the bank just breaks even), and again suppose that returns are set such that $\underline{y} < 2 R_b < \bar{y}$. In this case, $\underline{y} < 2 R_b$ means that the safe borrowers are unable to fully pay for an unlucky partner’s failure. Groups are now matched randomly. Since a fraction q of the population is made up of safe types, the chance that a (safe, safe) pair

emerges through random matching is q^2 .¹⁹ Similarly, the chance that a (risky, risky) pair emerges is $(1-q)^2$. And the chance that a (safe, risky) pair emerges is accordingly $1 - q^2 - (1-q)^2$, or, after simplifying, $2q(1-q)$.

The bank's expected gross revenues are then $2R_b$ from (safe, safe) pairs. This is because both repay the interest rate with certainty. Since the expected fraction of matches that are (safe, safe) is q^2 , the bank expects to get $2R_b$ in a fraction q^2 of cases. With probability $(1-q)^2$ the pair is (risky, risky), and the bank gets $2R_b$ if both are lucky. The chance that both are lucky is p^2 since p is the probability that either independently succeeds (again as in chapter 2). The probability that both risky borrowers fail is correspondingly $(1-p)^2$; in this case, the bank gets nothing back. And the chance that one is lucky while the other is not is $2p(1-p)$; in that case, the lucky partner can pay for both, so the bank gets $2R_b$ once more. Finally, with probability $2q(1-q)$, the bank faces a mixed (safe,risky) pair. We know that the safe partner always does well, so the question is: what happens to the risky partner? If the risky partner is lucky (which happens p fraction of the time), the bank again gets $2R_b$. But $(1-p)$ of the time the risky partner has bad luck. Note that here the safe partner cannot fully pay for the risky partner (by the assumption that $y < 2R_b$). Instead, the bank can only extract the amount y from the the safe partner by the assumption of limited liability (i.e., the bank cannot extract more than the safe borrower's current revenue).

In equilibrium, the gross interest rate R_b must be set so that the expected repayment per borrower is equal to the bank's full cost of funds k . Since we are analyzing loans to each member in a two-person group, the expected gross repayment must be at least $2k$. Now we can put all of this information together to yield:

$$\begin{aligned} & q^2 2R_b \\ & + (1-q)^2 (p^2 + 2p(1-p)) 2R_b \\ & + 2q(1-q)[p 2R_b + (1-p)y] \end{aligned}$$

¹⁹ The probability that two independent events occur is the product of probabilities. If you randomly chose someone from the population there would be a q chance that they would be safe and a $(1-q)$ chance that they would be risky. If you instead randomly chose two people from the population, there would be a $q \cdot q$ chance that they would both be safe and a $(1-q) \cdot (1-q)$ chance that they would both be risky. The chance that they would be a mixed pair is equal to the chance that they are not both safe nor both risky. That probability is $1 - q^2 - (1-q)^2$. After simplifying, this probability is equal to $2q(1-q)$.

$$= 2k.$$

or, simplifying by dividing by 2:

$$q^2 R_b + (1-q)^2 (p^2 + 2p(1-p)) R_b + 2q(1-q)[p R_b + (1-p) \underline{y}/2] = k.$$

The next step is to solve for the equilibrium gross interest rate R_b that makes the equation hold. The question is whether the R_b that emerges is lower than $k / [q + (1-q)p]$, which is the gross interest rate in the absence of group lending (found in chapter 2). After a bit more algebra (which we leave to readers as an exercise), we see that the break-even gross interest rate will indeed be lower than before. The bottom line is quite surprising: in principle, the group-lending contract can help lenders reduce interest rates—even where neither the bank nor the clients have information about who is safe or risky! In the process, adverse selection can be mitigated and a greater number of worthy borrowers can get access to credit.

The intuition is that risky borrowers, if lucky, can always repay their defaulting partners—whether safe or risky. But safe borrowers cannot repay for others due to the fact that their returns are lower and that all borrowers are protected by limited liability. Thus, defaults are *de facto* shouldered by risky borrowers only. Since risks are thereby passed on to risky borrowers specifically (rather than the average borrower), the bank is able to set interest rates that are low enough to win back the business of the safe borrowers. We end this section where we started, by reminding readers that the analysis only shows the *potential* for gains, and it draws on specific assumptions about the nature of risks and the role of limited liability. All the same, it is a striking example of the potential for group-lending contracts to make improvements—even in situations where it had been thought impossible.

4.3 Overcoming Moral Hazard

The previous section showed how group lending with joint responsibility can mitigate credit rationing due to adverse selection at the group formation stage. But as we pointed out in section 2.3, once loans have been granted, the bank may then face moral hazard

problems due to the difficulty of monitoring borrowers' actions. In this section we show how group lending with joint responsibility may circumvent moral hazard problems in lending, thereby further relaxing credit constraints. Here, we draw on the possibility that group members, who often live and work closely together, can impose social or economic sanctions on each other, possibilities that are impossible for an outside bank to impose.

4.3.1 *Ex Ante* Moral Hazard and the Role of Joint Responsibility

In important early work on the theory of group lending, Joseph Stiglitz (1990) set out an *ex ante* moral hazard approach to group lending. In a widely-cited article in the *World Bank Economic Review*, Stiglitz argues that the group-lending contract circumvents *ex ante* moral hazard by inducing borrowers to monitor each others' choice of projects and to inflict penalties upon borrowers who have chosen excessively risky projects. As Jean-Jacques Laffont and Patrick Rey (2003) argue, the fact that group members are affected by the actions--and inactions!--of other members means that they will take steps to punish anyone who puts in little effort and thus burdens the group with excessive risk.

To see how group-lending can address moral hazard, we go back to the *ex ante* moral hazard model of section 2.3.1, but with two borrowers that are linked by a group-lending contract. As in section 2.3.1, we assume that investment projects require a \$1 investment. A non-shirking borrower generates gross revenue y with certainty, whereas a shirking borrower generates gross revenue y with probability p and zero with probability $(1-p)$. Consider again a borrower's decision whether or not to put effort into her project. If R denotes the gross interest rate (interest plus principal) to be paid to the lender and c is the cost of effort, then a borrower's expected return if she puts in effort equals $(y - R) - c$, as before. Members of the group act to maximize group income, and anyone who deviates is punished with serious social sanctions.

In section 2.3.1, the borrower had the option to put in the requisite effort and getting net revenues of $(y - R) - c$. Or, alternatively, to take a gamble by shirking; in this second case, the borrower only succeeds p fraction of the time but does not have to bear the cost of effort. So, effort is only forthcoming if $(y - R) - c > p(y - R)$, which implies that the gross interest rate must be set so that $R < y - [c/(1-p)]$. Interest rates higher than

this level will encourage shirking. These inequalities are termed “incentive compatibility” constraints (or, simply, IC constraints), and they have a key role in understanding the function of contracts.

The group lending contract allows the lender to do better than above: interest rates can be raised higher without undermining good incentives. To see this, we consider a “group IC constraint”. We show that the maximum feasible interest rate that the bank can elicit from the group of borrowers without inducing default is higher because the IC constraint is “more relaxed” (i.e., easier to satisfy) than the individual IC constraint described in the paragraph above.

We again consider a two-person group. If both put in effort, they both pay back loans and both incur the costs of effort. Together, the return is $(2y - 2R) - 2c$. On the other hand, if they both shirk, they expect to be able to pay their full joint obligation $(2y - 2R)$ only p^2 fraction of the time. If the borrowers both shirk and one is lucky but not the other, the lucky one is responsible for the full repayment of both, leaving no surplus left over.

Thus, the group IC constraint under joint responsibility reflects the fact that positive rewards are only received when both projects succeed:

$$(2y - 2R) - 2c > p^2 (2y - 2R),$$

or equivalently $R < y - c/(1 - p^2)$. Since $p < 1$, it must be that $p^2 < p$, which means that $(1 - p^2) > (1 - p)$. Accordingly, the maximum achievable gross interest rate R under group lending with joint responsibility—i.e., $y - [c/(1 - p^2)]$ --is strictly larger than the maximum achievable interest rate in the absence of joint responsibility—i.e., $y - [c/(1 - p)]$.

The joint liability contract relies on the group’s ability to sanction individuals who try to shirk. In Stiglitz’s model, the sanctions are costless, but in subsequent work by others, monitoring and enforcement costs are derived as part of the decision framework. Given the contract, in principle both group members will never shirk, so it turns that out the sanctions are never actually used. In principle, all that is needed is the *threat* of their use.

4.3.2 *Ex post* Moral Hazard and the Role of Peer Monitoring

Now suppose that everybody works hard, so the kinds of concerns in section 4.3.1 are allayed. But now consider a problem that can occur after production has been completed and profits have been realized. The new concern is that borrowers may now be tempted to pocket the revenues without repaying the lender (i.e, to “take the money and run”). The problem then is that the bank cannot tell which borrowers truthfully cannot repay—versus those borrowers who are seeking to run away with their earnings.²⁰ To make the point sharp, we assume that, in the absence of peer monitoring, a borrower will default with certainty on her loan (whether or not she in fact has the resources to repay). Everything else equal, we saw in section 2.3.2 that this sort of *ex post* moral hazard eliminates the scope for lending as no bank will extend credit if it anticipates that the borrower will escape repayment.

Group lending with peer monitoring can, however, induce each group member to incur a monitoring cost k *ex post* to check the actual revenue realization of her peer. We assume that with this information, the partner can force the peer to repay. Let us assume that by incurring a cost k , a borrower can observe the actual revenue of her peer with probability q , and let d denote a social sanction that can be applied to a borrower who tries to divert due repayments. Then, if R denotes the gross interest rate set by the bank, a borrower will choose to repay if and only if:

$$y - R > y - q(d + R),$$

or equivalently:

$$R < [q / (1 - q)] d.$$

²⁰ Important papers on group lending with *ex post* moral hazard include those by Besley and Coate (1995) and Armendáriz de Aghion (1999). See also Rai and Sjöström (2004) and Jean-Jacques Laffont and Patrick Rey (2003) for theoretical approaches drawing from the economics of mechanism design, in which they derive optimal lending contracts in the case of moral hazard; these approaches show how the standard group lending contract can be improved upon depending on clients’ ability to make independent “side contracts between themselves.”

This in turn means that borrowers can contract any loans of size less than or equal to $[q / (1-q)] d$. In the absence of peer monitoring, we had $q = 0$ (zero chance of observing the borrower's actual revenue) and therefore no lending at all in equilibrium. Now, why do we have monitoring (implied by $q > 0$) in equilibrium? The answer to this question is somewhat similar to that developed by Banerjee et al. (see section 3.2.2 above). In their analysis of credit cooperatives, it was the insider's fear of losing her collateral w , which induced her to monitor her peer borrower. Here, it is the borrower's incentive to minimize the probability of suffering from joint responsibility which induces monitoring (provided the monitoring cost k is sufficiently small). Specifically, a borrower will choose to monitor her peer whenever the monitoring cost k is less than her expected gain qy from avoiding the need to assume responsibility for her peer's repayment. Thus, joint responsibility makes lending sustainable by inducing peer monitoring and overcoming enforcement problems associated with *ex post* moral hazard.

So, the group lending contract again does better than the traditional individual lending contract. But can the microlender do even better than that? Ashok Rai and Tomas Sjöström (2004) argue in an important theoretical contribution that the answer is: yes—and we return to the issue at the end of this chapter.

4.4 Evidence on Groups and Contracts

While the theories of group lending work on paper, how do they work in practice? Is the group lending mechanism in fact the key to the high loan repayment rates boasted by microlenders? Over the past few years empirical researchers have studied these questions, and they have arrived at a series of competing results. Some results support the theories presented above, while others point to tensions and constraints in the group lending approach.

Richard Montgomery (1996, p. 289) turns a critical eye to BRAC in Bangladesh, a Grameen Bank replicator (at least as far as its credit operations go). Montgomery argues that BRAC's implementation of group lending "can lead to forms of borrower discipline which are unnecessarily exclusionary, and which can contradict the broader (social) aims of solidarity group lending." This is an important reminder: the discussion so far has focused on ways that group lending can improve the bank's performance. We

have focused little on how the practice affects borrowers' lives, other than by assuming that improvements are made when group-lending improves access to credit for individuals lacking collateral. Montgomery's main concern is that group lending can create peer pressure that works against the poorest and most vulnerable members of the community. In attempting to keep repayment rates up, Montgomery contends, loan officers put sharp pressure on borrowers to repay, even when the borrowers faced difficulties beyond their control. He mentions stories of the "forced" acquisition of household utensils, livestock, and other assets of defaulting members. In one case, a woman's house was pulled down for failure to pay a housing loan.²¹ One response raised in chapter 6 involves providing insurance alongside credit, so that borrowers have a way to cope with major risks. Without such insurance, there is a legitimate question as to whether microfinance (whether implemented via group lending or via other methods) can make some borrowers more vulnerable than they had been.²² As we suggest in chapter 5, there may be other ways to get the upsides of group-lending without all of the downsides.

Montgomery also suggests that the "reality" of group-lending in Bangladesh is that the traditional five-person group ultimately plays a small role in ensuring repayment discipline. Instead it is the larger, village-level group that plays the key role. Montgomery writes with regard to this "village organization" (VO) that:

The VO leaders commonly treat overdue installments as a VO issue. If the individual continues to default on their installments, and the outstanding amount grows or the loan term expires, the VO leader and the group (VO) as a whole comes under pressure from the field staff. Rather than invoking the idea that four other members are jointly liable for the outstanding loan, field staff threaten to withdraw access to loans for VO members in general. The use of this sanction was freely admitted by the program staff in several of the five area offices in which field work was carried out; and it is because of the widespread use of this sanction that it

²¹ Examples are from Montgomery (1996), page 297.

²² Dale Adams, Emeritus Professor of the Ohio State Rural Finance Program and a microfinance skeptic, is fond of speaking of "microdebt" rather than "microcredit", signaling that loans carry burdens (as well as opportunities) for those who accept them.

is the VO, not the formal sub-groups within a VO, which becomes the joint-liability group in practice. In reality the 5-6 member joint-liability groups rarely exist, and especially in older VOs ordinary members cannot name the sub-group leaders stipulated in BRAC's formal blueprint of VO structure.²³

Similar stories have been told about Grameen Bank practices, and it happens often enough that one observer has called it “meeting day joint-liability.” The idea is that the loan officer is keenly aware of which borrowers in the larger, village-level group are finishing up their current loans and are about to request a next (often larger) loan. Those individuals are particularly susceptible to pressure to help with problem clients. Loan officers will thus be tempted to tell these soon-to-borrow-again customers that if help in dealing with the problem is not forthcoming, the anticipated loans may be delayed. To make the point sharper, it is not unheard of for the loan officer to refuse to leave the village until the books are completely squared. As Imran Matin (1997) has written, the staggered disbursal of loans helps to ensure that there is often someone in the larger group that is close to qualifying for a next loan—and thus particularly open to suasion.²⁴

The practice of “meeting day joint-liability” is not universal, and it is not necessarily a bad thing. Indeed, there is nothing sacred about the number five as the perfect group size. Elsewhere, solidarity groups stretch from 3 to 9 borrowers. And the village banking model used by FINCA, Freedom from Hunger, *ProMujer*, and others encompasses a single village-level group. While the adverse selection story of Ghatak (1999) hinges on the functioning of multiple groups within a village (so that borrowers can freely sort themselves into groups on the basis of risk), the moral hazard stories above do not depend critically on whether there is one group or more. Indeed, larger groups may be better able to deal with risks and less vulnerable to collusion.

Empirical researchers have tried to shine a bit of light on questions around the roles of groups, but getting clean results has not been easy. In the perfect world, empirical researchers would be able to directly compare situations under group-lending

²³ Montgomery (1996), pp. 296-297.

contracts with comparable situations under traditional banking contracts. The best test would involve a single lender who employs a range of contracts. But in practice most microlenders use just one main type of contract, leaving little variation with which to identify impacts. Where several different contracts are used, a different problem then emerges: why do some customers voluntarily choose one contract over another? Or why does a lender offer one version to some borrowers and a different version to others? Making comparisons thus opens up questions of whether “self-selection” or other aspects of the programs (e.g., management style, training policies, and loan officer behavior) are driving results. The best evidence would come from well-designed, deliberate experiments in which loan contracts are varied but everything else is kept the same. This can be achieved in a lab setting (see, e.g., Abbink, et al., 2002), but it has not yet been done in the field.

The lab experiment of Abbink, Irlenbusch, and Renner (2002), involves a game played in ten rounds. Participants are invited to the lab as part of a research experiment. In one case, the participants must register in groups of four, so that participants presumably sign up along with their friends. This case reflects the self-selection into groups at the heart of the Grameen Bank model. In the other cases, individuals register independently and are then placed into groups by the researchers, akin to the practice of the FINCA village bank in Ayachucho, Peru studied by Karlan (2002) in which FINCA forms groups from people who have independently come into the FINCA offices to join a list. The researchers aim to test the role of social ties by comparing outcomes of the self-selected groups relative to those of the groups put together by the researchers. The hypothesis is that stronger social ties should increase repayments.

The advantage of doing research in the lab is that the context can be kept exactly the same: the rules of the experiment, the way the participants are treated, and the eventual rewards received by the participants. Experimenters can then change just one aspect (the way groups are formed) and see what happens holding all else constant. The disadvantage, of course, is that the experiment proceeds in a deliberately artificial setting; for example, no mention was even made of “microfinance” for fear that it would trigger

²⁴ In the classic Grameen-style practice, typically two people in a five-person group get their loans first, then after a period the next two get loans, and finally after another wait, the last person gets their loan.

associations with certain kinds of behavior, actual loans are not made, and actual businesses are not operated. Moreover, the participants here were students at the University of Erfurt, Germany, not actual microfinance customers. And, on top of it, we have some reservations about how this particular experiment was designed.²⁵ All the same, experiments like this can help researchers understand the basic logic of contracts and develop hypotheses that can be pursued later in the field.

Each round of the Erfurt experiment begins with participants being given a “loan” that they must pay back. They are then told that either they have been lucky and their income is high, or else they are unlucky (less than 20 percent of the time) and their income is zero. With zero income, debts cannot be repaid by the unlucky individuals, so it falls to their group to repay all debts. If the group fails to fully repay for all, no one in the group can move forward to the next round of loans. The question is whether the lucky participants are willing to pitch in to help the unlucky members—and if those who are lucky will cheat and pretend that they have been unlucky (in order to avoid repaying). Groups that trust more and cheat less will do better, and the experiment is structured so that cooperation is unlikely without some basic trust between members—which suggests that the self-selected groups with pre-existing social ties should do better than the groups put together by the researchers. The researchers find, though, that, to the contrary, there is little difference in outcomes between the two groups; in fact, in some cases the self-selected groups do *worse* in terms of repayment rates, perhaps because shirking among friends is tolerated less than shirking among unconnected participants who can be expected to behave in a more self-centered way (the lower level of tolerance among friends destroys trust and, with it, reduces average repayment rates). Both types of groups, though, have slightly higher overall repayment rates than would be expected under traditional individual contracts. The experiment also reveals that women appear to

²⁵ One design feature that is questionable is that the participants are told that the experiment will stop after exactly ten rounds (if the group gets that far without defaulting). It is a well-known feature of finitely repeated games that in the tenth round strategic players will (in principle) act in a purely self-interest way, without concern for their fellow group members. If players are foresighted, they see that this will happen in the tenth round, and they will realize that they have nothing to lose by acting in a purely self-interested way in the ninth round too. So too for the eighth round, etc. Indeed, the whole thing should unravel and no cooperation should be possible from the first round forward. Given this, it is hard to know how to interpret the results of the Erfurt experiment. Clearly everything did not fall apart, and we discuss the results here as we think that this line of research has potential and the results are intriguing (even if the method is not fully satisfying).

be more reliable, that larger groups do worse, and that “dynamic incentives” play a major role in determining repayments. (We shall come back to gender-related issues in chapters 5 and 7, and we take up “dynamic incentives” in the next chapter.)

The finding that groups of strangers do as well as (and, in some cases, better than) groups of friends conflicts with arguments about the role of social capital and social sanctions in microfinance. But the finding has some support in theory,²⁶ and it is given support in the field by Bruce Wydick (1999) whose study of group-lending in Guatemala leads him to conclude that social ties *per se* have little impact on repayment rates: friends do not make more reliable group members than others. In fact, the participants he studies are sometimes softer on their friends, worsening average repayment rates (an interesting contrast to the experimental results in which friends appear to be tougher on each other, at least when dishonesty is perceived). Christian Ahlin and Robert Townsend (2003a) also find that proxies for strong social ties are associated with weaker repayment performance in evidence on group-lending in Thailand. Dean Karlan (2003), though, argues that social capital helps in Peru, and Mark Wenner (1995) finds that social cohesion is a positive force in groups in Costa Rica. Bruce Wydick too finds that social cohesion helps (as proxied by living close together or knowing each other prior to joining the microfinance group), even if *friendship* specifically creates tensions. Rafael Gomez and Eric Santor (2003) find that default is less likely if there is greater trust, social capital, and members have known each other before joining the groups.

In sum, the five empirical studies discussed above give a mixed picture. In thinking about the way forward, we pause here to take a more detailed look at the last of the studies. In any study based on survey data (based on actual borrowers and actual loans), the job for researchers is to convince readers that the comparisons of situations under different contracts are meaningful—that apples are not being compared to oranges. Gomez and Santor (2003) wrestle with comparability in their study of contracts used by two Canadian microlenders, Calmeadow Metrofund of Toronto and Calmeadow Nova Scotia of Halifax. Both programs make loans using individual-lending and group-lending methods. The individual loans tend to be larger (the median size is \$2700 versus \$1000

for group loans), but interest rates are identical at 12 percent per year plus a 6.5 percent up-front administration fee. As suspected, quite different types of people opt into group-lending over individual-lending. Group members are more likely to be female, Hispanic, and immigrant. Individual borrowers are more likely to be male, Canada-born, and of African descent; they are also more likely to have higher income and larger, older businesses, and to rely more on self-employment income. A simple comparison of performance across groups shows that group loans are more likely to be repaid (just over 20 percent of group-loan customers have defaulted on their loans versus just over 40 percent of individual-loan customers), but the comparison does not take into account other social and economic differences.

The approach taken by Gomez and Santor is to follow the “matching method” approach of Rosenbaum and Rubin (1983).²⁷ Using a sample of almost 1400 borrowers, the method involves first pooling all of the data and estimating the likelihood that a borrower will have a group loan (rather than a standard individual loan). Determinants include age, income, neighborhood, education level, and ethnicity. The estimates yield an index of the probability of taking a group loan, with the important feature that borrowers with the same level of the index also have similar observed characteristics. Reliable comparisons are thus achieved by only comparing borrowers with similar levels of the index. In principle, apples are compared to apples, and oranges to oranges. Using this method, Gomez and Santor find that borrowers under group contracts repay more often. The result, they argue, arises both because more reliable borrowers are more likely to choose group contracts and because, once in the group contracts, the borrowers work harder.

The estimation approach is simple and intuitive, but it rests on one vital assumption: that the choice of contract can be explained entirely by the variables in their equation (age, income, neighborhood, etc.). If there are important variables omitted from the equation (say, entrepreneurial ability or inherent riskiness), the method ceases to guarantee consistent estimates: riskier borrowers may more likely end up in individual

²⁶ As Christian Ahlin and Robert Townsend (2003a) note, the group lending models of Besley and Coate (1995) and Banerjee, Besley, and Guinnane (1994) predict that greater cooperation can undermine repayments as borrowers collude against the bank.

contracts, for example, and they may also be more likely to default. In this hypothetical case, the correlation between being in an individual-lending contract and having a worse outcome is not a product of behavior induced by the contract. Ideally, we would like to be able to investigate situations in which borrowers are sorted into contracts with some element of randomness—but such situations are rare.

Dean Karlan's (2003) study of the FINCA village bank in Ayacucho, Peru cleverly takes advantage of a quirk in the way that groups are formed that introduce some randomness into the process. There is only one main kind of contract (FINCA's village banking contract), but there is randomness in which group a borrower is placed. The FINCA contract involves groups of 30 women who meet weekly; each week, they receive new loans, pay installments on existing loans, and/or contribute to savings accounts. Unlike other models, the meeting is not held in the local neighborhood or village; instead meetings are held at the FINCA office in the town center. And, again unlike other models, it is FINCA that forms the groups in Ayacucho. FINCA broadcasts its intention to start village banks, and invites prospective borrowers to sign up. A list is posted on a wall, and once 30 names are listed, a group is formed. The next 30 people are formed into another group, and so forth. The staff find this the quickest way to form groups, and they hope to build social ties between strangers that will deliver independent benefits. In general, clients do not sign up as pre-formed groups, and most people do not know each other before FINCA puts them together. From an econometric standpoint, the fact that FINCA selects the groups minimizes biases due to unobserved characteristics.²⁷ Specifically, when comparing why one group had higher repayments than another group, concerns are alleviated that results will be biased due to peer selection based on unobserved strengths. Karlan's tests show that the composition of groups indeed looks similar to the general characteristics of the broader population—groups look like what you would expect from a random draw.

Karlan is most interested in the role of social capital—the links between clients that are foundations of trust and cooperation. Unlike real capital (cash, machines, and

²⁷ For more on the methods, see also Rajeev Dehajia and Sadek Wahba (1999) and P. Rosenbaum and Donald Rubin (1983). An easy-to-use estimator is available in the popular statistical package, Stata.

²⁸ Bias could creep back in when clients drop out of groups and are replaced by friends and neighbors of existing neighbors; Karlan thus limits analyses to initial members.

equipment), “social” capital cannot be observed and simply counted. To proxy for social capital, Karlan thus considers cultural similarity as indicated by language (Spanish only or Quechua—the most common indigenous language--only?), hair (braided, long, or short?), dress (indigenous *pollera* skirt or Western-style clothes?), and hat (indigenous-style hat or not?), as well as considering geographic proximity (percentage of group members living within a 10 minute walk of each other). These “social capital” measures correlate well with the level of social and business interactions and with who sits next to who at group meetings.

Do these measures of social capital make a difference to loan repayment rates? There are in fact two types of loan repayment rates. The first pertains to loans made by the central FINCA organization to the local group; these loans were all repaid on time during the period in question (1998-2000). The second pertains to loans made to group members from a pool of savings that was generated by the members themselves; here, repayment rates are much lower: around 20 percent. Karlan finds that larger scores on the measures of geographic proximity and cultural similarity predict lower default rates, a finding in line with the theory we sketched earlier in the chapter in which the threat of social sanctions aids repayment rates (and in line with, e.g., Stiglitz, 1990). Interestingly, while Karlan finds that default leads to drop-out from the program, the effect is attenuated for clients with more social capital. The finding suggests the possibility of beneficial risk-sharing: i.e., that clients who are forced to default due to circumstances beyond their control (as opposed to exhibiting moral hazard) are less likely to be forced to leave the program when the clients have strong social ties to the rest of the group.

Karlan’s results thus show that the group contract can harness local ties in ways that traditional lending contracts cannot. The limit of the results is that they can not nail down whether the improvements occur because of greater trust (and more effective use of social sanctions) as the stress on “social capital” in the paper’s title suggests—or, on the other hand, whether the improvements flow simply from the fact that people who are more similar and who live more closely may have an easier time monitoring each other (or perhaps both). The latter interpretation is consistent with Wydick (1999), who finds little support that stronger social ties help in group lending in Guatemala, but finds that repayment rates rise with variables that proxy for group members’ ability to monitor and

enforce group relationships (for example, repayments rise with knowledge of the weekly sales of fellow group members). The distinction between the two interpretations may not matter in practice (institutions may just be happy that the contracts help), but the unanswered questions point to next steps for research on contracts.²⁹

A different perspective on contracts is provided by the ambitious studies of Christian Ahlin and Robert Townsend (2003a and 2003b). They start with the theoretical models of group lending developed by Besley and Coate (1995), Banerjee, Besley, and Guinnane (1994), Ghatak (1999), and Stiglitz (1990). After putting the models into a comparable theoretical framework, Ahlin and Townsend take them to data, trying to determine which does a better job of explaining patterns in practice. Their data come from 262 joint liability groups of the Bank for Agriculture and Agricultural Cooperatives (BAAC) in Thailand, in addition to data on 2880 households from the same villages. Ahlin and Townsend do not seek to judge group lending versus alternative contracts. Rather, as with some of the other papers described here, their aim is to see what makes group lending work. Their answer is that there is no single universal answer. In the poorer regions of northeast Thailand, expected repayment rates increase when village social sanctions rise. But in the wealthier, central region, the extent of joint liability matters, and the higher are joint liability payments, the higher are default rates. Also, the greater is the extent of cooperation among group members (e.g., the more family members are in a group), the higher is default. These latter results suggest that too much social capital can be a bad thing when it fosters collusion against the bank.

Theory predicts that that a borrower's alternative options will influence outcomes as well. The more other good ways to borrow, the less a microfinance client will feel compelled to minimize chances of default—since defaulters can always fall back on their other options. Unlike the findings on social capital, there is fairly broad agreement here. Ahlin and Townsend (2003a), for example, find that everywhere in their sample, the better are a borrower's alternative options for borrowing, the worse are repayment rates. Mark Wenner's (1995) investigation of repayment rates in 25 village banks in Costa Rica affiliated with FINCA shows that delinquency rates are higher in better off towns where,

²⁹ Additional research by Karlan using experimental "trust games" with the same FINCA clients points to the beneficial role that social capital appears to be playing in Peru.

presumably, clients have more abundant outside options. The result is echoed by Manohar Sharma and Manfred Zeller (1996) in their study of three programs in Bangladesh (but not Grameen). An exception to these findings is found by Shahidur Khandker, et al. (1995, Table 7.2) in investigating the Grameen Bank and other Bangladesh banks. They find that both drop-out rates and repayment rates increase in better-developed villages (which tend to have better business opportunities).

A final empirical issue involves the role of diversity in groups. The theories that stress the positive roles of social capital and social sanctions suggest that less diverse groups will do better. Where collusion is a possibility, on the other hand, the opposite may hold: greater diversity may aid repayments by diminishing the chance for collusion. Loïc Sadoulet (2000) provides another reason that diversity can help: greater diversity means that group members' incomes are less likely to vary together, and thus group members' abilities to insure each other increases (i.e., there's a greater chance to provide mutual aid in times of need). Since insurance should help repayment rates, diversity helps.³⁰ And, if diversity helps, borrowers should try to form groups that are broad, which is exactly what Sadoulet and Seth Carpenter (2001) find in a study of groups in Guatemala. Ahlin and Townsend (2003a and 2003b), though, find that in Thailand it is positive correlations of income that, holding all else constant, appear to predict entry into group contracts.

Results from different parts of the world thus reveal different (sometimes opposing) relationships. Advancing understandings of group lending will thus entail better understanding the kinds of positive outcomes described in the first part of this chapter—along with understanding potentially negative scenarios as well.

4.5 Limits to Group Lending: Hidden Costs, Collusion, and Emerging Tensions

We started this chapter by reviewing the standard features of the group lending methodology introduced by the Grameen Bank in the 1970s. Theorists have been particularly interested in the ways that the model takes advantage of existing local

³⁰ Ghatak (1999) finds the opposite result: prospective borrowers will tend to seek out similar people to match with. If there are enough people to choose from, both Sadoulet and Ghatak could be right: safe borrowers seek to match with other safe borrowers (Ghatak) but, within the pool of safe borrowers, preference is placed on those with incomes that covary less with one's own income (Sadoulet).

information and social ties. But models that succeed in rural Bangladesh have not succeeded everywhere else. The evidence in the previous section shows a mix of results on what works and what does not.

Using social sanctions, in particular, has limitations. Typically, social sanctions involve excluding “problem” borrowers from privileged access to input supplies, from further trade credit, from social and religious events, or from day to day courtesies. Commercial banks hoping to move into the “microfinance niche” have particular difficulties invoking these kinds of mechanisms among their clients, but so do NGOs. For example, will the threat of social sanctions be credible in small village communities among very close friends and relatives? Or, at the other extreme, can social sanctions have teeth in urban environments where borrowers come and go and remain fairly anonymous to one another? Practitioners have thus had to tinker with contracts and redesign according to their contexts.

The tinkering and redesigning has had to address the costs inherent in group lending contracts, as well as the many advantages described above. The essence of group lending is to transfer responsibilities from bank staff to borrowers. Traditionally, loan officers select clients, monitor performance, and enforce contracts. Under group lending, borrowers share part of these burdens too. The gain for clients is that they obtain loans (and other financial services) at reasonable prices. But, given the choice, most clients would not opt to help start a bank and run it just in order to get loans. Jerry Ladman and Gonzalo Afcha (1990), for example, argue that in the case of the Small Farmer Credit Program (PCPA) in Bolivia, it was difficult to find potential borrowers to volunteer to lead their groups, and group leaders had to spend a great deal of time persuading borrowers to accept the group lending contract. In one village, group leaders had to put in four times as many hours in preparation before initial loan disbursement relative to the time needed under traditional individual lending procedures.³¹

Other concerns hinge on the group meetings that are at the core of group lending models. Attitudes are mixed. One complaint is that attending group meetings and monitoring group members can be costly, especially where houses are not close together.

In two of the three Chinese programs studied by Albert Park and Changqing Ren, for example, 8 percent of clients had to walk more than an hour to get to meetings. Overall, attending meetings and travel time took just over 100 minutes on average.³² In a survey of drop-outs from group lending programs in Uganda and Bangladesh, a Women's World Banking (2003) study found that 28 percent of drop-outs in Bangladesh left in part because of the frequency of meetings; this was so for 11 percent of former clients surveyed in Uganda. On the other hand, nearly all current clients of Women's World Banking affiliates in Uganda and Bangladesh report that they enjoy coming to meetings (Women's World Banking 2003, p. 5). In Uganda, the most-cited reason (65 percent) was that they liked the chance to share ideas and learn from each other; in Bangladesh, the most-cited reason (43 percent) was the social aspect of meetings.

A second issue relates to the fact that group lending works by transferring what are typically the bank's responsibilities onto the customers themselves. As we noted above, these responsibilities can carry hidden costs. Some borrowers may be tempted to think: I simply want a loan, why am I asked to help run the bank in return? But there is another aspect that goes beyond these kinds of costs. Group lending can bring added risks for borrowers, and if borrowers are risk averse, those risks can weigh heavily. The risk is embedded in the contract: a borrower is now not just at risk of defaulting on her own but she also faces the risk that her partners will default also. If monitoring and enforcing contracts is costless--as assumed in Stiglitz's (1990) treatment--borrowers can address moral hazard effectively and the risks are minimized. This is the great hope of the group lending contract. But, as noted above, monitoring is not costless, even for individuals living in close proximity. Typically, then, monitoring will be imperfect, opening way for moral hazard to enter back into the picture. But under the group lending contract, it is now the group that is exposed to the risk, not the bank. The threat of social sanctions can help, as we describe above, but in practice they are applied only imperfectly too.

³¹ One colleague who read this passage in a draft version of the chapter suggested that part of the problem might simply have been that the particular product was poorly designed—not that the group-lending concept was necessarily flawed.

³² The data are from a preliminary analysis of a survey of 3 programs completed by Albert Park and Ren Changqing.

The sum, as argued by Malgosia Madajewicz (2003a) in an important theoretical analysis, is that the benefits of group lending—which have been detailed in the first part of this chapter—are counter-balanced by costs. Those costs emerge when borrowers are risk averse and monitoring is costly. Moreover, the costs grow as the scale of lending grows, since the financial implications of default rise with the size of loans. Madajewicz argues that loan sizes are limited by what the group can jointly guarantee, so clients with growing businesses or those who get well ahead of their peers in scale may find that the group contract bogs everyone down. Below a certain scale, group lending dominates individual lending. But her analysis shows that at a certain size of business, individual lending will be preferred by customers. In an investigation of data from Bangladesh, Madajewicz (2003b) estimates that the switch toward the greater net benefits of individual loans already happens for households holding 1.25 acres. Such households would not be considered to be “functionally landless”, but they are mainly poor nonetheless.³³

One implication is that better-off clients tend to seek individual loans as they move forward, pushing Bolivia’s BancoSol and the Grameen Bank, both group lending pioneers, to introduce new individual-lending contracts for successful clients. A related issue is that some clients simply prefer not having to be obligated to others. As the Women’s World Banking (2003, p. 3) study reports:

This issue was tested further through the question: “Which do you prefer, to have the security that the group will help you out when you are not able to pay back each week, or to assume complete responsibility for your own loan and not having to pay for someone else’s loan?” Most customers of both institutions indicated a desire to be independent and to forsake the security of the group. In Bangladesh, 76% of the affiliate’s current borrowers and 82% of dropouts answered that they would want to assume total responsibility for their own loan. In Uganda, 87% of the affiliate’s

³³ Jonathan Conning (2000) also provides an analysis of implications of costly monitoring by borrowers, describing when and how group lending can dominate individual lending--and vice versa.

current borrowers and 84% of dropouts expressed a similar desire for independence.

A third issue is that under some conditions, borrowers in group lending contracts may collude against the bank and undermine the bank's ability to harness “social collateral”.³⁴ As we saw in the section above, stronger social ties within a group can push up repayment rates in some places, while, in others, social ties increase the likelihood of default.

Jean-Jacques Laffont and Patrick Rey (2003) take up these tensions from a theoretical perspective and come to a somewhat optimistic conclusion. In their investigation of moral hazard and group-lending, close ties and information-sharing among borrowers open the way for contracts that improve on traditional individual-lending contracts. But, on the other hand, the scope for collusion against the lender increases when borrowers share knowledge and social ties. If borrowers do *not* collude, Laffont and Rey show (in a stylized model) that group-lending contracts are superior to individual-lending contracts (because the contracts take advantage of borrowers' knowledge and social ties—as described at the start of the chapter). But even better contracts exist in principle. These include using yardstick competition (judging one member's performance according to the performances of others) and information revelation mechanisms (such as cross-reporting arrangements).

But what if borrowers collude? In that case, Laffont and Rey show that group lending is superior to these alternative mechanisms. The contract delivers outcomes that are not as good as could be obtained if the lender had full information on borrowers, but it beats alternatives. Their bottom line is that having more information (either on the part of borrowers or on the part of the lender directly) leads to contracts that improve on standard individual-lending contracts, even when borrowers collude against the lender.

A final issue is whether the group lending contract is more efficient than alternatives *even when it is successful on its own terms*. At the end of section 4.3.2, we raised the question: even if the group lending contract does better than the traditional

³⁴ Collusion is also an important possibility considered in the theoretical studies of Besley and Coate (1995) and Laffont and N'Guessan (2000).

individual lending contract, can the microlender do even better than that? Ashok Rai and Tomas Sjöström (2004) argue that the answer is: yes (as do, in somewhat different contexts, Jean-Jacques Laffont and Patrick Rey, 2003). The criticism of the group lending contract as we see it on paper (and as we have described it above) is that punishments are too harsh. For example, in the widely-replicated original Grameen Bank contract with five-person groups, when one borrower defaults, all four others are cut-off from future lending too. It is that threat that drives the “peer monitoring”, “peer selection”, and “peer enforcement” mechanisms above. But what if the defaulter got into trouble because her husband fell ill? Or her cow died? Or prices dropped for the goods she sells? What if the problem occurred despite good monitoring, selection, and enforcement?

Rai and Sjöström’s particular criticism does not hinge on the morality of the situation, but rather on its efficiency (in the sense used in chapter 2 above); in the dispassionate language of economics, the punishment implies a “deadweight” loss. They argue that by using a system of cross-reports (see the end of the next chapter for more), punishments need not be levied so bluntly. Rai and Sjöström argue that rather than writing a contract and passively following the rules, the bank (and borrowers) can take active steps to gather more information when crises emerge. Their idea of cross-reports is to elicit truthful information about what has happened (e.g., was default due to shirking or to a deeper problem?). This information can be elicited by the microlender by soliciting reports from the problem borrower and her neighbors and showing leniency when all of the independent reports agree with each other. Some overly harsh punishments can thus be avoided. The proposed system of cross-reports is just one way to improve on contracts, and it works well on paper in a specific theoretical context. With modification it might work in practice too, but, even without cross-reports, microlenders are taking steps to address the inefficiencies.

We take the Rai and Sjöström criticism seriously, and microlenders act as if they do as well. Our first-hand observations in Latin America and Asia indicate that group contracts are seldom enforced exactly as they should be on paper. When asked, loan officers respond that they see no reason to automatically punish everyone for the problem of a single person. Instead, typically loan officers typically spend a great deal of time

investigating and managing “problem” cases. In doing so, staff call on defaulters’ neighbors for advice and information (in the spirit, loosely, of cross-reporting). And, once the problem has been investigated (and if the defaulter’s peers are found to be relatively blameless), microlenders’ staff try to get as much of the problem loan repaid as possible and then (if called for) drop just the one defaulter from the group and replace her with an alternative borrower. This is a natural route to improving efficiency (and equity), even as it undermines the strict reading of group lending contracts.

In a notable break, Grameen Bank’s “Grameen Bank II” proposal recognizes the tension between what works on paper and what happens in practice by formally introducing mechanisms through which loan officers can address the problems of individual borrowers without invoking punishments for the entire group (Yunus, 2002). The heart of Grameen Bank II is comprised of two types of loans. Borrowers first start with a Basic Loan (in Bangla, this is an “Easy Loan”). The new system allows loans of any duration—from three months to three years—and allows for installments to be smaller in some seasons and larger in others. The weekly repayment practice remains, however. Then, if borrowers get into trouble, they will be offered a Flexible Loan (with the penalty of a sharp drop in their loan size limit). The Flexible Loan has easier terms spread over a longer period, and it allows the borrower to get back on track, eventually returning to Basic Loan status. Half of the loan is provisioned for at the time of switching status to the Flexible Loan. Only when the customer fails to repay the Flexible Loans are they expelled and the loan is fully written off as bad debt. Some see this proposal as a major departure from group lending by the pioneer of the group lending contract.³⁵

4.6 Summary and Conclusions

This chapter took up one of the major innovations of the microfinance movement--group lending. From the lender’s perspective, the beauty of the contract is that it’s a way to

³⁵ As of the middle of 2004, the new flexibility provided by Grameen Bank II has not been implemented widely in practice, perhaps because loan officers are slowly getting used to the idea of increased flexibility. As chapter 6 describes, Grameen Bank II also brings new savings methods—which may be as important a break for the bank as are the proposed new lending methods.

Chapter 5**Beyond Group Lending**

Introduction

5.1 Creating dynamic incentives

5.1.1 Threatening to stop lending

5.1.2 Progressive lending

5.1.3 Competition and incentives

5.2 Matching household cash flows with repayment schedules

5.3 Complementary incentive mechanisms

5.3.1 Flexible approaches to collateral

5.3.2 Financial collateral

5.3.3 Making repayments public

5.3.4 Targeting women

5.3.5 Information-gathering by bank staff

5.3.6 Cross-reporting

5.4 Summary and conclusions

Exercises

Introduction

The “discovery” of group lending opened up possibilities for microfinance.³⁶ It is by far the most celebrated microfinance innovation, and with good reason. Group lending showed how unconventional contracts can work where “tried and true” banking practices failed again and again, and the shift in understandings led to other new ideas that borrowed as much from traditional moneylenders as from modern banking practices. Today, group lending is just one element that makes microfinance different from conventional banking.

Many of these other new ideas are also used by institutions practicing group lending. But the mechanisms are not intrinsically linked, and institutions are increasingly finding that they can pick and choose different elements. A case in point is “progressive

³⁶ This chapter draws on Beatriz Armendáriz de Aghion and Jonathan Morduch (2000).

lending”, which is a staple of the “classic” Grameen Bank model but which does not hinge on group lending *per se*. Progressive lending refers to the practice of promising larger and larger loans for groups and individuals in good standing. Other innovations already present in the classic Grameen model include repayment schedules with weekly or monthly installments, public repayments, and the targeting of women. In addition, microlenders have adopted more flexible attitudes to collateral. The emerging new contracts do not necessarily involve groups, and they have been especially helpful in areas with low population densities or highly diverse populations--and in situations where more established clients seek greater flexibility.

Bangladesh’s ASA, with its obsession with maximal efficiency, has weakened joint liability in its lending approach, for example, and even the Grameen Bank has proposed to soften joint liability in “Grameen Bank II”, which (on paper at least) allows problem loans to be routinely renegotiated without invoking group pressure.³⁷ In Bolivia, BancoSol has moved a large share of its portfolio out of “solidarity group” contracts into individual contracts. “Solidarity group” contracts are still used for small loans (from \$50 to \$2000) that are offered to less-established clients, but individual contracts (up to \$30,000 but averaging \$1000) are the norm for established clients.³⁸ Bank Rakyat Indonesia, another microfinance leader, eschewed group loans from the start, and it is joined on that path by urban microlenders in Latin America and Eastern Europe.

Table 5.1 provides comparative data for the 147 programs surveyed in the *MicroBanking Bulletin*. Of these “top performers”, 73 are individual-lenders and the rest either lend through Grameen-type groups of 3 to 9 borrowers, or through the larger groups associated with the village banking approach.³⁹ Relative to lenders using group-lending methodologies, microlenders focusing on individuals tend to (a) be smaller and serve better-off clients, as reflected by average loan size; (b) be more self-reliant as proxied by the percentage of their financial costs covered—102 per cent relative to just 89

³⁷ Renegotiation occurs by transferring problem borrowers from standard “basic” loans to “flexi-loans” with longer terms and smaller installments, but by early 2004 the practice was not yet common at Grameen.

³⁸ Data are from www.bancosol.com.bo/en/historia.e.html, as posted in May 2003.

³⁹ Village banks operate by placing everyone in the village into one large group with mutual responsibility. Group meetings are often used for training sessions as well as financial matters. For more on village banking, see www.villagebanking.org and Karlan (2003).

Table 5.1
Financial Performance Comparisons by Lending Methodology

	Individual	Solidarity Groups	Village Banks
			10 or more borrowers per group
Definition	1 borrower	Groups of 3 to 9 borrowers	
Observations	73	47	27
Scale			
Number of borrowers	9,610	47,884	16,163
Average loan size	\$973	\$371	\$136
Outreach			
Average loan size / GNP per capita (%)	88	46	20
Fraction female (%)	46	73	89
Financial performance			
Return on assets (%)	-1.2	-4.1	-7.2
Return on equity (%)	0.6	-12.5	-10.7
Operational self- sufficiency ratio (%)	121	102	107
Financial self-sufficiency ratio (%)	102	89	89
Portfolio yield (real, %)	21	30	49
Portfolio at risk > 90 days	3.7	3.6	3.5
Efficiency			
Operating expense / loan portfolio (%)	20	37	61
Cost per borrower (US\$)	155	93	62
Number of borrowers / total staff	147	155	160
Number of borrowers / loan officer	508	356	309

Source: Microbanking Bulletin, Issue 8, October 2002, Table aA (www.mixmbb.org).

for group-lending microentrepreneurs; (c) serve a smaller population of women clients—on average 46 percent of the clients of individual-microlenders are women versus 73 percent for group-lenders and 89 percent for village banks; and (d) charge lower interest rates and fees as reflected in the real portfolio yield: 49 percent for village banks, 30 percent for group-lenders, and 21 percent for individual lenders. On this latter point, however, it should be noted that village banks and group-lenders also have considerably higher expenses relative to loan size. While individual-lenders devote 20 cents of each dollar lent to operational costs, group-lenders must devote 37 cents, and village banks 61 cents.

The bottom line is that the group-lenders and village banks tend to serve poorer clients and have higher costs relative to loan size. As microlenders have matured and diversified, their push to serve better-off clients and reduce costs has opened the door to individual-lending approaches. But individual-lending approaches also have appeal in sparsely-populated regions, areas with heterogeneous populations, and areas marked by social divisions, where peer monitoring costs are high and social punishments for non-compliance more difficult to implement. Individual-lending approaches may thus be critical in serving some very poor areas as well.⁴⁰

Below, we first discuss the recent trend towards bilateral contracting and its emphasis on dynamic incentives via progressive lending techniques. By isolating these lending methods, we aim to shed light on alternative variants of the classic group lending model as described in the previous chapter. This in turn can open the scope for microfinance to expand to areas where barriers were thought to be too high. We also discuss the use of collateral requirements and the replacement of joint liability clauses with public repayments as a simpler way of maintaining peer pressure, and discuss how these innovations are re-shaping the microfinance landscape. At the chapter's end, we revisit the group-lending methodology and the challenges it faces as the microfinance industry moves forward.

⁴⁰ The work of *SafeSave* in the slums of Dhaka is one example.

5.1 Creating Dynamic Incentives

Even without recourse to peer monitoring, collateral, or social sanctions, microlenders can give incentives to borrowers by threatening to exclude defaulting borrowers from future access to loans. In this way, microlenders have a weapon that was unavailable to failed state-run banks of the past. Those banks were often pressured to extend loans based on political exigencies and could not be counted on to supply a steady flow of financing to small entrepreneurs. One striking finding about India's troubled Integrated Rural Development Program, for example, was that only 11 percent of all IRDP borrowers borrowed more than once (Pulley, 1989). If you suspect that you'll only ever take one loan from an institution, the chances that you'll go to great lengths to repay it fall sharply, and it is not surprising that IRDP's repayment rates fell below 50 percent over time.⁴¹ Microlenders ratchet up incentives even further by giving borrowers in good-standing access to ever-larger loans, creating the promise of turning start-up businesses into steady enterprises.

In this section we present a simple model of debt without collateral to analyze how bilateral contracts work. We then explore the role of "progressive lending" as an additional tool. While a thick, competitive microfinance market ought to be a microfinance dream, we describe cases in which competition has undermined dynamic incentives in microfinance (and led to microfinance crises in Bolivia and Bangladesh). And we describe why credit bureaus are needed to improve matters.

5.1.1 Threatening to stop lending

Nearly all moneylenders surveyed by Irfan Aleem (1993) rely principally on two devices for eliciting debt-repayments from their clients: developing repeated relationships with the borrowers and making sure that existing borrowers do not contract new loans with other lenders.⁴² The two devices make the threat of not refinancing a customer a powerful weapon. We begin by analyzing the theory of these "non-refinancing threats".

⁴¹ A credit agency can help address this problem, such that banks can investigate credit histories of prospective clients, but we know of no such agencies serving microfinance populations.

⁴² See, Aleem (1993) Table 7-2, p. 137.

Suppose that monitoring costs are very high so that lenders cannot induce repayments via peer groups.⁴³ As before, we maintain the assumption that borrowers do not have collateral. Moreover, we assume for the moment that social sanctions cannot be used as a way of putting pressure on borrowers to fulfill their contractual obligations. Starting from these basic assumptions, we present a stripped-down version of a model by Patrick Bolton and David Scharfstein (1990). The model is inspired by the “sovereign debt” problem of the 1980s which involved lending relationships between “foreign” commercial banks and sovereign nations.⁴⁴

Assume that there are two periods of production and an investment project that requires \$1. At the end of each period the borrower can generate a gross return $y > \$1$, calculated before repayment of the loan with interest, provided that her current project is financed by the bank. At the repayment stage, however, the borrower may decide to default strategically by simply not repaying the loan. In order to deter the borrower from “taking the money and running”, the bank can extend a second period loan contingent upon full repayment of the first-period obligations. The borrower’s penalty for default is thus that she will not be able to invest in the second period. Is this threat enough to elicit payment from the borrower?

Suppose that the borrower decides to default. Her expected payoff in this case will be $y + \delta v y$, where δ is the borrower’s discount factor, and v is the probability of being refinanced by the bank despite having defaulted. The discount factor captures the fact that most people weigh payoffs in the future less than payoffs today. To fix ideas, we assume for simplicity that the borrower needs the bank in order to finance a second-period investment, even in the case where he pockets the entire first-period return realization.⁴⁵

⁴³ See Armendáriz de Aghion (1999) for a framework where peer monitoring costs are explicitly taken into account. Specifically, if peer monitoring is exceedingly costly, individual (i.e., bilateral lender-borrower) contracts are shown to dominate over group-lending contracts.

⁴⁴ In the sovereign debt case, there is no international court where foreign creditors can enforce claims on a country, so there can be no use of collateral either. See Jeremy Bulow and Kenneth Rogoff (1989a and 1989b).

⁴⁵ This turns out to be an important assumption. If the borrower could default and hold onto enough principal to easily finance future business operations, the threat of non-refinancing would be considerably weakened. See Bond and Krishnamurty (2001) for a discussion of assumptions needed for threats of non-refinancing to have teeth when this is the case.

Now suppose that, having done well with her investment, the borrower decides to repay. In this case, her payoff will be $y - R + \delta y$, where R is the gross interest rate payable to the bank (principal plus interest). Here, the bank refinances the borrower's second-period investment for sure, setting $v = 1$. As we argue below, this is an equilibrium strategy.

Clearly, because of the *finite* number of periods (two in this case), the borrower has no incentive to repay at the end of the second period. So if she repaid in period 1 and is refinanced with certainty, her net expected payoff in period 2--evaluated in period 1--is equal to δy .⁴⁶ Similarly, if she defaulted in period 1 and is consequently refinanced with probability $v < 1$, her expected payoff in period 2 (evaluated as of period 1), is equal to $v\delta y$.

Now moving back to period 1, it is easy to see that the borrower will decide to meet her first-period debt obligation if and only if $y + v\delta y \leq y - R + \delta y$. This is an “incentive compatibility” (IC) constraint in the jargon of contract theory, a concept we used in section 4.3.1. As we saw in chapter 4, the constraint determines the largest feasible interest rate that the bank can elicit from the group of borrowers without inducing default. The constraint says that the bank should make sure that the borrower's net present payoff is at least as large when she does not default as when she does. And the obvious way that the bank can do this is by setting an interest rate that is not “too high”.

From this, we use the incentive compatibility constraint to derive the maximum gross interest rate R that the bank can elicit from the borrower at the end of the first period is equal to $\delta y (1 - v)$. The expression is maximized by setting $v = 0$ for defaulters, that is by fully denying access to future refinancing.⁴⁷ Thus, the maximum repayment that the bank can request after the first period is simply $R = \delta y$, which is the borrower's opportunity cost of defaulting strategically. It will never pay for the borrower to repay

⁴⁶ The model rests on the assumption that the bank can credibly commit to provide a second-period loan, even though it anticipates this new loan will be defaulted upon, which may seem unrealistic. However, it will all depend on the interest rate that the bank charges, which in this set-up will be endogenously determined. Note that the probability of default will be substantially reduced in an infinite horizon model. In particular, we know by the “folk theorem” of game theory, that if the discount factor, δ , is large enough, strategic defaults will never be observed in equilibrium. See, for example, Drew Fudenberg and Eric Maskin (1986).

⁴⁷ This expression reduces to $\delta y (j - v) < \delta y (1 - v)$ if a non-defaulting borrower is refinanced only with probability $j < 1$.

more than δy in this set-up.⁴⁸ If, say, the borrower's discount factor is 0.90 and the borrower's gross return is 160 percent, the maximum feasible gross interest rate is 144 percent (or a maximum net interest rate of 44 percent). When operating costs are high, the constraint may well bind. And banks will be even more constrained when borrowers have low discount factors or perceive a relatively high chance of getting refinanced despite default. As described below, competition without coordination—say, without a credit bureau that keeps tabs on defaulters from other banks—may serve in effect to push the effective refinancing probability v above 0.

This simple framework also suggests why maintaining the appearance of stability is important for lenders. If borrowers begin to think that the bank could go under in future periods, they are more likely to default now, since it is not clear whether there will be a future flow of loans. Whether based in fact or not, such speculation can trigger a “debtor run” which becomes a self-fulfilling prophesy. Philip Bond and Ashok Rai (2002), for example, describe a ballooning of defaults faced by Childreach, a microlender in Ecuador, in response to rumors that the organization faced a looming financial crisis.

5.1.2. Progressive Lending

Table 5.2 shows that the Grameen Bank not only provides a continuing series of loans, but that they quickly increase in size. The table shows data for three borrowers randomly chosen from a 1991-92 sample of thirty Grameen Bank borrowers who each had had six loans to date. The first borrower doubled the value of her loan by the fifth loan; the second borrower had doubled the size by the fourth loan. The final column shows average loan sizes for the entire sample, growing from Tk. 2124 for first loans (\$57 in 1991) to Tk. 4983 (\$ 135) for sixth loans. For the lender, progressive lending cuts average costs since servicing a Tk. 2000 loan is not twice as expensive as servicing a Tk. 4000 loan. Progressive lending also enables the lender to “test” borrowers with small loans at the start in order to screen out the worst prospects before expanding the loan scale (see Parikshit Ghosh and Debraj Ray, 1997).

⁴⁸ Note that the maximum enforceable repayment $R = \delta y$ satisfies the “individual rationality constraint” of the borrower, namely: $y - R + \delta y \geq 0$. This constraint states that an individual borrower must find it profitable to enter into a contractual obligation with the bank – otherwise, they refuse to borrow in the first place.

Table 5.2
Loan Size Increases (Taka)

Grameen Bank, Bangladesh				
Loan number	Borrower A	Borrower B	Borrower C	Full Sample Average
1	2000	2000	3500	2124
2	2500	2500	4000	2897
3	3000	3000	3000	3656
4	3500	4000	4000	4182
5	4000	4000	5000	4736
6	4000	5000	4000	4983

Source: Authors' calculations from Bangladesh Institute of Development Studies-World Bank 1991-92 Survey. Data are in current taka (in 1991, \$1 = Tk. 37; in 1986, \$1 = Tk. 30). The final column averages loan sizes over the full sample of Grameen Bank borrowers in the data set (excluding loans used for land/building), and sample sizes diminish with loan number; starting from the first row downward, there are 319, 286, 250, 168, 89, and 30 observations.

From the analysis above, progressive lending has a third, important role with regard to incentives. Microlenders can elicit even larger repayments by offering loans of larger size to borrowers that repay their debts. Specifically, progressive lending schemes increase the opportunity cost of non-repayment and thereby discourage strategic default even further. To see this, suppose that the bank decides to increase the size of its short-period loans by a factor $\lambda > 1$ between period 1 and period 2, and that the production technology has constant returns to scale. The opportunity cost of strategic default will then increase by the same factor between the two periods. In particular, by not repaying the gross interest rate R , the borrower now suffers a loss $\lambda \delta y > \delta y$. This in turn relaxes the incentive compatibility constraint, and the bank can now achieve a maximum interest rate equal to $R' = \lambda \delta y > R = \delta y$. Interest rates can be raised while keeping the borrowers happy.⁴⁹

Note though that, as before, the analysis rests on an assumption that may not be fully tenable—that if a borrower defaults in the first period, she nonetheless needs a loan

⁴⁹ One more step is actually needed. It has to be checked that the interest rate satisfies the borrower's "individual rationality" constraint—i.e., is it worth it for the borrower to borrow at that rate?

to be able to invest in the second period. In principle, borrowers may be able to keep at least part of the principal from the first period and use that to invest in the second. If so, dynamic incentives are harder to maintain; in this case, borrowers can expect a return of $y - R' + \lambda \delta y$ if they pay their first-period debt. If they do not, their return is $y(1 - \varphi) + \varphi \delta y$, where $\varphi < 1$ is the fraction of the first-period gross return that is invested in the second period. Suppose that, if the borrower defaults, their choice is to hold back a fraction $\varphi = R/y$. That is, from first-period gross returns, they save for the next period exactly the amount that they would have paid to the bank (had they chosen to repay the loan with interest). In this case, the household will not default if $\lambda > \varphi$. Since loan sizes are growing ($\lambda > 1$) and since not all of the loan is retained ($\varphi < 1$), this inequality must hold: the borrower will not default. But incentives will erode if loans shrink in size, or if the borrower can scale-up their own resources faster than the bank can (for more on this, see Bond and Krishnamurty, 2001).

This leads to another observation. A borrower who is disposed to strategically default will wait until loan sizes have grown substantially before ultimately choosing to renege on the loan contract. The lender (if also acting strategically) will in turn carefully determine loan schedules in order to minimize default. More specifically, consider a multi-period debt relationship between the lender and the borrower. If the growth factor λ is large (i.e., loans increase in size very quickly), the borrower has incentives to default earlier than they would when compared to a “flatter” loan schedule. The incentive problem imposes an upper bound on the desirable growth rate of loan size over time. On the other hand, reputation considerations on the borrower’s side (which are absent from the simple model spelled out above) should mitigate this effect by reducing the borrower’s incentive to default (see for example, Joel Sobel, 2002).

5.1.3 Competition and Incentives

Economists usually view competition as a good thing, and most theoretical models assume that there is perfect competition. So far, we have assumed in fact that microlenders are either perfectly competitive or that they simply wish to break even. But in this section we argue that strong competition can undermine dynamic incentives. If a microlender is a monopolist, their threat to cut access to defaulters has greatest bite since

they are the only source of credit. Dynamic incentives can weaken when alternative lenders enter the market (assuming that the defaulter has a chance to borrow from them instead). Not only that, but competition can weaken reputation effects.⁵⁰

Problems with competition have emerged most notably in two countries where microfinance was first to take hold: Bolivia and Bangladesh. Craig McIntosh and Bruce Wydick (2002) also report on problems of competition in Uganda, Kenya, Guatemala, El Salvador, and Nicaragua. The Bolivian crisis took root when aggressive providers of consumer credit entered the market. In this case, the new entrants were outsiders, notably Acceso FFP, a large Chilean finance company.⁵¹ Acceso came in with streamlined operations and over 1000 highly motivated employees (most of whose pay came in the form of incentives rather than base salary). Within three years, Acceso had 90,000 loans outstanding, a level that BancoSol had not reached in its twelve-year history. In 1999, the worst year of the crisis, BancoSol lost 11 percent of its clients, and loan overdue rates for regulated microlenders fell from 2.4 percent at the end of 1997 to 8.4 percent by mid-1999. BancoSol saw its return on equity fall from 29 percent in 1998 to 9 percent in 1999.⁵²

The immediate problem with competition in Bolivia was borrowers taking multiple loans simultaneously from different lenders. The borrowers then became over-indebted, paying one lender's installments by taking a loan from another, leading to a spiral of debt and, too often, financial peril. Carmen Velasco, co-executive director of *Pro Mujer*, tells of visiting a client in Cochabamba who had loans from two different institutions and was sinking under the weight. The client's husband reported a proposed solution—the next day they planned to seek a loan from BancoSol to help pay off the first two loans!⁵³ While our discussion above concerns problems that occur when borrowers can turn from one lender to another in sequence (rather than simultaneously), the root of the problem is similar. As long as borrowers believe that they have multiple options, no single lender will have the power to clamp down and maintain full discipline.

⁵⁰ See Hoff and Stiglitz (1998).

⁵¹ The Bolivian experience is well-told by Elisabeth Rhyne (2001, chapter 7), from which this account is taken.

⁵² Data on number of clients are from Rhyne (2001), p. 142. Data on overdue rates are from pp. 148-9, and data on BancoSol's return on equity are from p. 149.

⁵³ The story is related in Rhyne (2001) p. 145.

Pro Mujer declared that clients holding loans from other banks were henceforth ineligible to borrow, but following up on all financial activities of clients and their families is costly in practice. The general situation in Bolivia improved, though, as regulators tightened rules, the Chilean financiers retreated, and the early microfinance providers like BancoSol and *Pro Mujer* took extra steps to keep their clients satisfied. Looking forward, the most effective solution would be a credit bureau which keeps track of the credit histories of all borrowers across the nation.

The Bolivian crisis occurred around the same time as the crisis in Bangladesh. The middle and late 1990's saw the explosive growth of the Grameen Bank, ASA, BRAC and Proshika. While it is impossible to accurately count (because borrowers from a given institution also borrowed from others), around ten million new microfinance clients signed on over the decade. The main microfinance providers had agreements not to work with the same clients, but that did not prevent a crisis of simultaneous borrowing along the lines of what occurred in Bolivia. In Bangladesh the problem has been dubbed "overlapping", and Imran Matin and Iftekhar Chaudhury (2001) report that by the end of the decade, there was more than one microlender operating in 95 percent of 80 villages surveyed by researchers at the Bangladesh Institute for Development Studies. Matin (n.d.) reports on a BIDS study that estimates that 15 percent of all borrowers took loans from more than one institution. The result, coupled with a broader pattern of lending more than clients could fully absorb, was a repayment crisis that took Grameen Bank's reported repayment rates from above 98 percent to below 90 percent, with greater difficulties in densely-served areas like Tangail district.⁵⁴

The lesson from these experiences is not that monopolies should be protected. In both Bangladesh and Bolivia, competition has brought a healthy round of general rethinking that would have not otherwise happened so soon.⁵⁵ The chief lesson is instead that cooperative behavior among microlenders can help to mitigate the problem. Programs would be aided by the creation of credit bureaus to better share information on credit access and performance history of borrowers. Having credit bureaus enables

⁵⁴ Grameen Bank *Annual Report* 1995 and *Annual Report* 2000. Imran Matin (1997) tells a richly-observed story of how "overlapping" led to severe difficulties in villages in Tangail.

lenders to address over-indebtedness and to make borrowers face the consequences of strategic defaults (which is not to say that it would be simple to set up credit bureaus in countries like Bangladesh, where there is no system of social security numbers or national ID numbers). No-one can force microlenders onto joining a credit bureau, but the argument in favor of fierce competition cannot be defended without the presence of an adequate regulatory framework.⁵⁶

In Bolivia, regulated financial intermediaries like BancoSol are required by law to report both names and national identification card numbers of delinquent borrowers to the Superintendency of Banks and Financial Institutions (Gonzalez-Vega, et al 1997). In return, all regulated financial intermediaries are allowed to view the information provided by the others, and informal arrangements are used to share information with non-regulated microlenders. These measures strengthen dynamic incentives, but lenders must fend for themselves in dealing with “overlapping” clients.

5.2 Frequent Repayment Installments

One important issue that has so far been mainly overlooked by academics is a curious (or at least non-“standard”) aspect of microfinance contracts. This is that lenders often expect loans to be paid in small installments, starting soon after the initial disbursement. In the Grameen Bank model, the installments are weekly. Similarly, in Bolivia between 1987 and 1995 the microlenders Caja Los Andes demanded weekly repayments from about half of its clients. Another 42 percent made repayments every other week (i.e., bi-weekly), and the remaining 6 percent made monthly installments. For its competitor, BancoSol, over one-third of clients were asked to repay weekly, about one-quarter paid bi-weekly, and the rest paid monthly.⁵⁷

While having several installments is not unusual for consumer loans made by commercial banks, it is atypical for loans made (at least on paper) for investing in businesses. In “standard” business loans made by traditional commercial banks, the

⁵⁵ Grameen Bank’s “Grameen Bank II” is the most notable example—although it remains too new to assess and observers have concern that the ease with which loans can be rescheduled under the new system will unduly exacerbate moral hazard.

⁵⁶ The need for credit bureaus is made forcefully by McIntosh and Wydick (2002) who show cases where, in principle, competition can worsen the lot of the poorest households. Competition can, in particular, make it difficult to cross-subsidize the poorest borrowers.

process is just as you would think: entrepreneurs borrow, invest and grow their businesses, and then--once sufficient profits have been earned--repay the loan with interest. Here, it is quite common to expect repayment to start the next month or week!

Table 5.3 provides more data from Bolivia collected by a research team from the Ohio State University. For both Caja Los Andes and BancoSol, the weekly repayment schedules were demanded on smaller-sized loans, while the larger loans carried bi-weekly or monthly installments. On average, it is poorer households that are being asked to repay in more frequent installments, since it is poorer households that tend to take smaller loans.

Table 5.3

Loan Terms and Conditions in Bolivia

BancoSol and Caja Los Andes, 1995

Repayment frequency	Median amount initially disbursed (\$)	Median term to maturity (months)	Effective annual real interest rate (percent per year)
<i>Caja Los Andes</i>			
Monthly	37	1	35
Weekly	62	3	35
Weekly	106	5	34
Bi-weekly	309	5	33
Monthly	309	6	26
Monthly	309	6	23
<i>BancoSol</i>			
Weekly	62	3	59
Bi-Weekly	72	4	53
Monthly	82	6	48

Source: Gonzalez-Vega, et al (1997), Table 15, pp. 49-50. Amounts are in US dollars at the exchange rate of 4.93 bolivianos per dollar. The effective annual real interest rate is calculated as 12 times the internal monthly rate of return of the contract (in real terms) for loans with median size and median term to maturity. The data reflect loans denominated in bolivianos only; both lenders also provided dollar-denominated loans--in much larger sizes (e.g., the median size for Caja Los Andes was about \$2500) with monthly or bi-weekly installments, lower real interest rates (30 percent per year or below), and year-long terms to maturity.

⁵⁷ Data on Bolivia are reported by Gonzalez-Vega, et al. (1997), p.74.

The puzzle is why the repayment should be scheduled this way. One explanation is that it creates an early warning system. By meeting weekly, credit officers get to know their clients well by meeting face-to-face on a regular basis. This information can provide loan officers with early warnings about emerging problems and offer bank staff a protocol by which to get to know borrowers more effectively—and clamp down more quickly when needed. Personalized relationships and regular opportunities for monitoring are thus established, just as with local moneylenders.⁵⁸ Drawing on their research in Bolivia, Gonzalez-Vega, et al (1997, p. 74) stress the value of the early warning feature, asserting that “The most important tool for the monitoring of borrowers in these lending technologies is requiring frequent repayments followed by immediate reaction in the case of arrears.” The observation is reinforced through an example:

After the creation of BancoSol, the proportion of its clients making monthly repayments increased. A couple of years later, BancoSol revised this policy, most likely in response to higher arrears in 1992-93. Thus, the proportion of loans with weekly repayments increased from 27 percent in 1993 to 47 percent in 1995.⁵⁹

Ani Silwal (2003) also notes the correlation between repayment troubles and the frequency of required installments. He compares repayment performance in nine “village banks” in Nepal and finds that 11 percent of loans were not repaid by the end of the loan period when installments were weekly, while twice that rate (19.8 percent) were delinquent when loans were paid in a single lump-sum payment at the end of the loan’s maturity (which was generally 3-4 months). Similarly, when BRAC in Bangladesh experimented with moving from weekly repayments to twice-per-month repayments, delinquencies soon rose, and BRAC—just like BancoSol—quickly retreated to their weekly scheme.⁶⁰

⁵⁸ A theoretical formalization of this notion would follow the treatment of repeated lending contracts described in Parikshit Ghosh and Debray Ray (1997).

⁵⁹ Quoted from Gonzalez-Vega, et al (1997), p. 74.

⁶⁰ Morduch interview with Fazle Abed, Founder and Chairperson of BRAC, Dhaka, December 2002.

But puzzles remain. After all, the “early warning system” explanation does not answer why it could make sense to demand repayments *before* investments are likely to have borne fruit. Moreover, as Gonzalez-Vega, et al (1997, p. 74) argue, “While frequent repayments are critical in keeping the probability of default low, they increase the transaction costs incurred by borrowers and thereby reduce the quality of service to the client.” On the face of it, having to pay more frequently does seem to impose an added constraint on borrowers. But we suggest below that this is too simple. For borrowers that have difficulty saving, the frequent repayment schedules can *increase* the quality of service to the client.

Before we get to that, we suggest why it could make sense for the bank to demand initial installments to be repaid so soon after loans are disbursed. One answer is that it helps the bank select less risky clients. The frequent repayment schedule reduces the bank’s risk by selecting borrowers that are more likely to be able to repay loans even if their investments fail. This is because households must have some *other* stream of income on which to draw in order to repay the early installments.⁶¹ So, requiring frequent and early installments mean that the bank is effectively lending partly against that stream of outside income, not just the proceeds from the project. The bank is therefore taking advantage of the borrower’s ability to obtain funds from family members or from household activities apart from the given investment project.

For example, if before borrowing the household has a net income flow of \$10 per week after expenses from the husband’s wage job, the microfinance institution can fairly safely lend the wife an amount under \$520 (52 weeks times \$10) to be repaid in a year with the confidence that the household in principle has resources to repay even if the project fails. The example assumes that the husband is happy to help pay off the loan, and to the extent that’s not so, the bank would have to reduce its calculations of maximum feasible loan size for the wife. But the example captures the flavor of the way that loan officers assess the repayment ability of their clients. Strikingly, in most of the programs surveyed by Craig Churchill (1999), lenders estimate repayment capacity

⁶¹ Of course, part of the early installments can be (and often is) paid directly from the not-yet-invested principal of the loan. This makes the effective loan size smaller. The practice does not fully answer the puzzle at hand, since it cannot explain the bank’s logic in requiring that the first installments are paid so

without taking into account expected revenues from the loan in question, and they take into account income flows provided by all household members.⁶²

We have to push a bit further, though, to more satisfactorily explain the requirement of frequent installments. One question is: why not do as above and estimate repayment capacity based on household income (rather than expected investment income) but not require frequent installments? An answer is that the repayment schedule is the easiest way for the microlender to “capture” those other household income flows (which are earned throughout the year) and guarantee that they are put toward paying off the bank loan.

A related part of the story is that frequent installments will be particularly valuable for households that have difficulty holding onto income. This takes us back to issues of savings constraints addressed in the context of ROSCA enforcement in chapter 3--and about which we will say more in chapter 6. If borrowers must wait months before they repay loan installments, part of their earnings may be dissipated as neighbors and relatives come by for hand-outs, spouses dip into the household kitty, and discretionary purchases command attention. Months later, funds may no longer be there to pay the bank. A repayment schedule with frequent installments instead takes the money out of the house soon after it is earned. The essential insight is that everyone gains by matching repayment schedules as closely as feasible to the cash flowing into borrowers' households. In this way, loan products become saving product too, and the result is the initially-puzzling hybrids that we see in practice.⁶³ It is also why we asserted above that

soon. The bank, of course, might not be acting fully logically, but we suspect that there is more to it than that.

⁶² Sanjay Jain and Ghazala Mansuri (2003) offer a different but related story. They argue that if borrowers must resort to informal lenders (rather than the flow of other income coming into the household), then the microlender can piggy-back on the informal lender's informational advantage. In other words, if you can't get a microloan without also getting a short-term loan from the moneylender to pay for the initial microloan installments, then only people judged to be creditworthy by moneylenders will demand microloans. The microlender gains due to this implicit screening mechanism. The mechanism is plausible in theory, but we do not know of any evidence that gives it empirical credence. Instead, other family income is most typically used to pay for initial installments, and it is unclear that this would provide the same kind of helpful piggy-backing described by Jain and Mansuri.

⁶³ Our discussion here is influenced heavily by conversations with staff at Bank Rakyat Indonesia about how they determine loan terms. And by Stuart Rutherford's (2000) volume, *The Poor and Their Money* which considers lending mechanisms in the context of savings problems. We present a more “formal” discussion in Armendariz and Morduch (2000).

for borrowers that have difficulty saving, the frequent repayment schedules can *increase* the quality of service received.

The calculation of optimal repayment schedules will then involve the timing and amount of the income that is earned by the household, the difficulty that households have holding onto that income, the bank's desire for early warnings of troubles, and both the bank's and customers' transactions costs associated with collecting repayments. All else the same, if households can save without difficulty and transactions costs are high, the optimal number of installments falls. PRODEM, a rural lender in Bolivia, for example, requires monthly installments because they find that weekly installments are too costly in the low population density areas in which they work (Gonzalez-Vega, et al., 1997). But where saving is hard and transactions costs are relatively low, weekly repayments are more likely to appeal. The latter scenario will hold with poorer households, where the opportunity cost of time is relatively low, and where the mechanisms to enforce financial discipline are relatively limited. These tendencies are reinforced by the fact that small-scale business like petty trading tend to generate a flow of revenue on a daily or weekly basis, making frequent collections especially desirable in the absence of satisfactory savings facilities. In wealthier households, however, opportunity costs are likely to be higher and revenue costs less frequent, militating toward less frequent loans installments. These arguments are in line with the pattern of weekly versus monthly installment schedules seen in Table 5.3, in which bigger loans, which tend to go to wealthier clients, are more likely to be repaid in larger but less frequent installments.

One notable problem is that these regular repayment schedules are difficult to impose in areas focused on highly seasonal occupations like agricultural cultivation. Indeed, seasonality poses one of the largest challenges to the spread of microfinance in areas centered on rainfed agriculture, areas that include some of the poorest regions of South Asia and Africa. (Another major challenge in lending in agriculture is covariant risk, where a bad drought, a pest infestation, or the like can devastate an entire region, debilitating the microlender too.)

The Grameen Bank's proposed "Grameen Bank II" attempts to address this issue in part by maintaining weekly repayment schedules (for all of the reasons discussed above) but allowing loan officers to vary the size of weekly installments according to

season (Yunus, 2002). In low seasons borrowers can ask to pay less in return for paying more during high seasons.

We close this section with a question: since many lenders appear to judge repayment capacity without taking into account expected revenues from the investment that the loan is intended for, why don't the borrowers simply save up the money needed, rather than taking out a loan with interest? The answer must partly hinge on discount rates (borrowers would rather have assets sooner if possible) and partly on savings constraints (saving up is not so easy). We suspect that if more households did have better ways to save, the demand for loans would fall considerably. Which takes us to a provocative thought. As Stuart Rutherford (2000) notes, the requirement of frequent installments not only builds recognition of saving difficulties into loan products, it also means that some customers with particular problems saving may logically look to the new microfinance loan products as an alternative way to “save”—i.e., as a useful mechanism to help convert the small, frequent bits of money that enter the household into a big lump that can be used for a major purchase or investment. For these customers, that the particular financial product is structured and labeled as a “loan product” may be of secondary concern.

5.3 Complementary Incentive Mechanisms

In the rest of the chapter we describe additional means used by microlenders to secure repayments. We describe important mechanisms now in use and one interesting proposal (on “cross-reporting” strategies) that could, in theory, improve on or supplement existing schemes.

5.3.1 Flexible Approaches to Collateral

One premise of microfinance is that most clients are too poor to be able to offer collateral. Loans are thus “secured” through non-traditional means like group lending. But in practice some microfinance lenders do require collateral, the best known being Bank Rakyat Indonesia (BRI). In rural Albania, for example, microlenders require tangible assets such as livestock, land, and housing to be put up (in addition to any assets purchased with loans), and the programs have been vigilant in enforcing agreements if

clients fail to repay. In urban Albania, a borrower's home or business is typically required as collateral (Benjamin and Ledgerwood, 1999).

Microlenders like BRI take a non-traditional view of collateral. While BRI requires collateral in general, the bank is flexible in the assets that it will accept, and in practice collateral is not a major constraint when seeking poor clients. A survey completed in 2000, for example, shows that 88 percent of non-customers had acceptable collateral of some sort.⁶⁴ All the same, the survey shows that non-customers have much less in the way of assets to use as collateral. Table 5.4 below shows that the median value of collateralizable assets held by BRI borrowers is roughly 2.5 times the median value of those held by a random sample of non-customers drawn from the same area. In order to push still further, BRI has introduced products that require no collateral at all for loans up to Rp. 2 million (\$225 in 2003), offered at the discretion of the unit manager.⁶⁵

BRI's view is that the re-sale value of collateral is far less important than the judgment that the pledged items should be particularly problematic for households to give up. Thus, household items may be considered as collateral if they have sufficient personal value for borrowers, even if they are worth relatively little in the hands of BRI. The idea breaks with the traditional banker's view that collateral should be valuable enough so that banks can sell the collateral to cover the costs of problem loans. In other words, for BRI the value of collateral is determined by the *notional* value of the asset, not the expected sale value. Land without a certificate of title, for example, may be nearly impossible to sell without the cooperation of the borrower and the local community. It thus has very little value to BRI if the client is hostile. But BRI still sees such collateral as being potentially valuable. In part, it is an indicator of borrower intent and a guarantee that borrowers have resources to use if they should get into repayment difficulty.⁶⁶

Table 5.4

⁶⁴ The survey of customers and non-customers was completed by Bank Rakyat Indonesia and analyzed by Morduch.

⁶⁵ Personal communication with Don Johnston, a resident advisor to BRI in Jakarta, January 29, 2003.

⁶⁶ BRI's policy is consistent with the view of collateral as a lever to improve credit contracts. In some cases, requiring collateral may be a lender's way of obtaining assets from the poor. Debraj Ray (1998), for example, argues that in India moneylenders sometimes require collateral and are pleased when borrowers default since it allows asset transfers from poor borrowers to wealthier moneylenders. This is not the case in microfinance.

Collateral value (Rupiah x 10,000,000)

	25th percentile	Median	75th percentile
<i>Value x 10,000,000</i>			
BRI borrower	1.1	2.3	4.1
BRI saver only	0.9	1.9	3.8
Non-customer	0.4	.91	2.1

Note: cell size for BRI borrowers, $n = 175$; for BRI saver only, $n = 170$; and for non-customers, $n = 741$. On June 1, 2000, 10 million rupiah were equivalent to \$1160.

Source: BRI survey, 2000. Calculations by Morduch.

More formally, we extend the framework above to show how collateral requirements discourage borrowers from defaulting on debt obligations. Let w be the collateral that the bank confiscates at the contracting stage. Returning to the set-up in section 5.1, take $v = 0$ which, again, is the optimal re-financing strategy from the bank's standpoint. Then, the borrower's incentive compatibility constraint becomes: $y - w \leq y - R + \delta y$, or, equivalently: $-w \leq -R + \delta y$. This, in turn, implies that the bank's maximum gross interest rate can be as large as $R = v\delta y + w$. Thus, with collateral requirements the bank is now able to charge a higher interest rate while not fearing a greater probability of default. But note that the bank does not need to take possession of and *sell* the collateral for this constraint to bind, it only needs to deny the borrower access to the collateral.

5.3.2 Financial Collateral

The flexible approach to collateral described is one solution when borrowers lack assets. Another solution is to hit the problem straight on—to provide ways for borrowers to build up financial assets and then to base lending on those assets. Many microlenders, for example, require that borrowers show that they can save regularly for a period before they become eligible to borrow. Demonstrating the ability to save demonstrates characteristics like discipline and money management skills that correlate with being a

good borrower. But saving also leads to deposits in the bank, and that can help directly by providing security for loans.

At *SafeSave* in the Dhaka slums, the first loan product developed required that borrowers held a savings account for three months before borrowing was allowed. The maximum size of the loan was determined as (current savings balance) + (ten times the smallest monthly net inflow of savings over the previous three months).⁶⁷ While loans are outstanding, savings withdrawals are restricted in some *SafeSave* loan products.

At Grameen Bank, the policy at the end of 2003 was that borrowers holding loans must deposit between 5 and 50 taka per week into obligatory personal savings accounts (between about 10 cents and one dollar in December 2003), with the amount depending on their loan size.⁶⁸ For most loans, an obligatory deposit equal to 2.5 percent of the loan value is also deducted off the top of the loan and placed into the borrowers' personal savings accounts. Another 2.5 percent is put into a "special savings" account. On top of this, borrowers taking loans larger than 8,000 taka (about \$145) are required to open a Grameen Pension Savings account with a monthly deposit of at least 50 taka. The GPS requires monthly deposits for a term of five to ten years. Borrowers in good standing can withdraw from their personal savings accounts at any time, provided they visit the branch with their passbook. The "special savings" accounts, though, have heavier restrictions—for example, withdrawals are not allowed for the first three years. And the GPS is a fixed term account which, if it goes into arrears, is closed and the funds are returned with reduced interest. Loan ceilings are predicated in part on the size of these various loan balances.

How well can these kinds of deposits function as collateral? On one hand, if borrowers get into repayment trouble, the microlender can, in principle, hold onto the deposits to minimize their exposure to the full extent of the default. Saving up is not easy, so borrowers will surely be careful when their nest egg is at risk. On the other hand, if the outstanding loan is larger than the funds on deposit, the lender remains exposed to the possibility of default on the difference. From this vantage, the use of financial collateral does little more than to effectively reduce the capital that borrowers

³² Product data are from personal communication with Stuart Rutherford, January 2004. Similar data are available at www.safesave.org.

have available to them, since the borrower's savings are tied up with the lender and not available to be invested by the borrower. Since borrowers have to pay higher interest rates on the money that they borrow than they receive as interest on their deposits, the scheme can also add substantial, "hidden" costs to borrowing.

This discussion assumes, though, that borrowers see a dollar as a dollar, a peso as a peso, and a taka as a taka. In other words, it assumes that money saved is "counted" the same as money borrowed. But if borrowers attach special worth to money saved over time, the microlender might be able to capitalize on financial collateral and its "special" place in the borrower's heart and mind—and in the process to provide larger loans with lower risk. It is often noted, for example, that individuals will prefer to borrow—even at relatively high interest rates—than to draw down the savings that they have diligently built up over years.⁶⁹ The bottom line is that using financial collateral can be an effective way to facilitate lending, but it hinges on special assumptions about borrower psychology and constraints that are unlikely to hold for everyone or at all times.

5.3.3 Making Repayments Public

In an important break from its original model, ASA of Bangladesh ultimately weakened its insistence on the group lending mechanism in its credit practices. Customers often still meet as groups, though, making public repayments. Similarly in "Grameen Bank II" the focus shifts from the group to individual relations between borrowers and loan officers. Still, though, customers meet as groups and make public repayments.

A telling story on the importance of public repayments comes from a Grameen Bank replication in Kenya that ran into trouble before instituting monthly public meetings with borrowers. Originally, the lender had instructed borrowers to deposit their installments directly into a bank account, but the incidence of default soared. Repayment

³³ Data are from Stuart Rutherford, personal communication, January 2004.

³⁴ Morduch personal communication with Monique Cohen, President of Microfinance Opportunities, an organization based in Washington that is focused on better understanding how microfinance customers use financial services (March 2004).

rates come under control only after bank officials started meeting in villages with borrowers each month, collecting installments face-to-face.⁷⁰

Public repayment schemes have several advantages for the lender. First, without the ability to secure collateral, microlenders can use the avoidance of social stigma as an inducement for individual borrowers to promptly repay loans (Rahman, 1998). Public repayments heighten the ability to generate stigma--or, more powerfully, the *threat* of stigma. Second, by meeting as a cluster of borrowers in scheduled locations, and at scheduled times, some transactions for bank staff might be reduced, even if it adds to clients' costs. Third, the group is often a useful resource through which staff can directly elicit information about errant borrowers and create pressure as needed (i.e., "cross-reports" described below). Fourth, group meetings can facilitate education and training, which may be particularly helpful for clients with little business experience and/or low literacy levels. The education might aid financial performance or it might be valued intrinsically as a way to improve levels of health and knowledge. Fifth, it is often said that the comfort of clients (many of whom have had no prior experience with commercial banks) is enhanced by encouraging them to approach the bank with their neighbors. And, sixth, by keeping transactions in the open, public repayments can help enhance internal control for the bank and reduce opportunities for fraud.⁷¹

5.3.4 Targeting Women

The Grameen Bank has bound microfinance to creating opportunities for poor women. Much that is written on Grameen focuses on gender issues, and we devote Chapter 7 to issues around gender. But Grameen did not start with such a strong focus on women. The bank lent originally to large numbers of men, in addition to women, keeping both groups and centers segregated by sex. When the focus shifted, in the early 1980s, the move was mainly a response to growing repayment problems in male centers, and by the end of that decade well over 90 percent of clients were women. At the end of 2002, 95 percent of clients were women.

³⁵ This story is related in Rai and Sjöström (2004), drawing on Espisu, et al., (1995). An alternative explanation of the story offered by Stuart Rutherford is that "people pay when they are asked to, and tend not to pay if they're not asked (the oldest rule in banking)."

As we describe in chapter 7, women seem to be more reliable than men when it comes to repaying their loans (before conditioning on other variables like social status and education). Mahabub Hossain (1988), for example, argues that women in Bangladesh are more reliable customers, citing evidence that 81 percent of women had no repayment problems versus 74 percent of men. Similarly, Shahidur Khandker et al., (1995) find that 15.3 percent of male borrowers were “struggling” in 1991 (that is, missing some payments before the final due date), while only 1.3 percent of women were having difficulties. In Malawi, David Hulme (1991) finds on-time repayments for women customers to be 92 percent versus 83 percent for men, and David Gibbons and S. Kasim (1991) find that in Malaysia the repayment comparison is 95 percent for women versus 72 percent for men.⁷²

The evidence suggests that it may thus be profit-maximizing for banks to lend to women, independent of other concerns about gender. Why women often seem to be more reliable customers is up for debate. Helen Todd’s (1996, p. 182) time in two Grameen villages in Tangail leads her to argue that it has to do with women being “more cautious” than men, who are more likely to have trouble sustaining membership over the long term. Based on a later village study, Aminur Rahman (1988) finds that women instead tend to be much more sensitive to the verbal hostility of fellow members and bank employees when repayment difficulties arise, while men are more likely to be argumentative and non-compliant. In Indonesia, a manager of a Grameen Bank replicator argued that women were better customers because they tended to stay close by the home rather than going out to work. This makes women, on average, easier to find when troubles arise and gives them little way to escape pressures; men, on the other hand, more easily remove themselves (physically) from difficult situations.⁷³

In terms of the dynamic incentives analyzed above, women will be more likely to repay (than men) if they have fewer alternative sources of credit. Since men may have

³⁶ Thus, a lender like SafeSave, that bases its operations on one-on-one visits by staff to client homes rather than public transactions, has one less lever to use in maintaining internal control.

³⁷ The data from Hossain (1988), Hulme (1991), and Gibbons and Kasim (1991) is taken from Hulme and Mosley (1997) as cited in Wright (2000), p. 23.

³⁸ Morduch interview with George Oetomo, General Manager for Operations, Yayasan Dharma Bhakti Parasahabat (www.ydbp.com), March 2003.

greater access to formal credit and to informal credit from traders and moneylenders, men may have weaker repayment histories than their wives and sisters.

These observations are surely not universal and are apt to change over time. And not all successful microlenders focus on women. Bank Rakyat Indonesia, for example, does not especially target women, but they still boast near-perfect repayment rates. Concerns with gender should thus be seen within the broader context of a lender's approach and objectives, as well as wider social, cultural, and economic constraints—issues taken up further in Chapter 7.

5.3.5 Information-Gathering by Bank Staff

In the 19th century German credit cooperatives, borrowers were asked to obtain a loan guarantee from a neighbor. By inducing joint liability, the loan guarantee was a precursor to group lending. More recent experience shows that even without a formal loan guarantee, incorporating neighbors in credit decisions can improve bank performance.

In another step away from traditional bank practices, many microlenders spend considerable time talking with prospective borrowers' neighbors and friends when making lending decisions. One microlender in Russia, for example, relies heavily on staff visits to applicants' businesses and homes, rather than just on business documents (Zeitinger, 1996). In rural Albania, applicants must often obtain a loan guarantee and character reference from a member of the local "village credit committee." Similarly, Churchill (1999, p. 55) describes practices at Bank Rakyat Indonesia:

At the BRI units, most loan rejections are based on character, not the business assessment. Rejection occurs if the credit officer learns that the applicant is not respected in the community or has misrepresented himself in the application. Almost without exception, the unit staff interviewed for this research identified the neighbor's assessment of the applicant's character as the most important means of predicting a new applicant's future repayment behavior—more important than the business assessment.

At ADEMI in the Dominican Republic, credit officers also check the stability of home life, based on their finding that “troubled homes often become troubled borrowers” (Churchill, 1999, p. 56). At Financiera Cálpia in El Salvador, agricultural extension workers are important informants about some borrowers’ character, and accordingly credit officers build ongoing relationships with extension workers.

Thus, even where group lending is not used, novel mechanisms are in place to generate information. Credit officers get out of their branch offices and get to know the neighborhoods in which they work. Microlenders find that the views of shopkeepers, bartenders, school teachers and other central figures in communities can be as helpful in assessing borrower’s creditworthiness as a stack of business plans.⁷⁴

5.3.6 Cross-Reporting

Gathering information from neighbors can be helpful at many stages in the loan process, not just at the application stage. One problem faced by microlenders using the threat not to refinance defaulters is that it’s a strong penalty. It’s particularly strong when coupled with group lending, since, in principle at least, the entire group should be cut off when any member fails to repay. Ashok Rai and Tomas Sjöström (2004) argue that these punishments are inefficiently tough, and that “cross-reporting” can improve performance.⁷⁵

Cross-reporting refers to statements made by one borrower about another. If Mrs. Haq is willfully refusing to repay (despite having the necessary resources), the bank can take appropriate action if Mrs. Rahman speaks up about it. If Mrs. Haq’s troubles are not self-imposed, Mrs. Rahman can provide helpful input then too (preventing the bank from coming down too hard on Mrs. Haq). Rai and Sjöström describe how cross-reporting can be reliable and improve efficiency. While their focus is on improving group lending schemes, cross-reporting can have wider applications.

³⁹ Churchill (1999) describes similar monitoring and information-collection mechanisms in individual lending programs run by the Alexandria Businessman's Association in Egypt and the Cajas Municipales of Peru, and he is the source for the information on Financiera Cálpia cited above.

⁴⁰ Armendáriz de Aghion (1999) provides an alternative view.

In order to work, the bank must credibly commit itself to a system of reward for truthful reports, and the bank must itself check on its borrowers' monitoring activities. One fear is that formalizing such a system may create tensions among individual borrowers or a strong incentive for them to collude. Still, "cross-reporting" seems promising in a variety of settings, and, as Rai and Sjöström argue, it is already an informal feature of banking relationships, especially coupled with group lending.

5.4 Summary and Conclusions

Group lending with joint responsibility is far from the only innovation in microfinance. Successfully creating dynamic incentives and creating products that are built around households' cash flows have been as important. Good dynamic incentives are created through attractive long-term relationships. When forward-looking customers know that default means risking losing the relationship, incentives to work hard are strengthened. Helping customers to manage cash flows is also critical, since it helps banks to give banks access to customer resources before they are spend or otherwise dissipated. Weekly or monthly repayment schedules, although a sharp break from traditional banking practices, have been particularly critical in allowing customers to repay loans in manageable bits. Strategic microlenders often attempt to break repayment installments into pieces that are small enough that customers can, if needed, repay loans from household funds other than profits from the given investment project. The bank's risks are considerably reduced as a result.

In order to work effectively in sparsely populated rural areas, in highly transient urban areas, and with more mature clients, it has been necessary to develop additional mechanisms. Even where group-lending has been central (e.g., in the densely-populated villages of Bangladesh), the additional mechanisms have been put to good use. These additional mechanisms include flexible approaches to collateral (where what matters most is the value that the customer attaches to losing the item, rather than the value that the lender expects to recover from selling the item) and having public repayments, even when joint responsibility is not a part of credit contracts. It is not clear in the end how important group-lending is to the continued success of microfinance. We expect that the

future will see much more innovation, and the beginning point should be better understandings of existing mechanisms.

But, to date, the innovations described above have been studied far less than group lending, and we know of few systematic attempts to sort out which mechanisms have most power in practice, nor how the mechanisms operate together. Progress could be made by experimenting with different mechanisms in a way that would allow researchers to properly infer causality--say, by using different methodologies in different, randomly-chosen branches. Microlenders will understandably be reluctant to give over their decision making to a random number generator, but building some elements of randomization into research and development can allow more systematic product testing and piloting—and cleaner answers on what really drives microfinance performance.

Chapter 8

Measuring Impacts

Introduction

8.1 How Microfinance Affects Households

8.2 Evaluation Basics

8.3 Addressing the Selection Problem in Practice

8.3.1 Using Prospective Data in Northeast Thailand

8.3.2 Attrition Bias: Problems When Using “new borrowers” as a Control Group in Peru

8.3.3 Longitudinal data: USAID AIMS Studies in India, Peru, and Zimbabwe

8.3.4 Using a quasi-experiment to construct instrumental variables: Bangladesh Studies

8.3.4.1 Estimates from the 1991/92 cross-section

8.3.4.2 Estimates from the full panel, 1991/92 – 1998/99

8.4 Summary and Conclusions

Exercises

Introduction

Anecdotes on the benefits of microfinance abound, and inspiring stories from around the globe have helped to turn microfinance from a few scattered programs into a global movement. Consider the story of Mrs. Braulia Parra, who lives with a family of seven in a poor neighborhood in Monterrey, Mexico, in a home with cardboard walls and dirt floors.²³ Illiterate and inexperienced in the workplace, Mrs. Parra took her first \$150 loan from ADMIC, a local microlender. The loan allowed her to buy yarn and other sewing supplies to make hand-sewn decorations. Each week she sells about 100 hand-made baskets, dolls and mirrors, going door-to-door in her neighborhood. After ten loans, Mrs.

²³ This story was taken from accion.org/insight/meet_meet_our_borrowers.asp in mid-2003. The site also contains stories of other ACCION customers.

Parra had earned enough to install a toilet in her modest home, as well as an outdoor shower. Building a second floor was next in her sights.

Anecdotes like this are not a substitute, however, for careful statistical evidence on impacts from large samples. For every Braulia Parra, was there another customer that fared poorly? The number of careful impact studies is small but growing, and their conclusions, so far, are much more measured than the anecdotes would suggest.²⁴

Microfinance is touted as a way to raise incomes for the very poor, but USAID-sponsored studies of SEWA Bank in India, Zambuko Trust in Zimbabwe, and Mibanco in Peru, for example, found that on average borrowers had net income gains only in India and Peru. In Zimbabwe, there were no measurable increases in average incomes relative to those in control groups (Snodgrass and Sebstad, 2002).²⁵

This should not be surprising: the anecdotes are culled to show the potential of microfinance, while the statistical analyses are designed to show typical impacts across the board. Inevitably, some customers will thrive, others will be unchanged, and some may slip backwards. One study of Bolivia's BancoSol, for example, reports that staff estimated that in any given cohort roughly 25 percent showed spectacular gains to borrowing, 60-65 percent stayed about the same, and 10-15 percent went bankrupt (Mosley, 1996b). Increasing income is, of course, not the only metric by which to judge microfinance. Microfinance participation can affect households in many ways.

Researchers have analyzed a range of social and economic outcomes beyond household income and consumption, including business profits, nutrition, schooling, fertility, contraception, risk, asset-holdings, and a range of measures of "empowerment" and changes in social consciousness.²⁶ In the USAID study of Zimbabwe, for example, clients were shown to diversify their income sources more than others, a potentially important means of risk diversification.

²⁴ Ledgerwood (2001, pp. 49-50), for example, concludes that "Few [microlenders] invest much in impact analysis, and the literature on microfinance and microenterprise development has been remarkably short on discussions of the subject."

²⁵ Even in Peru, a second look at the data shows that the results are not 100% robust. As we describe below, Alexander (2001) shows strong, positive results on income even after controlling for household-level unobservables, but the results are not robust when econometrically treating the problem of reverse causation from income to credit using instrumental variables methods.

²⁶ See Jennifer Sebstad and Gregory Chen (1996) for an overview of the range of outcomes that have been evaluated.

No matter what are the outcomes of interest, the most difficult part of evaluating impacts is to separate out the causal role of microfinance. That requires stripping out the various “selection” and “reverse causation” biases common to nearly all statistical evaluations. Even if earnings from microfinance participation are funding new houses, further education for children, new savings accounts, and new businesses, we have to ask whether these changes are more remarkable than that would have happened without microfinance? If we see that richer households have larger loans, we have to ask whether the loans made the households richer--or whether richer households have easier access to credit (or both)? Ultimately, the question that every careful evaluation seeks to answer is whether borrowers would have done just as well without the programs?

In practice, it's a surprisingly difficult question to answer cleanly. One major problem is that many microfinance clients already have initial advantages over their neighbors. In examining village bank programs in Northeast Thailand, for example, Brett Coleman (2002) finds that households that will later become microfinance borrowers tend to already be significantly wealthier than their nonparticipating neighbors before the village bank starts its operations. Moreover, the wealthiest villagers are nearly twice as likely to become borrowers as their poorer neighbors, and the wealthiest are also more likely to use their power to obtain much larger loans than others. Gwen Alexander (2001) similarly finds that microfinance borrowers in Peru start off considerably richer than their nonparticipating neighbors.

In a small sample from Bangladesh, Syed Hashemi (1997) also finds important underlying differences between borrowers and non-borrowers in villages served by Grameen Bank and BRAC. Over half of those who chose not to participate did so because they felt that they could not generate adequate profits to repay loans reliably. Another quarter opted out due to religious and social sanctions that restricted the ability to participate in meetings outside of the home with non-family males. If sufficient care is not taken to control for such self-selection into microfinance programs, estimated “impacts” on income and “empowerment” will be misleading. The microfinance interventions will seem more positive than is indeed the case.

Unfortunately, this is not an esoteric concern that practitioners and policymakers can safely ignore. It is not just a difference between obtaining “very good” estimates of

impacts versus “perfect” estimates--the biases can be large. In evaluating the Grameen Bank, for example, Signe-Mary McKernan (2002) finds that not controlling for selection bias can lead to overestimation of the effect of participation on profits by as much as 100 percent. In other cases discussed below, controlling for these biases reverses conclusions about impacts entirely.

In many ways, the problems and solutions are no different from evaluations of health and education interventions, say. So microfinance researchers can learn much from the broader literature on evaluation. But, by the same token, the best empirical work on microfinance holds lessons for researchers with interests beyond the financial sector. We take up important issues around evaluation methods, challenges, and solutions below.

8.1 How Microfinance Affects Households

First, researchers have to ask: what are we trying to measure? Microfinance may affect household outcomes through a variety of channels. Most immediately, microfinance may make households wealthier, yielding an “income effect” that should push up total consumption levels and, holding all else the same, increase the demand for children, health, children's education, and leisure. But running microenterprises may also take time (and make that time relatively more valuable than other activities), yielding “substitution effects” that may counter-balance the effects of increased income. With increased female employment, for example, time spent raising children can become costlier in terms of foregone income, pushing fertility rates downward.²⁷ The need to have children help at home (to compensate for extra work taken on by parents) could decrease schooling levels, and leisure, in this case, may fall if the return to working rises sufficiently. Only evaluating impacts on business profits, for example, may thus miss out on other important changes within the household.

The fact that it is often women that are earning the income is not incidental. As described in Chapter 7 on gender and microfinance, another way that microfinance can affect household outcomes is by tipping the balance of decision-making. With added

income, women may gain clout within the household, using it to push for greater spending in areas of particular concern to women.

Micro lenders may also make direct, non-financial interventions that affect client outcomes. Some programs use meetings with clients to advise on family planning and to stress the importance of schooling and good health practices, taking advantage of group meetings to hold communal discussions and training sessions. Village banks that are run on the “credit with education” model developed by the NGO Freedom From Hunger have made this a mainstay of their approach, for example, and other micro lenders like Latin America’s *ProMujer* have added training and education components in various ways (Dunford 2001). Taking these kinds of extra benefits into account, McKernan (2002) finds that being a member of the Grameen Bank is associated with a 126 percent increase in self-employment profits *after* accounting for the direct benefit of access to capital.²⁸ The increase, she presumes, is due to increased social and human capital derived from group meetings.

The multiplicity of channels means that it is typically impossible to assign a given measured impact to the strictly financial elements in microfinance, although there have been attempts to analyze programs that are essentially similar but which differ in a specific, limited way. In order to separate out the role of education programs, for example, ideally one would want to run programs without the “credit with education” training sessions and compare them to similar programs that use the integrated approach. Stephen Smith (2002) does this with data on Project HOPE’s “health banks” in rural Ecuador and urban Honduras. He finds that the health interventions did indeed improve health care for the participants relative to the health care received by those in credit-only programs, and the health interventions did not diminish the the banks’ financial performance. There is also hope that health interventions like this might have impacts on

²⁷ Pitt et al. (1999) show evidence that these substitution effects may be weak in the case of fertility in Bangladesh, since most micro-enterprises are based in the home, making it possible to simultaneously raise children and run new businesses without the added burdens that jobs outside the home would entail.

²⁸ Grameen does not use the “credit with education” model, but they do incorporate some social components into their activities, and the very act of meeting in village groups may have some intrinsic benefits for participants. McKernan’s estimates also imply that a 10 percent increase in capital will, on average, yield a 20 percent increase in profit—a result that is so large that it leads us to wonder about robustness of the specification. Malgosia Madajewicz, in her Ph. D. dissertation, suggests that McKernan’s results weaken when capital is disaggregated into a fixed capital component and a working capital component.

household income and spending by reducing the incidence of illness and raising productivity, but results on that score are mixed in Smith's sample. Much could be learned by following Smith's example to gauge the impacts of business training, marketing, "consciousness-raising", and other interventions that are sometimes delivered alongside financial services.

8.2 Evaluation Basics

To be concrete, we will focus on attempts to measure the causal impact of microfinance on borrower income. Income can be attributed to many sources. Most immediately, those sources are your job, your business, your pension, etc. But here we take one step backward in order to focus on more basic sources like your age, education, and experience. These attributes are generally measurable. Another category of attributes is far harder to measure, like your entrepreneurial skills, your persistence in seeking goals, your organizational ability, and your access to valuable social networks. In this latter category, we also include "shocks" like whether you had a bad flu last winter or whether you had a falling out with your boss. Another set of attributes has to do with where you live—for example, in a city or village (measurable) or in a place with a thriving local market (measurable, but typically not actually recorded in surveys). A final broad category includes income determinants that tend to be broadly felt, like political upheavals, rampant inflation, or economic booms.

Calculating the impacts of microfinance requires disentangling its role from the simultaneous roles of all of the attributes above. The challenge is made harder by the fact that the decision to participate in a microfinance program—and at what intensity—will likely depend on many of those same attributes. Thus, there is likely to be a high correlation between microfinance participation and, say, your age and your entrepreneurial ability. Since researchers can record your age, there are simple ways of controlling for age-related issues. But since entrepreneurial ability is typically unmeasured, researchers need to be careful in making comparisons or else the impact of being a better entrepreneur could misleadingly be interpreted as an impact of microfinance access.

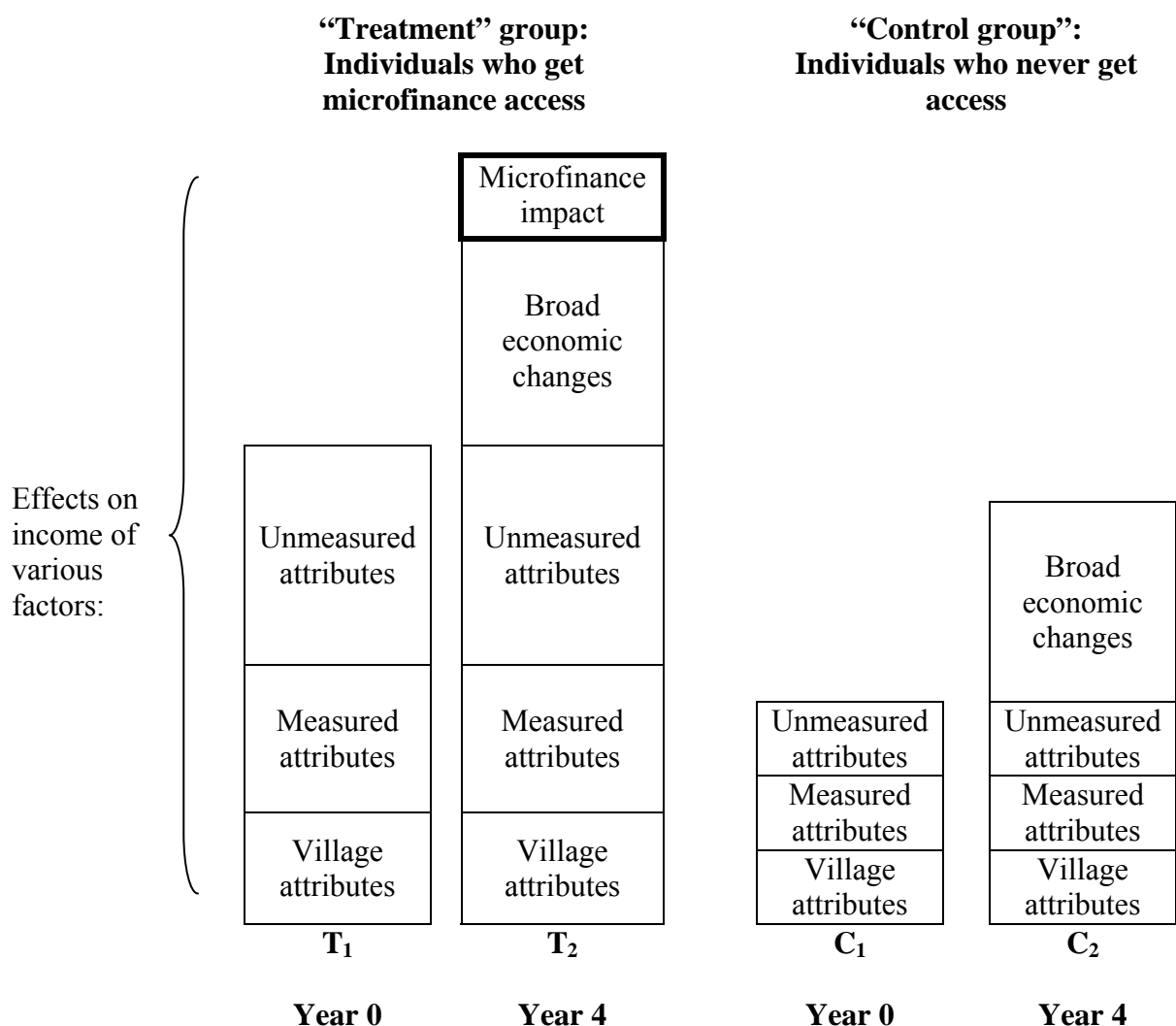


Figure 8.1: Sources of Income for "Treatment" and "Control" Groups

With this in mind, we use Figure 8.1 to consider various evaluation approaches. The ultimate goal is to isolate and measure the "microfinance impact" in the bold box. The impact is felt by a "typical" person who gains access to a microfinance program. We term this position T_2 , taken to be four years after the program started. Before access to the program, in year 0, this person's income is reflected by position T_1 . The difference between T_2 and T_1 is a useful place to start as it nets out the roles of those measured and

unmeasured individual attributes that do not change over time, as well as location-related issues. But while the difference captures the microfinance impact, it also reflects broader economic and social changes that occur between year 0 and year 4 and that are independent of microfinance. It would thus be misleading to attribute the entirety of the $T_2 - T_1$ difference to the microfinance impact. The problem is that we cannot parse it without more information.

Identifying a control group is thus critical. Figure 8.1 shows a plausible control group from an area without access to microfinance. It would be very unlikely to find a population that was exactly identical to the “treatment” population, and we see here that base income levels start at a lower level for the control group in this example. Thus, comparing, say, the difference between T_2 and C_2 will help address biases due to the broadly-felt economic and social changes, but it will not take the differing base levels into account. Isolating the true microfinance impact requires comparing the difference $T_2 - T_1$ with the difference $C_2 - C_1$, which is a so-called difference-in-difference approach.

Given the set-up in Figure 8.1, the difference-in-difference approach is adequate to deliver accurate measures of microfinance impacts. But we have made an implicit assumption that we now need to put on the table. We have taken the impacts of personal attributes like age, education, and entrepreneurial ability to be unchanging over time. Thus, their effects net out when we look at $T_2 - T_1$ and $C_2 - C_1$. But in reality, these characteristics may change over time (perhaps a borrower gets more education or strengthens their social networks—for reasons unrelated to microfinance). And/or these characteristics may directly affect changes over time, so they do not net out as assumed. More capable entrepreneurs will likely have greater earnings *growth*, for example, and not just a higher base level of income. When the relevant variables are not measurable, the problem is mitigated by making sure that control groups are as comparable to treatment groups as possible.

To find comparable treatment groups, we need to consider who joins microfinance programs in the first place. Figure 8.2 gives a plausible scenario, where the focus is just on entrepreneurial ability. Participants tend to have more entrepreneurial ability and non-participants tend to have less. Participants thus have higher incomes—and potential for income growth—before the microfinance program even arrives.

Comparing microfinance borrowers in a given village to their neighbors who decide not to participate is thus apt to run into problems. The former already has an advantage, reflected by the average income level I_P , relative to their non-participating neighbors with average income level I_{NP} . As noted above, the concern is that unmeasured attributes like entrepreneurial ability may affect both income growth and initial income levels.

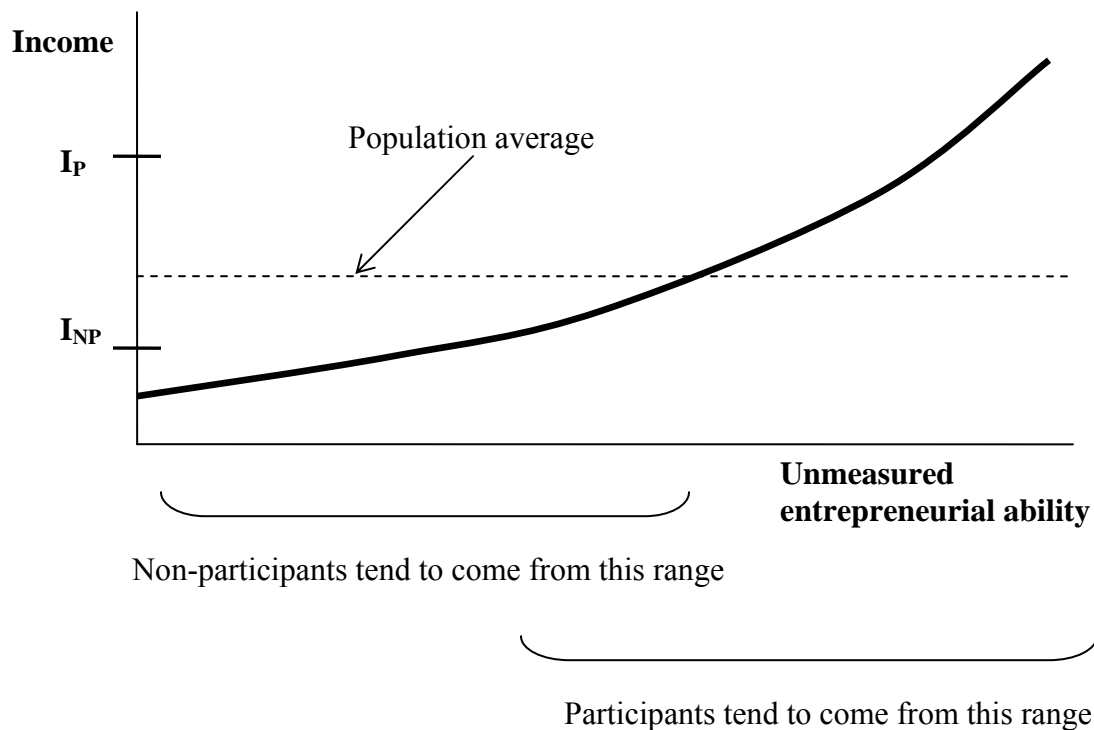


Figure 8.2: The Hypothetical Relationship Between Unmeasured Entrepreneurial Ability and Income in a Given Village

So, imagine that we had access to data from another village that was identical to the one depicted in Figure 8.2, except that the second village lacked a microfinance program. It would seem to provide a perfect control group, but how should it be used? Figure 8.2 shows that comparing the income of participants in the treated village

to the population average in the control village will also create problems since the former group is “self-selected” while the latter is not. The problem of course is that by definition there are no “participants” in the control village since it has no program yet.

Two solutions present themselves, and we discuss them and other ideas below. The first idea is to change the question. We have been asking: what is the effect of microfinance participation? But instead we could ask: what is the effect of microfinance *access*—whether or not villagers ultimately end up participating? To answer this second question, which may well be more relevant from a policy standpoint, it is only necessary to compare outcomes for the entire population in the treatment village (or, more easily, a random sample drawn from the entire population) against a sample drawn from the control village. A second approach, used by Brett Coleman (1999), is to try to identify future borrowers in the control villages and to compare the actual microfinance participants to the set of future participants. A third approach, that is common but problematic, involves comparing older borrowers in a given village to newer borrowers who are just joining the program. The main difficulty here involves non-random attrition, an issue discussed in section 8.3.2 below.

8.3. Addressing the Selection Problem in Practice

In the sections below, we consider a series of related approaches to impact evaluation. The overview is not exhaustive and we do not aim to provide a full survey of impact surveys to date. Rather, we aim to point to key methodological issues and to gather several important results. The results to date are decidedly mixed, with some evidence of modest positive impacts of microfinance on income, expenditure, and related variables, while other studies find that positive impacts disappear once selection biases are addressed. There have been few serious impact evaluations of microfinance so far, though, so a collection of definitive results is still awaited. All the same, the existing studies provide useful insights and directions for future research.

8.3.1 Using Data on Prospective Clients in Northeast Thailand

A number of recent studies use novel research designs to address selection biases. One approach is to use information on borrowers before the microfinance program enters.

Brett Coleman (1999) and (2002) takes advantage of a particular way a microfinance program was implemented in Northeast Thailand, providing a unique way to address selection bias. He gathered data on 445 households in 14 villages. Of these, eight had village banks operating at the start of 1995. The remaining six did not, but village banks would be set up one year later. Interestingly (and critically for the evaluation), at the beginning of 1995, field staff from the village bank programs organized households in these six villages into banks, allowing the households to self-select according to the village bank's standard procedures. But then the households were forced to wait one year before getting their first loans.

The unusual procedure on the part of the programs allows Coleman to analyze who joins and who does not *before* the village banks start running. Moreover, it allows him to estimate the following regression equation:

$$Y_{ij} = \mathbf{X}_{ij} \alpha + \mathbf{V}_j \beta + M_{ij} \gamma + T_{ij} \delta + \eta_{ij}, \quad (8.1)$$

where the variable to be explained Y_{ij} is a household-level outcome--income or profit--for household i in village j . The regression approach allows a refinement of the difference-in-difference approach discussed in the previous section. Here, “dummy variables” (i.e., variables that only take the values of 0 or 1) are used to control for location and participation status. Other variables control for factors like age and education.²⁹ The variables \mathbf{X}_{ij} capture household characteristics (and a constant term); and \mathbf{V}_j is a vector of village dummy variables that control for all fixed characteristics of the village. The two variables of most interest are M_{ij} and T_{ij} . The first is a “membership dummy variable” that equals 1 for both actual members of the village banks and those villagers who have opted into the programs (in the “control” villages) but who have not yet received loans. Coleman argues that M_{ij} controls for selection bias, so that δ , the coefficient on T_{ij} , is a consistent estimate of the causal treatment effect. In his application, the variable T_{ij} is the number of months of village bank credit was available to (actual) members, which is exogenous to the household.

²⁹ For more on regression approaches, see, for example, Peter Kennedy's (2004) *Guide to Econometrics*.

Controlling for selection makes an important difference. Coleman (1999) finds that average program impact was not significantly different from zero after controlling for endogenous member selection and program placement. When he extends the estimating framework to differentiate between impacts on “rank and file members” and members of the village bank committee (who tend to be wealthier and more powerful), he finds that again most impacts were not statistically significant for “rank and file members”, but there were some noted impacts for committee members, particularly on wealth accumulation.

Coleman cautions, though, that the results need to be put in the context of the larger financial landscape. Thailand is relatively wealthy (at least compared to Bangladesh, say), and villagers have access to credit from a range of sources, some at low interest rates from government-backed sources. Strikingly, survey households held over 500,000 baht in wealth on average and had over 30,000 baht of “low-interest” debt (excluding village bank debt). Thus, the village banks’ loans of 1500 to 7500 baht may be too small to make a notable average difference in the welfare of households (in fact complaints about small loan sizes had prompted some women to leave the banks). Coleman argues that one reason that wealthier borrowers may have experienced larger impacts because they could commandeer larger loans.

8.3.2 Attrition bias: Problems When Using “New Borrowers” as a Control Group in Peru

A problem in trying to replicate Coleman’s approach is that it’s not often that a researcher comes upon programs that of to the trouble of organizing villagers but then delay credit disbursement for a period. So instead, researchers have tried to capture the flavor of the approach by comparing “old borrowers” to “new borrowers” within the same area. Typically this is done with cross-sectional data, yielding an approach that is simple and relatively inexpensive (and which does not require surveying non-borrowers). The procedure has been promoted by USAID through its AIMS project (more on this below) and by other microfinance organizations (see Karlan, 2001).

Assuming that the characteristics of people that enter into programs are unchanging over time, the method should control for the fact that borrowers are not a

random group of people. But assuming that the relevant characteristics are similar over time requires a leap of faith. Why did the new borrowers not sign up earlier? Why were the older borrowers first in line? If their timing of entry was due to unobservable attributes like ability, motivation, and entrepreneurship, the comparisons may do little to address selection biases—and could, in fact, exacerbate bias.

Dean Karlan outlines two additional problems based on his experience evaluating village banks organized by FINCA Peru. Assume that the conditions of selection *are* constant over time such that the same kinds of people become clients today as became clients five years ago. All seems well, but there are still two potential biases, both of which are most pronounced when assessing impacts using cross-sectional data. Both are also due to drop-outs.

Drop-outs are an ongoing reality microfinance. Sometimes borrowers leave because they are doing so well that they no longer need microfinance, but more often it is the borrowers in trouble that leave. Wright (2001) gives evidence that dropout rates are 25 to 60 per cent per year in East Africa. In Bangladesh, Khandker (2003) estimates rates for three leading lenders of 3.5 per cent per year between 1991/92 and 1998/99, which is much smaller than the numbers cited by Wright, but which nonetheless can add up over time. Gonzalez-Vega et al (1997, pp. 34-35) give parallel data for Bolivia. They investigate the fraction of people who ever borrowed from a given microlender that are still active borrowers at the time of their survey (the end of 1995). The resulting proxy for retention rates shows that just half of BancoSol clients were still active. In rural areas, two-thirds of borrowers from PRODEM were still active, possibly reflecting the fact that there are fewer alternative lending sources in the country-side.

It is likely that these “older borrowers” (i.e., those who remain active) have the positive qualities of survivors, while “new borrowers” have yet to be tested. If the failures are more likely to drop out, comparing old to new borrowers will overestimate impacts. We suspect that this pattern is most often the case, but, as suggested above, the prediction is not clear-cut. If it is mainly the successes that move on (leaving weaker clients in the pool), the sign of bias will be reversed, under-estimating causal impacts.

The second problem is due to non-random attrition independent of actual impacts. If richer households are more likely to leave, the pool of borrowers becomes poorer on

average. Then it could look like microfinance borrowing makes one poorer, when in fact it may have no impact at all. Conversely, when poorer households leave in greater numbers, impacts will be over-stated.

Karlan argues for hunting down the dropouts and including them in the analysis along with the other older borrowers, though it may be costly. A cheaper improvement would be to (i) estimate predictors of dropout based on observable information on older borrowers, then (ii) form a prediction of who among the new borrowers is likely to (later on) drop out, and (iii) use the prediction to weight the new borrower control group. The method is not perfect, though: in particular, drop-outs who made their decision based in part on the size of impact are not addressed by the re-weighting scheme.

8.3.3 Longitudinal data: USAID AIMS Studies in India, Peru, and Zimbabwe

Some biases can be mitigated by using data collected at several points in time, allowing “before versus after” comparisons as described in section 8.2. Under certain conditions, the approach controls for both non-random participation and non-random program placement. But when those conditions are not met, the approach is subject to biases due to unobservable variables that change over time—hard-to-observe characteristics like entrepreneurial spirit and access to markets that are likely to be correlated with borrowing status.³⁰

The most ambitious longitudinal studies to date are those sponsored by the United States Agency for International Development in the late 1990’s, with the hope to demonstrate methods and generate benchmark results.³¹ Teams analyzed impacts on members of the Self-Employed Women’s Association, a labor organization (and microlender) serving women in the informal sector in Ahmedabad, India; Mibanco, an ACCION International affiliate in Peru, and the Zambuko Trust in Zimbabwe. Baseline data was collected and then the same households were re-surveyed two years later. Case studies were also conducted alongside the statistical analyses.

³⁰ The reliability of methods based on differences is reduced as the time periods get closer together, reducing temporal variation. Differencing noisy data can also exacerbate measurement error; in the “classical” case this leads to attenuation bias. Noisy recall may thus bias downward coefficients which show program impacts. See James Heckman and Jeffrey Smith (1995) and Angus Deaton (1997) for more detailed discussion of methods.

The teams selected clients randomly from lists provided by the programs. The trick was then to identify control groups. In India and Peru, the control group was a random sample drawn from non-participants in the same regions who met program eligibility criteria. In Zimbabwe, enumerators instead used a “random walk procedure” in which they set off in a given direction to find non-client households for the control group. As Barnes, et al (2001, p. 19) explain, “For example, when the client’s business was in a residential area, from the front of the house the interviewer turned right, went to the first road intersection, turned right and walked to the third intersection and then turned left; from there the interviewer asked a series of questions to identify who met the criteria for inclusion in the study.” The criteria used to match treatments and controls were gender, enterprise sector and geographic location, as well as additional criteria added by Zambuko Trust: “a) never received credit from a formal organization for their enterprise, b) be the sole or joint owner of an enterprise at least six months old, and c) not be employed elsewhere on a full-time basis” (Barnes et al 2001, p. 19).

The data have potential, and the researchers followed drop-outs as best they could to avoid the attrition biases described above. With two years of data, the researchers could have analyzed impacts by investigating how changes in microfinance participation affect changes in outcomes. But, surprisingly, the AIMS researchers chose *not* to analyze variables converted to changes over time, which would have eliminated all biases due to omitted variables that do not change over time (i.e., to analyze differences-in-differences as described in section 8.2). The stated rationale is that the “differencing” procedure also eliminates the chance to analyze the roles of variables like gender and enterprise sector which are also fixed through time, and so alternative methods (analysis of covariance) were used (Dunn 2002). In our view, the costs of that choice far outweigh the benefits.

To see the differencing method (i.e., the method not used), we can modify equation (8.1) to specify that the variables are measured in a given time period t :

$$Y_{ijt} = \mathbf{X}_{ijt} \alpha + \mathbf{V}_j \beta + M_{ij} \gamma + C_{ijt} \delta + \eta_{ijt}, \quad (8.2)$$

³¹ An earlier set of longitudinal studies is described in Hulme and Mosley (1996). Quality control problems have diminished their relative value as more careful studies have been completed (see Morduch 1999c).

As before, we are interested in estimating the value of δ , but here it is the coefficient on the value of loans received. (The two variables, value of loans and length of membership, are typically very similar since loan sizes and length of time borrowing often move closely together.) The dependent variable, Y_{ijt} , is a household-level outcome (income or profit) for household i in village j at time t . The variables \mathbf{X}_{ijt} capture household characteristics at t (and a constant term), and \mathbf{V}_j is a vector of village dummy variables *that are assumed to be unchanging over time*. The dummies will capture village-level features like distance to the closest major city, proximity to major transportation and markets, and the quality of local leadership. Similarly, we assume that the individual-specific variable M_{ij} , the variable that captures non-random individual selection into the program, is also unchanging over time. It may reflect, for example, an individual's energy level, management ability, and business savvy. In this case, though, we do not assume that it is observable. Thus, there is a potential bias stemming from its omission when equation (8.2) is estimated.

The problem can be addressed by estimating in differences. Assume that we have the same variables collected in period $t+1$:

$$Y_{ijt+1} = \mathbf{X}_{ijt+1} \alpha + \mathbf{V}_j \beta + M_{ij} \gamma + C_{ijt+1} \delta + \eta_{ijt+1}. \quad (8.3)$$

Then, we can subtract equation (8.2) from (8.3) to obtain:

$$\Delta Y_{ij} = \Delta \mathbf{X}_{ij} \alpha + \Delta C_{ij} \delta + \Delta \eta_{ij}, \quad (8.4)$$

where Δ indicates the difference in the variables between periods t and $t+1$. Here, the village dummies drop out, as do the fixed (and unobservable) individual-specific characteristics (which was the concern that prompted the AIMS researchers not to follow this method). The benefit, though, is considerable: a consistent estimate of the impact δ can be obtained, which is the most important aim.³²

³² All fixed household-specific variables drop out as well (like education level, for example) so their effects cannot be independently estimated in equation (8.4), which was a concern of the AIMS researchers (although one that was weighted too heavily in our view). There are two important caveats here. The first is that estimating equation (8.4) can exacerbate attenuation bias due to measurement error (it can make

It turns out that the omitted unobservables in equations like (8.2) do make a large difference, and not addressing them undermines the credibility of the AIMS impact studies. When Gwen Alexander (2001) returns to the AIMS Peru data and estimates the equations in differences (akin to equation 8.4), she finds that estimated impacts on enterprise profits fall. In fact when she controls for reverse causality by using an instrumental variables approach (more on this below), the estimated impacts shrink and are no longer statistically significant. Selection bias is clearly a major problem, but results might be different if the two surveys had been collected more than two years apart or if other instrument variables had been used. Below we address why finding instrumental variables continues to be a challenge.

8.3.4 Using a quasi-experiment to construct instrumental variables: Bangladesh Studies

A different way of approaching the problems above would have been to search for an “instrumental variable” for microfinance participation. The instrumental variables method allows researchers to address problems posed by measurement error, reverse causality, and some omitted variable biases. The instrumental variables strategy involves finding an additional variable (or set of variables) that explain levels of credit received but which have no direct relationships with the outcomes of interest (like profit or income). Then a proxy variable can be formed based on the instrumental variable and it can be used to tease out the causal impact of credit access.

The interest rate is a potential instrumental variable--or simply “instrument”--since it can explain how much credit a borrower desires while not being a direct determinant of income in itself (that’s testable, at least). The trouble is that interest rates seldom vary within a given program, and the statistical techniques are impossible without some variation. And, while it is true that interest rates vary when comparing clients of different institutions--both formal and informal--it is likely that the variation partly reflects unobserved attributes of the borrowers, undermining the use of interest rates as instruments. Lender characteristics are also candidates for instrumental variables. Like all

positive coefficients shrink toward zero). Second, time-varying unobservables are not addressed. Both concerns suggest that instrumental variables methods are required for consistent estimation.

other community-level variables, though, they will be wiped out when including village dummy variables in specifications if there is no variation in program access within a village. In short, the instrumental variables approach can be powerful, but finding convincing instrumental variables for credit has been frustrating.

But when there is within-village variation in program access, rules determining eligibility can be the basis of an evaluation strategy, an approach employed in a series of studies of microfinance in Bangladesh. Over the years 1991 and 1992, the World Bank and the Bangladesh Institute of Development studies surveyed nearly 1800 households in 87 villages in Bangladesh; most villages were served by microlenders but 15 were not. In 1998 and 1999, teams were sent back to find the same households, but by then all of the villages were served by microlenders.³³ After losing some households through attrition, 1,638 households were left that were interviewed in both rounds.

In a sign of the rapid spread of microfinance in Bangladesh, about one quarter of the sample included a microfinance customer within the household in 1991/92, but by 1998/99 the figure had jumped to about half.³⁴ The jump makes program evaluation more difficult, but not impossible. To complicate matters, about 11 percent of customers were members of more than one microfinance institution in 1998/99.

8.3.4.1 Estimates from the 1991/92 Cross-Section

The first round of data has, on its own, generated a series of papers, and the most important results have been compiled in Shahidur Khandker's (1998) volume, *Fighting Poverty with Microcredit*. Completing impact studies with just a single cross-section requires ingenuity and some important assumptions, and the task was made more challenging by the desire to estimate impacts of borrowing by men and by women separately. The studies are sophisticated in their use of statistical methods in compensation for the fundamental limitations of the data set. One large limitation arises because the researchers were eager to generate results with the first wave of the data rather than waiting for the second. That the studies use heavier statistical artillery than other microfinance studies does not necessarily mean that they deliver results that are

³³ The survey focused on customers of Grameen Bank, BRAC, and RD-12, a government program, but by 1998/99 a variety of other lenders were operating within the survey area, including ASA and Proshika.

more reliable or robust than other studies. In fact, as we describe below, the studies are open to questions about the validity of the underlying assumptions that prop up the statistical framework.

On the face of it, it would seem that it would be impossible to get far with just a single cross-sectional data set and without a special set-up like that of Coleman (1999). But the way that the microlenders in Bangladesh implement their programs opens a door for researchers. To capture the basic insight, Figure 8.3 shows two hypothetical villages, one with a program (the treatment village) and one without (the control village). The villages are separated into distinct groups based on their eligibility and participation status—we discuss how eligibility is determined in a minute. The groups within the thick black lines are eligible to borrow (or, in the case of the control village, *would be* eligible). As a first step, researchers could compare the incomes and other outcomes of microfinance participants to non-participants just using data from the treatment village, but it is impossible to rule out selection biases of the sort described in section 8.2. It is also possible to use the control villages to compare participants from the treatment villages served by microfinance to the eligible households from the control villages, but even here there are potential selection biases since the participants are still a select group.

A more satisfactory approach is to compare eligible households (all households within the thick black lines) across the two types of village. Here, the goal is to estimate the impact of microfinance *access* rather than actual participation. The benefit is that a clean estimate of the average impact of access may be more useful than a biased estimate of the impact of participation. Moreover, if there are no spillovers from participants to non-participants, it is possible to recover a clean estimate of the impact of participation from the estimate of access (by simply dividing the latter by the fraction of households in the village that participate). (The assumption that there are no spillovers is strong, though, and Khandker, 2003, finds evidence against it.)

The cost of the latter approach is that, while selection biases at the household-level are addressed, it does not address biases stemming from non-random program placement. As mentioned above, villagers served by microlenders may seem to do poorly relative to control groups just because the microlender chooses to work in isolated,

³⁴ Data on the surveys and household characteristics are taken from Khandker, 2003.

disadvantaged villages. In other cases, villages may be doing better than average even without the microlender, so the bias would go in the other direction (estimated impacts would be too high).

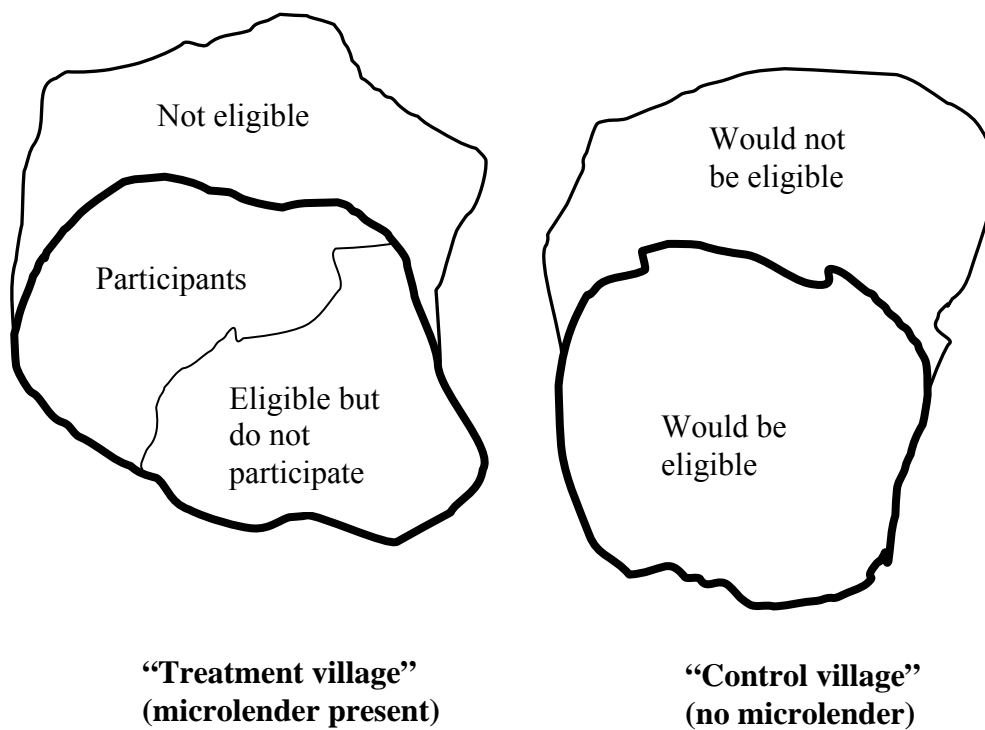


Figure 8.3: Example of Impact Evaluation Strategies Using Eligibility Rules

A potential solution is in hand, though, provided by the particular way that the microlenders being studied determine eligibility for program access. Mark Pitt and Khandker (1998) develop a framework for estimating impacts using the 1991/92 cross-section. The starting point is the observation that the three programs being studied—Grameen Bank, BRAC, and the state-run RD-12—all share the same eligibility rule. In order to keep focused on the poorest, the programs restrict their services to the “functionally poor”, and this is implemented through a rule that households owning over half an acre of land are not allowed to borrow. The individual programs place some additional restrictions, but the half-acre rule is the common criterion. So, in terms of Figure 8.3, the functionally landless are encompassed by the thick black lines, and the non-eligible lie outside. The fact that there are ineligible households within villages with programs means that there is another control group that can help to alleviate concerns that the microlenders choose villages that are special in one way or another.

An improved estimation strategy—but *not* the one adopted by Pitt and Khandker—is to compare “differences-in-differences” as described in section 8.2. It involves comparing the outcomes of households with microfinance access to the outcomes of households that are ineligible but living in treatment villages. The strategy then turns to the control villages, where the ineligible are compared to those who “would be” eligible. Finally, those two comparisons are pitted against each other. The result tells us whether households with access to microfinance are doing better than their ineligible neighbors, relative to the difference in outcomes between similar households in control villages versus *their* ineligible neighbors.

One can do even better by implementing the strategy in a regression framework so that a broad range of household characteristics is also controlled for. In the regression framework, the difference-in-difference strategy would be implemented as:

$$Y_{ij} = \mathbf{X}_{ij} \alpha + \mathbf{V}_j \beta + E_{ij} \gamma + (T_{ij} \cdot E_{ij}) \delta' + \eta_{ij}, \quad (8.5)$$

The idea is much like that of equation (8.1), but two important changes are made. First, E_{ij} is a dummy variable that reflects whether or not a household is eligible to borrow from a microlender (whether or not there is in fact a microlender present in the village). The

variable equals one if a household is within the thick black lines in either village in Figure 8.1. The other important change is the variable $(T_{ij} \bullet E_{ij})$, which is the product of E_{ij} and a dummy variable for whether the household is in a treatment village; it equals one only if the household is within the thick black lines in the village with a microlender. The coefficient on the dummy variable gives the average impact of credit access--after controlling for being functionally landless, living in a particular village, and having specific household characteristics.

Morduch (1998) takes the approach in equation (8.5) and finds no sharp evidence for strong impacts of microfinance on household consumption, but he finds some evidence that microfinance helps households to diversify income streams so that consumption is less variable across seasons. The estimates, though, rely on the assumption that the village dummy variables perfectly capture all relevant aspects about the villages that would lead microlenders to decide where to locate. In this setting, though, the village-level dummies only control for unobservables that affect all households in a village identically (and linearly). Non-random program placement thus remains an issue if, as is plausible, the functionally landless are noticeably different from their wealthier neighbors (noticeable to bank staff but not the econometrician), and if the programs take this into account when deciding where to locate. In that case, the dummy variable $(T_{ij} \bullet E_{ij})$ could pick up the effects of these inherent differences, biasing estimated impacts.

Morduch (1998) also takes a closer look at the eligibility rule on which the strategy rests. As Pitt and Khandker (1998) point out, it is important that land-holdings are exogenous to the household—i.e., that households are not, for example, selling land in order to become eligible to borrow. If that was the case, selection biases would creep back in even when estimating using equation (8.5) since unobservably promising borrowers would be taking special steps to switch their eligibility status. Pitt and Khandker cite the fact that in south India in the 1980's, village land markets tended to be thin, and most land was acquired through inheritance. In that case, land-holdings were exogenous to the household and unlikely (or at least much less likely) to be correlated with unobserved potential. But Bangladesh in the 1990's is not, of course south India in the 1980's, and land markets in the study area turn out to be fairly active—and this is

evident upon closer inspection of the landholding module of the data set. On the other hand, Morduch (1998) finds no evidence that households are selling land in order to meet microfinance eligibility criteria. If anything, successful borrowers are buying land, and one explanation for Morduch's inability to find significant impacts on household consumption could be that funds were instead going to land (and other asset) purchases.

The reason that households are not selling land to gain access to microfinance raises another, different tricky issue. It turns out that the microlenders were not following the eligibility criteria strictly, so many households owning over a half an acre were nonetheless borrowing in 1991/92. As a result, there was no reason to sell land to become eligible. Khandker (2003) acknowledges the problem and finds that 25 percent of borrowers were over the half acre line in 1991/92 and 31 percent were over in 1998/99.³⁵ Mark Pitt (1999) follows up on the issue and suggests that households with more land have worse quality land, so they still may be poor, even if they are not strictly "functionally landless". The problem when estimating of equation (8.5) is that the "eligible" households in the control villages were surveyed on the basis of a strict interpretation of the half-acre rule, while the "eligible" households in the treatment villages include the mis-targeted households. Morduch (1998) adjusts the samples in order to maintain comparability, and Pitt (1999) does robustness checks to show that the Pitt-Khandker (1998) results change little when mis-targeting is taken into account.³⁶

These issues should be borne in mind when turning to the Pitt and Khandker (1998) framework. We start by noting that equation (8.5), which can be run using ordinary least squares, is closely related to the following instrumental-variables approach: estimate instead:

$$Y_{ij} = \mathbf{X}_{ij} \alpha + \mathbf{V}_j \beta + E_{ij} \gamma + C_{ij} \delta'' + \eta_{ij}, \quad (8.6)$$

³⁵ In a demonstration of how loosely the targeting rules were taken, Khandker (2003) shows that in 1998/99, 22 percent of households with over 2.5 acres in fact included microfinance borrowers, as was true for 42 percent of households holding between one acre and 2.5 acres.

³⁶ Had the eligibility rules be followed to the letter, it would have been possible to apply a regression discontinuity design approach, comparing outcomes of households just below the line to those just above.

where C_{ij} is the amount of credit received and $T_{ij} \bullet E_{ij}$ is employed as an instrumental variable.³⁷ Estimating equation (8.6) using ordinary least squares would bring trouble since households who have received more and larger loans can be expected to be different in unobservable ways from those who have received fewer loans (leading to a variant of selection bias associated with loan size). The instrumental variables method addresses the problem and leads to a clean estimate of δ'' , the average impact of credit access (subject to the same caveats about village dummy variables noted above).

Before moving on to the method used by Pitt and Khandker (1998), note that the instrument $T_{ij} \bullet E_{ij}$ is a dummy variable which only reflects credit access. The estimate of δ'' thus does not draw on variation in *how much* credit is received, it only depends on whether credit is received. The step taken by Pitt and Khandker is to expand to a larger set of instruments by, in effect, using as instruments $X_{ij} \bullet T_{ij} \bullet E_{ij}$. The step yields as many instruments as there are X's (the X's include education and various aspects of household demographics). The move means that the estimate of δ'' takes advantage of variation in how much credit households receive.

An important identifying assumption is that the specification in equation (8.6) is correct, such that education and demographics affect household outcomes in exactly the same way for the whole sample; otherwise, biases enter back in. In other words, it is assumed that there are no important nonlinear relationships in the ways that age, education, and the other variables influence outcomes of interest.³⁸ Another critical identifying assumption stems from their use of a Tobit equation to explain credit demand in the step when they are effectively creating the instrumental variables used in the final regressions. The Tobit provides a way to efficiently handle variables with many zero values (like credit), but it requires that all microfinance impacts must be assumed to be identical across borrowers, an assumption that is often made out of necessity but which stretches plausibility. It also implies (implausibly) that marginal and average impacts of credit are equal. Estimating using a simpler two-stage least squares method would lead to consistent estimates without requiring these assumptions, but the method is less

³⁷ The equation will then be exactly identified: there is one endogenous variable and just one instrument.

³⁸ Pitt and Khandker (1998) demonstrate that their results are robust to allowing flexibility in the specification for the land holdings variable but do not show results with flexible treatments of other variables.

efficient (i.e., coefficients would tend to have larger standard errors). By using the Tobit, the efficiency of the estimators is improved.

Pitt and Khandker take one more step, which is to investigate credit received by men separately from credit received by women (motivated by the concerns raised in chapter 7). To do this, they take advantage of the fact that microlending groups are not mixed by gender in Bangladesh. In the 87 villages surveyed in 1991/92, 10 had no female groups and 22 had no male groups (and 40 had both, leaving 15 villages with no groups). Identification in this case comes from comparing how the roles of age, education, etc. for men with access to male groups compare to their roles for men without access. Similarly, for the characteristics of women with and without access.³⁹

Pitt and Khandker's most cited result from the 1991/92 cross-section is that household consumption increases by 18 taka for every 100 taka lent to a woman. For lending to men, the increase is just 11 taka for every 100 taka lent. Men, it turns out, take more leisure when given the chance, explaining in part why household consumption rises less when they borrow. Non-land assets increase substantially when borrowing is by women, but not by men. Schooling of boys increases in general with borrowing, but schooling of girls only increases when women borrow from Grameen—but not when women borrow from the other programs. It cannot be ascertained from the estimates why loans to women have higher marginal impacts than loans to men. Pitt and Khandker interpret this as an indication of a lack of fungibility of capital and income within the household, which is plausible assuming that their basic result is right. Another, very different interpretation is given by the fact that loans to males tend to be larger, so that the smaller relative impacts may be explained, at least in part, by the standard theory of declining marginal returns to capital. However, marginal returns would have to be very sharply diminishing, since loan sizes are in the same general ballpark.⁴⁰

The 1991/92 cross-section has also been used to analyze non-credit program impacts, fertility and contraception choices, and impacts on seasonality and nutrition (for

³⁹ The fact that a man is in a village with no male groups may say something about the unobserved qualities of the men and the strength of their peer networks in that village, so identification relies on the assumption that group structures are exogenous to individuals.

⁴⁰ In 1991/92, men borrowed slightly more on average from Grameen (15,797 taka versus 14,128 taka for women). For BRAC, males cumulatively borrowed 5,842 taka versus 4,711 taka for women, and for BRDB, males borrowed 6020 taka versus 4118 taka for women (Morduch, 1998).

an overview, see Morduch, 1999b). Khandker (1998) has used the basic impact numbers described above, imperfect as they be, to estimate broad impacts on poverty and to complete cost-benefit analyses (see chapter 9 below for a more detailed discussion). The work is ambitious, but, as the discussion above suggests, the underlying set-up is far from perfect. The imperfections are not the fault of the researchers, but they do necessitate more structure, greater econometric sophistication, and a heavier load of assumptions than would otherwise be necessary. The second round of data collected in 1998/99 provides hope that simpler methods may be able to deliver more robust, transparent results, but initial results were just being circulated at we write this book.

8.3.4.2 Estimates from the full panel, 1991/92 – 1998/99

With the two rounds of data, Shahidur Khandker (2003) estimates an equation along the lines of equation (8.4). As with the work on the cross-section, he modifies the equation slightly, though, to allow for separate impacts when women borrow versus when men borrow, and in other specifications he explores spillovers to non-borrowers who live in the same villages as borrowers. As noted earlier, the control villages from 1991/92 all have programs by 1998/99, so simple before-after comparisons in treatment versus control villages are not possible. In addition, the extent of mis-targeting became more severe by the end of the 1990s.

The panel data allow us to see trends that help to put the microfinance revolution in Bangladesh into perspective. Table 8.1 puts together data from Bangladesh in Khandker (2003). If we just look at the top panel of the table, we can see that in program villages, microfinance participants saw important declines in poverty rates (as measured by moderate poverty), from a rate of about 90 percent in 1991/92 to about 70 percent in 1998/99, roughly a 20 percentage point decline. But eligible non-participants saw a similar decline (roughly 19 percentage points), as did non-eligible non-participants (roughly 20 percentage points). Pessimists may thus argue that the poverty declines for microfinance participants would have happened even without microfinance. Optimists, on the other hand, will argue that the impacts of microfinance have been far-reaching, spilling over to non-participants as well. This, they will argue, explains the broad and similar progress in villages with programs.

Table 8.1
Falling Poverty in Bangladesh
Program Participants versus Non-Participants

	Headcount for moderate poverty			Headcount for extreme poverty		
	1991/ 92	1998/ 99	Differ- -ence	1991/ 92	1998/ 99	Differ- -ence
<i>Program area</i>						
All program participants	90.3	70.1	20.2	52.5	32.7	19.8
Target non-participants	91.1	72.0	19.1	58.9	44.0	14.9
Non-target non-participants	69.8	50.8	19	23.6	19.3	4.3
Total	83.7	65.5	18.2	45.0	31.4	14.6
<i>No Program in 1991/92</i>						
All program participants	90.8	71.6	19.2	56.6	43.8	13.2
Target non-participants	87.4	82.9	4.5	57.0	51.2	6.8
Non-target non-participants	72.7	53.2	19.5	35.5	26.0	9.5
Total	80.3	67.7	12.6	46.6	38.3	8.3

Notes: Program and non-program area is based on 1991/92 program placement.

All villages had programs by 1998/99.

Source: Khandker (2003), Table 14, and calculations by the authors.

If the results for program villages are compared to results for those without programs in 1991/92, we see similar patterns: poverty rates all fell by around 19 to 20 percentage points, except in this case, eligible non-participants only saw a poverty decline of about 5 percentage points. Khandker's conclusions, based on his new set of econometric estimates, balances the optimistic and pessimistic vision: he argues that microfinance contributed to roughly one third to one half of these poverty declines. Overall, Khandker finds that at most, lending 100 taka to a woman leads to an increase in household consumption by as much as 8 taka annually. This is considerably less than the 18 taka increase that he found in the earlier cross-section, but it is still meaningful. Khandker's (1998) much-cited finding that microfinance might cause as much as a 5 percent per year drop in poverty thus appears to be far too optimistic, and we have already discussed caveats above about the cross-sectional estimation on which that calculation was based.

The World Bank-Bangladesh Institute for Development Studies surveys have yielded a broad range of interesting data, and they will surely generate a series of interesting studies and much discussion. Of important note, the emerging results from the full panel are much more muted than the initial results based just on the cross-section. Those initial results have also raised methodological issues, and, given the complicated scene on the ground in Bangladesh (where microlending has spread far and wide, leaving little scope for identifying control groups), we suspect that the ultimate resolution of how large an impact microfinance can have will be settled by data from elsewhere.

8.4 Summary and Conclusions

The microfinance movement was born of the ideal to create new banks with social and economic missions. Completing impact evaluations is an important way to determine if those missions are being achieved. As described above, there is no study yet that has achieved wide consensus as to its reliability, and this reflects the inherent difficulty in evaluating programs in which participation is voluntary and different customers use the services with varying degrees of intensity.

Still, a set of solid impact evaluations are within reach. Incorporating experimental designs into the program implementation will be one way to achieve more

reliable estimates, and useful lessons can be drawn from the experimental design of Mexico's PROGRESA education and health program.⁴¹ The discussion above shows that it matters to get details right, and that, for analytical purposes, having one very reliable evaluation is, in the end, more valuable than having one hundred flawed evaluations.

The challenges in evaluation arise because no microfinance program lends to random citizens. Instead, they carefully select areas in which to work and clients to whom to lend. When the characteristics that make borrowers different from non-borrowers are observable, the relevant conditioning variables (age, education, social status, etc.) can be accounted for in impact evaluations. Often, though, what makes clients different is not measured—borrowers may, for example, have a more entrepreneurial spirit, enjoy better business connections, or be more “focused” than non-participants. Because these kinds of unobservable attributes are correlated with having credit, what seems like an impact of getting access to credit may in fact largely reflect these unobservable attributes. Estimated impacts of microfinance will be biased if nothing is done about the problem, and the biases can be large.

An important source of selection bias stems from where institutions and their branches are located. Are they set up specifically to serve the under-served in atypically isolated areas? This may lead to apparent negative impacts if control areas are not similarly isolated. Alternatively, the programs may set up where there is good complementary infrastructure (highways, markets, large towns), biasing estimates upward. When evaluating large programs, programs may be placed in different areas for different reasons, so comparisons with control areas need to be made carefully. Some approaches, like those based on comparisons of outcomes at more than one point in time, can address those characteristics of program location that do not change over time. But they have limitations too—and often unobservable characteristics do change over time.

Still, while some observers have despaired at the impossibility of generating reliable evaluations, their despair is misplaced and too pessimistic. It is true that rigorous statistical evaluations are seldom easy. But an often-heard early concern—that since money is fungible within the household it is impossible to trace the impact of a particular loan to a particular change in enterprise profits—turns out to be a minor limitation (this

⁴¹ See the references on PROGRESA and further discussion (in a different context) in chapter 10.

has been called the “attribution dilemma” by Ledgerwood, 2001). Even if a given loan cannot be attached to a given change in profit, it is still possible to evaluate how profits change with capital (i.e., to measure the marginal return to capital) and how borrowing affects household-level variables like income, consumption, health, and schooling. In many ways, these are more interesting policy questions anyway, relative to narrow issues around sources of microenterprise profit.

Useful evaluations need not be enormous in scale, involving surveys of thousands of households. All else the same, the larger the sample, the better. But some of the smaller studies discussed here turn out to yield more reliable evidence than larger studies that are imperfect in one dimension or another.

There is currently a movement afoot to design evaluations that are rough but that let practitioners quickly gauge their broad “impacts” by tracking indicators of outcomes for borrowers. The approach, led by the Imp-Act project based at the Institute of Development Studies at Sussex, will surely provide users with a great deal of helpful data that will lead to program refinements, but they should be distinguished from “impact assessments” of the kind described above.⁴² Those studies attempt to answer the question: What would have happened to the participants had the program not existed?

Our argument is not that practitioner-friendly steps should be abandoned. Far from it: the Imp-Act tools are helping organizations to better understand their clients, to improve targeting, and to develop appropriate products and marketing. Rather, our argument is that the approach is not a sufficient way to learn from microfinance. Obtaining more careful, credible impact studies that can garner universal acceptance is also vital to push conversations forward. Reliable studies need not be complicated, they only need to be well-designed, as, for example, Brett Coleman’s interesting study in Thailand suggests.

The road does not end with impact evaluations, however. Even with a spotless, perfect impact evaluation, interpreting the results is another matter, and one that has received even less attention. Consideration of the worth of program typically stops too soon. A clear showing of a positive net impact does not necessarily mean that a program

⁴² Similar practitioner-friendly tools have been created by USAID’s AIMS project and by CGAP.

is a good candidate for support. Cost-effectiveness matters too. As described in chapter 9, the microfinance programs that are being evaluated should be judged against the costs and benefits of alternative approaches, including other ways of doing microfinance.