

**Microfinance Impact Assessments:  
The Perils of Using New Members as a Control Group**

**Dean Karlan**

**December 2001**

---

*Contributions to this research made by a member of The Financial Access Initiative and Innovations for Poverty Action.*

The Financial Access Initiative is a consortium of researchers at New York University, Harvard, Yale and Innovations for Poverty Action.

NYU Wagner Graduate School  
295 Lafayette Street, 2nd Floor  
New York, NY 10012-9604

T: 212.998.7523  
F: 212.995.4162  
E: [contact@financialaccess.org](mailto:contact@financialaccess.org)

[www.financialaccess.org](http://www.financialaccess.org)

Innovations for Poverty Action applies rigorous research techniques to develop and test solutions to real-world problems faced by the poor in developing countries.

Innovations for Poverty Action  
85 Willow St, Building B, 2nd Floor  
New Haven, CT 06511

T: 203.772.2216  
F: 203.772.2428  
E: [contact@poverty-action.org](mailto:contact@poverty-action.org)

[www.poverty-action.org](http://www.poverty-action.org)

**Microfinance Impact Assessments:  
The Perils of Using New Members as a Control Group**

Dean S. Karlan  
M.I.T. Department of Economics

November 21<sup>st</sup>, 2001

BIO:

Dean Karlan is a Ph.D. Candidate in Economics at the Massachusetts Institute of Technology. He has a joint MBA/MPP from The University of Chicago, and was a consultant to FINCA International from 1992-1995.

Thanks to Iris Lanao for discussions and for inviting me to observe the training of a cross-sectional impact assessment led by AIMS for FINCA-Peru, and to Dale Adams, Monique Cohen, Barbara McKnelly and Gary Woller for discussions and comments for this paper. All errors and opinions are my own.

## **I. Introduction**

Microfinance institutions aim to reduce poverty. Some assess their impact through a cross-sectional impact methodology which compares veteran to new participants, and then calls any difference between these two groups the "impact" of the program. Such studies have risen recently in popularity because they are cheap, easy to implement, and often encouraged by donors. USAID, through its AIMS project, encourages this methodology with its SEEP/AIMS practitioner-oriented tools<sup>1</sup>. This paper intends to inform practitioners about the perils of using such a strategy, and suggests a couple solutions to some of the larger problems with this approach.

This approach makes many assumptions that are untested, and others that are tested and false. For example, this approach assumes that dropouts have, on average, identical income and consumption levels to those who remain. Furthermore, this approach assumes that dropouts are not made worse off by participating in the program. This approach also assumes that when lending groups form they do not sort themselves by economic background. These assumptions not only are brave theoretically, but are contradicted by existing empirical research. This paper suggests a method to address the attrition biases, and suggests further research be conducted on the other implicit assumptions before expending resources on a plausibly unreliable assessment methodology.

This paper proceeds as follows: Section II describes the cross-sectional methodology as implemented by USAID and the SEEP/AIMS practitioner-oriented methodology. Section III discusses problems created by dropout, Section IV discusses problems created by the selection process into an MFI, and Section V discusses problems created by the dynamic nature of credit policy. Section VI discusses potential solutions to some, but not all, of the problems. Section VII concludes.

---

<sup>1</sup> The author bases the analysis of the AIMS tools on his personal observation of the evaluation tools being implemented by AIMS for FINCA-Peru, and the draft version of the practitioner tools manual.

## **II. Cross-Sectional Impact Assessments**

A valid control group is the holy grail of any microfinance impact assessment and must have participants who possess the same "entrepreneurial spirit" of those in the treatment group that receive the loans. The cross-sectional approach claims to overcome this problem since both its control and treatment group consist of individuals who have opted to participate in the MFI. The new entrants are the control group, whereas the veteran participants with two or more years experience with the MFI are the treatment group. The methodology then attributes any difference between these groups to the MFI, since the new entrants have received little to no treatment from the MFI, but the veterans have received two or more years of loans.

The AIMS practitioner-oriented tools developed by USAID explain this process in detail (USAID, 1999). In this approach, survey takers measure current income and consumption of members, both old and new, in an MFI. Then the analysis compares the income and consumption levels between old and new members. If the mean spending on food, for example, is higher for veteran members than new entrants, then the methodology concludes that participants in the microfinance program led to higher food consumption for its participants.

Advocates like this approach because of two operational advantages: no need to identify and survey non-participants in order to generate a control group and no need to follow clients over time as in a longitudinal study.

## **III. Dropout**

Dropout causes two major problems. I will call the first the incomplete sample bias, and the second the attrition bias. The incomplete sample bias is created because those who dropout presumably were impacted differently, and potentially worse, than those who remained. Since an impact assessment should examine the impact of the program in its entirety, not just of its success cases, these individuals must be considered as well. The attrition bias is created because those who dropout are different from those who remain, irrespective of the program impact (e.g., the wealthier participants stay and the poorer

dropout). Both are serious problems and somewhat easy to address, but the standard AIMS practitioner tools do not resolve it.

### Incomplete Sample Bias

For simplicity, think of two types of participants, those who benefit from participation and those who are made worse off. Those who benefit invest the loan proceeds in their business and generate more additional income than the interest they pay on their loan. These people stay in the program. Those who are made worse off fail to invest the money well and then dropout of the program. By only including those who remain in the program in the treatment group, those with negative impact are ignored. The cross-sectional impact analysis would find a positive impact, whereas the true impact depends entirely on the relative size of these two groups and their impacts.

The above scenario assumes that dropout is generated by failure. Now assume that dropout is generated by success. After successfully improving their business, learning to manage their money, and develop their own savings base, clients no longer need the credit and hence leave the program. In this scenario, the cross-sectional impact analysis would underestimate impact since the successes are ignored in the analysis.

### Attrition Bias

Again for simplicity, think of two types of participants, rich and poor. Suppose for the moment that the program has no impact whatsoever, neither positive nor negative, on any participant. Who drops out? If the rich dropout, the "veteran" pool will consist only of the poor types. Then, a comparison of veterans to new participants will conclude a negative impact, since the veterans are only poor but the new participants are a mix of rich and poor. On the other hand, if the poor are more likely to dropout, the "veteran" pool will consist only of the rich types. Then, a comparison of veterans to new participants will conclude a positive impact, since the veterans are only rich but the new participants are a mix of rich and poor. Note in both of these stylized cases, there was no impact whatsoever; hence, dropout is not "failure" in this case, merely bad fit. Yet the cross-sectional methodology produced a positive impact (if the poorer individuals are

more likely to dropout) or a negative impact (if the richer individuals are more likely to dropout).

#### **IV. Selection**

A selection bias refers to the problem of attributing causation to a program with voluntary selection. Those who participate in microfinance programs are more entrepreneurial in spirit, more resourceful in business, and hence more likely to overcome life's problems one way or another. Attributing their success to microfinance then becomes difficult. The cross-sectional impact assessment purports to overcome this problem since those in both the treatment and control groups selected into the program. This claim only examines the selection bias statically, and fails to realize the full dynamics of the decision to participate. Why did those in the treatment group join two years ago whereas those in the control group just joined? The answer is important. Does one join only at certain points in life? Or if peer selection determines participation, why was one person chosen two years ago and the other not until recently?

##### Timing of Decision Problem

Why does someone join a credit program now rather than 2 years ago? I do not know, but I intuit that there is a reason, and it is significant. Imagine that individuals join after coming to an epiphany that they must grow their business in order to pull themselves out of poverty. Or perhaps participants join when everyone in their household is healthy, and hence does not need constant care in the home. Such a situation suggests that perhaps access to credit is not the problem, but rather access to good health care. If ample opportunities exist for credit and savings in their community, then attributing the improvement in their lives to the microfinance institution would be erroneous. Their epiphany or their family's health should get full credit.

One way to address this problem is to analyze the alternatives for credit and savings that clients have in their communities. Since social networks can create both credit (e.g., informal loans) and savings (e.g., Rotating Savings and Credit Associations, ROSCAs) opportunities, evaluating a client's next-best alternative is not an easy task. Further

research to understand the informal opportunities to borrow and save is essential for understanding the seriousness of the timing problem.

### Peer Selection Problem

Imagine banks form like a draft for sports players. The best candidates get drafted first, the good-but-not-best candidates get drafted second, and so on and so on. Theory suggests (Ghatak 2000) and evidence supports<sup>2</sup> (Hatch 1997) that individuals are selected into banks in just this way, assortatively by quality of participants, where wealth is used by peers as a proxy for quality. Hence, one group to form in a community contains the best off; the second will be slightly less well off, etc. Again, without any impact at all, a naïve cross-sectional analysis would find veterans have higher wealth than new participants, and would attribute this difference to program impact. In fact, if one is targeting the poorest of the poor, then finding positive impact suggests failure since it suggests that perhaps the wealthier are always served first. This issue is heightened by the SEEP/AIMS practitioner-tools since their tools specifically instruct practitioners to use 2-year old banks for the 2-year old veteran pool, 1-year old banks for the 1-year old veteran pool, and new banks for the new entrant pool.

The point of the above story is not limited to the stylized case provided. Take the following scenario as another potential situation. The poorest join first because they are the ones willing to take the risk of participating in this unknown project. Next come the better off clients who only moved once they saw the product tested. Then come the middle tranche. In this scenario, comparing new entrants to veterans will underestimate impact, since the veterans will have started out poorer than the new entrants.

## **V. Institutional Dynamics**

Microfinance institutions change their strategies and/or client identification process, and such changes could affect materially the composition of a new versus veteran participant pool. If any such change systematically alters the relative wealth or income of the new

---

<sup>2</sup> Specifically in the case of FINCA International, Hatch found that older banks invited wealthier individuals to participate than younger banks, and that new banks in old areas were poorer than old banks in old areas.

versus veteran participants, again a naïve cross-sectional analysis would erroneously attribute differences to impact. I will discuss two plausible scenarios, both of which I have witnessed in the field.

### Program Placement

Microfinance institutions typically have a multi-year strategy for which communities to enter and why. Suppose, quite reasonably, that a young microfinance institution prefers to start out cautiously, and hence enters slightly more well off communities. Then, after achieving comfort with the local culture, economy, and business practices, the MFI branches out to the poorer neighborhoods. In this situation the veteran participants would all be wealthier than the new participants even if the program has no impact. Hence, a naïve cross-sectional analysis would erroneously attribute impact to the program success. The SEEP/AIMS practitioner-oriented tools try to address this issue by instructing practitioners to choose similar neighborhoods. Assuming the similar communities exist, this is possible, but if the implementation plan follows the pattern described above, similar-enough neighborhoods simply might not exist. This becomes a timing issue for the practitioner: at what point in the implementation of the plan will the practitioner learn that no valid control communities exist?

### Changes in Credit Requirements

Just like banks, MFIs often respond to changes in the economy by tightening or loosening their credit requirement. In a recession when even micro entrepreneurs are hurt, MFIs might be more stringent about the credit criteria for participating. Or perhaps they are more lenient. If tighter credit requirements effectively filter out the poorest of the poor, then individuals who join during a recession will be better off than those who join in a normal or boom time. Or if policy became more lenient, individuals who would not have received credit now do. If the impact assessment is being conducted in the middle of a recession, and two years prior the economy was either normal or in a boom, then the new participants will be more well off than the veteran participants. In this situation, a cross-sectional analysis will underestimate the true impact of the program. Or if policy became more lenient, the analysis will overestimate the true impact of the program. The point

here is not which direction the bias is, but rather that this approach to impact assessment demands that no such policy change is made, whereas reality dictates that policy does change as the economy changes.

## **VI. Solutions**

The dropout biases are particularly important when attrition is high. Both dropout problems are solvable within the constraints of the one-shot, cross-sectional AIMS approach. Although the current tools they offer do not address the problem, a change to the sampling technique can solve both problems. Conceptually, the two samples are not the same: the veteran group only consists of those who remain, whereas the new member group consists of members who will dropout. One can alter the veteran group to include those who dropout, or can alter the new member group to include only those expected to remain. The first approach is far better, and solves both of the problems. The second approach requires some econometric work, and only solves the second problem.

As discussed in Section III, one major issue is that those who dropout probably were impacted differently than those who remain, and any analysis which ignores them is akin to cherry-picking one's successes, ignoring one's failures, and then claiming victory. The solution requires conducting the "veteran" survey on a sample of members who were in the bank two years ago, some of which are still present but others of which have dropped out. Then, the analysis which compares consumption and income levels across veteran and new groups would include the complete "veteran" pool. This approach solves both the incomplete sample and the biased-dropout problems. It would be important when implementing this approach to sample randomly the veterans to interview (not just pick those easiest to contact) and to pursue them diligently. A recent study in Indonesia found that the extra effort to pursue the difficult-to-find pays off tremendously, as these individuals are significantly different from those who remain in their neighborhoods and are easy to reach (see, for example, Thomas et al., 2000).

The second approach requires combining the data on the veteran members and the dropouts to attempt to find predictors of dropout. The predictors must be observable when someone enters since they will be used to predict which new members will drop

out. For instance, distance to the meeting place, number of family members in the lending group, age of business, history of prior credit use, history of prior savings, are all observable and plausibly predictive of dropout. Using this information, one would then use econometric tools to predict who will remain amongst the new members, and then weight the new entrant sample according to their probabilities of remaining. As long as poverty is correlated with some of the observable information used to predict dropout, this solves the second problem noted in Section III. However, this does not solve the first problem discussed, since we have simply modeled who will dropout, not who will have the biggest impact. The veteran sample still contains only those with positive impacts and ignores those with negative or no impact.

## **VII. Conclusion**

The impact evaluation debate rages on in microfinance. Some believe all impact evaluations are useless, but targeting evaluations are appropriate to ensure at the minimum that the clients are the intended recipients. Others believe that mid-level impact evaluations, such as the one analyzed here, are useful and informative. As this paper highlights, the dropout biases inherent in a cross-sectional impact evaluation are problematic but solvable. However, the selection and institutional dynamics problems are more difficult. Depending on the circumstances in a given project and economic setting, these issues suggest that any findings cannot be attributed easily to the project, and hence the cross-sectional approach is not appropriate. A solid understanding of the selection process, economic environment and institutional dynamics is important in deciding whether or not to employ this mid-level, cross-sectional approach.

An alternative to mid-level impact assessments would be a two-prong approach, with many “targeting” evaluations, and a few methodologically rigorous longitudinal evaluations. The “targeting” evaluations would be small, frequent tools which monitor client targeting (but do not claim to measure impact), combined with institutional analysis which examines, from a management perspective, the efficiency and flexibility with which a program delivers its services. The longitudinal studies would have proper

control groups<sup>3</sup>, which follow all members, including dropouts. Such projects could inform the rest of the microfinance community about proper targeting, impact, and mechanism design issues. Ideally such studies also would test different product designs, so that one could assess differential impact of one product over another. Organizations which conduct such studies would be contributing to a public good, wherein other MFIs can learn from their study and learn how to target better and design better products so as to achieve their primary goal, poverty alleviation, more effectively.

Creating a control group in a longitudinal study does not necessarily imply impositions to operations. This author, for instance, is currently working on a longitudinal impact studies in urban South Africa, where the control group is randomly created and hence strong methodologically. The process is of little to no cost, and even a benefit, to operations. The strategy took advantage of the natural organizational limitations of a project as it entered a new area. Not all MFIs are in the situation to do what is necessary to conduct such a study, but if enough are, and the studies are conducted, then we as a community can learn more about whether MFIs can alleviate poverty, who we can help the most, and how we can best help them.

---

<sup>3</sup> Proper control groups are particularly difficult to create for microfinance impact studies since the entrepreneurial spirit of participants is presumably quite unique. Hence, merely finding “similar” individuals as a control group does not solve this problem.

## References

- Sebstad, J., Neill, C., Barnes, C., and Chen, G. (1995). Assessing the Impact of Microenterprise Intervention: A Framework for Analysis *AIMS Project Web Site* <http://www.mip.org>
- The SEEP Network. (1999). Tools Manual for Practitioner Impact Assessment, Draft. <http://www.mip.org>
- Ghatak, M. (2000). Screening by the Company You Keep: Joint Liability Lending and the Peer Selection Effect. *Economic Journal*, Vol. 110, Issue 465, July.
- Hatch, J. (1997). Personal conversations with John Hatch from FINCA International.
- Thomas, D., Frankenberg, E., Beegle, K., and Teruel, G (2000). “Lost but not forgotten: Attrition in the Indonesian Family Life Survey.” *Rand Working Paper DRU-2059*.