



**Assessing the Impact of
Microenterprise Services (AIMS)**

Management Systems International
600 Water Street, S.W.
Washington, D.C. 20024-2488
Tel: (202) 484-7170 \$ Fax: (202) 488-0754
E-mail: aims@msi-inc.com

**REVIEW OF METHODOLOGICAL
APPROACHES TO THE STUDY
OF THE IMPACT OF
MICROENTERPRISE CREDIT PROGRAMS**

June 1996

Submitted to:

Monique Cohen, Ph.D.
Office of Microenterprise Development
Economic Growth Center, Global Bureau
USAID

Submitted by:

Gary L. Gaile and Jennifer Foster
Management Systems International

This work was funded by the Microenterprise Impact Project (PCE-0406-C-00-5036-00) of USAID's Office of Microenterprise Development. The Project is conducted through a contract with Management Systems International, in cooperation with the Harvard Institute for International Development, the University of Missouri, and The Small Enterprise Education and Promotion Network.

TABLE OF CONTENTS

EXECUTIVE SUMMARYv

REVIEW OF METHODOLOGICAL APPROACHES TO THE STUDY OF THE IMPACT OF MICROENTERPRISE CREDIT PROGRAMS1

- I. INTRODUCTION1
- II. DESCRIPTION OF THE STUDIES3
 - A. Purpose of the Studies3
 - B. Specific Studies4
 - 1. Buckley's study of the informal sector in Kenya4
 - 2. Pitt and Khandker's study of credit impacts in Bangladesh5
 - 3. Lapar et al.'s study of selectivity bias in the Philippines7
 - 4. Buvinif et al.'s study of credit impact in Ecuador8
 - 5. Schuler and Hashemi's study of the impact of credit on women's empowerment and contraceptive use in Bangladesh8
 - 6. Nelson and Bolnick's study of credit impacts in Indonesia10
 - 7. Dearden and Khan's study of credit impacts on fertility in Bangladesh11
 - 8. Zeller et al.'s multicountry study of the impact of credit on food security11
 - 9. Pulley's evaluation of the Integrated Rural Development Program (IRDP) in India13
 - 10. Hulme et al.'s studies of credit impact in Sri Lanka and Bangladesh14
 - 11. Mustafa et al.'s impact study of BRAC's Rural Development Program15
- III. SAMPLE DESIGN AND EXECUTION17
 - A. Endogeneity and Selection Bias17
 - B. Sampling Issues19
- IV. TEMPORAL ISSUES21
 - A. Longitudinal Studies and Recall Methods21
- V. ANALYTICAL TECHNIQUES23
 - A. Econometric Modeling23
 - B. Quasi-Experimental Techniques23

VI.	CONTROL METHODS	25
VII.	LOCATIONAL CONSIDERATIONS	27
VIII.	ALTERNATIVE METHODS, QUESTIONNAIRE CONCERNS AND FIELD ISSUES	29
	A.	Alternative Methods 29
	B.	Questionnaire Concerns 29
	C.	Field Issues 30
IX.	CONCLUSIONS	31
	A.	Methodological Recommendations 31
		1. Sampling 31
		2. Temporal Issues 32
		3. Analytical Techniques 32
		4. Control Methods 32
		5. Other Issues 33

BIBLIOGRAPHY 35

ANNEXES:

Annex 1:	Variables Used in Studies Reviewed
Annex 2:	Sampling Frames in Studies Reviewed

EXECUTIVE SUMMARY

Purpose and Scope

The purpose of this desk study is to review some of the more rigorous assessments of the impact of microenterprise credit programs in order to inform the design of core impact assessments to be conducted by USAID's Assessing the Impact of Microenterprise Services (AIMS) Project. The review covers eleven studies carried out in Asia, Africa and Latin America. It focuses on sample design and execution, temporal issues, analytical techniques, and control methods for eliminating alternative explanations for changes. After discussing the ways in which previous studies have addressed each topic, the authors provide recommendations for the AIMS impact assessments.

Description of the Studies

Methodology was not the principal goal of most of the studies selected for review, yet, in each, considerable effort and innovation was devoted to methodological issues. Most of these studies were undertaken explicitly to evaluate one or more microenterprise programs. A wide range of variables were covered in the 11 studies. (These are listed in Annex 1.) A few studies centered on a limited number of impact variables, while others looked for impacts at both the household and enterprise levels. While all studies employed quantitative measures and techniques, a couple of them also used qualitative methods. Most were based on data that were collected more than once, but all within a 24 month timeframe. Almost all of them involved a comparison group.

Sample Design and Execution

Because of the issues of fungibility and selectivity bias, sample design and execution in microenterprise impact studies is complex and critical. Selection bias arises both in terms of the program clients and the location of the programs. The fungibility issue concerns the fact that financial and other resources, including credit and/or the profits from a microenterprise activity, may move between and among various household activities, making it difficult to track impacts. The review of studies shows that there is consensus that some form of quasi-experimental design is appropriate (assuming that an experimental design is not an option). Recommendations include:

- \$ sampling should occur with control groups from within program sites and a control population chosen from matched non-program sites;
- \$ statistically-equated control groups may be used for individual controls;
- \$ in program sites, eligible non-borrowers should be used as control groups; and
- \$ an overall sample size of about 500 should allow for effective use of control variables and for dealing with problems associated with longitudinal analysis.

Temporal Issues

Impact studies are very sensitive to temporal issues. The point at which impacts first begin to occur, and the length of time that impacts are sustained (as well as the rate of change) are subjects

of debate. There is consensus that longitudinal analysis is required. Recommendations include:

- \$ the research design should include a longitudinal study with an 18-24 month interval between data collection rounds ;
- \$ recall methods can be used to enhance the longitudinal profile;
- \$ seasonality should be a consideration in research design;
- \$ in-depth interviews may reveal "time lines" of credit impacts; and
- \$ there may be neglect of long term credit impacts.

Analytical Techniques

Quasi-experimental design coupled with multivariate statistical analyses are the predominant analytical techniques used in the studies reviewed. Econometricians have used these techniques as complements to econometric modeling. Econometric modeling has the advantage that it is readily generalizable, but also the disadvantages that rigorous assumptions, are required that cannot always be met, and such modeling has a restricted audience. Recommendations include:

- \$ multivariate techniques can control for selection bias and endogeneity issues;
- \$ choice of techniques should be a function of the type of data collected and their distributional characteristics;
- \$ an expanded list of variables should be covered, including social, contextual and locational variables; and,
- \$ data cleaning and checks on data validity should be part of the research design.

Control Methods

None of the studies reviewed successfully controlled for the fungibility of resources between household and enterprise. Selection bias also presents control complications. Linked with both sampling design and analytical techniques, recommendations on control methods include:

- \$ statistically-equated control methods are sufficient to address most control issues;
- \$ gender is a critical control variable;
- \$ continued efforts to control for fungibility must be made; and
- \$ control methods should be a function of the data available.

LOCATIONAL CONSIDERATIONS

Location is given minimal consideration in most impact studies, yet it plays a major role. The location of the program is a major determinant of success. The relative location of clients is likely to be important. Locational changes (e.g., road improvement) also have an impact on program performance. Finally, use of carefully paired, non-program locales as a control method will significantly improve methodological rigor.

OTHER ISSUES

Too little attention has been paid to alternative methodologies, such as qualitative methods and

counterfactual analysis. Similarly, such questionnaire concerns as survey fatigue and the need for back translation have received scant notice. Also, concern for issues related to gathering information in the field is rarely expressed in the studies. Finally, issues such as politics, favoritism, corruption, accountability, and leakages are rarely part of impact research design.

CONCLUSIONS

Several issues complicate selection of an appropriate methodology for studying the impacts of microenterprise program credit. Two predominate. The first is the issue of fungibility, since credit and other resources may be used for both enterprise and household purposes. The second is the issue of selectivity bias, since both the borrower and the lender "select" participation, which means that loan recipients are decidedly non-random.

Debate surrounds program evaluation methodology. Both quantitative and qualitative methodologies have been used, both have positive and negative aspects, and both have achieved acceptance. Valid evaluations can be achieved through a variety of approaches.

The papers reviewed for this study indicate that significant "norming" has occurred in the field of microenterprise program impact research on many issues, such as reducing selection bias and improving controls. Some issues, such as fungibility, remain problematic. Methodology "drives" some studies. More care must be taken to fully specify study objectives and to allow these objectives to dictate the types of data that are collected and the methodology that is used.

REVIEW OF METHODOLOGICAL APPROACHES TO THE STUDY OF THE IMPACT OF MICROENTERPRISE CREDIT PROGRAMS¹

I. INTRODUCTION

The purpose of this desk study is to further the design of the core impact assessments of USAID's Assessing the Impact of Microenterprise Services (AIMS) Project by reviewing some of the more rigorous assessments of the impact of microenterprise credit programs. This review describes the purpose of each study, the study population, the questions asked and/or hypotheses tested, and the variables included in the analysis. It then examines the design of each study in the following areas of inquiry:

- \$ sample design and execution;
- \$ temporal issues;
- \$ analytical technique; and
- \$ methods for eliminating alternative explanations for changes.

This examination identifies (i) the approaches taken, (ii) the rationale for (or strengths and weaknesses of) each approach, (iii) the key issues and challenges encountered by the researchers in using the approach, (iv) important lessons for the design of the AIMS core impact assessments, and (v) recommendations for the design of the core impact assessment, where feasible. This study also identifies issues to be included in the AIMS exploratory field studies.

The AIMS Project seeks to gain a better understanding of the processes by which microenterprise program services strengthen businesses and improve the welfare of microentrepreneurs and their households. This will be achieved through a set of desk studies, field focused research, core impact assessments and use of tools developed under the project to monitor the impact of microenterprise programs. The core impact assessments (CIAs) shall be methodologically rigorous longitudinal assessments of three different programs. The CIAs are expected to mix qualitative and quantitative methods. It is anticipated that the quantitative data will be collected during two main rounds, with a two year interval between the rounds. The design of the CIAs will be informed by exploratory field studies of the selected programs and the recommendations in this paper as well as the other AIMS desk studies.

This review and the recommendations are also meant to inform others interested in assessing the impacts of microenterprise credit programs. It covers key methodological issues in the design of rigorous impact studies and the analysis of the data.

¹The authors of this study would like to thank Carolyn Barnes, Peggy Clark, Monique Cohen, Mark Pitt, Muneera Salem-Murdock and Jennefer Sebstad for their useful comments on an earlier draft of this study.

II. DESCRIPTION OF THE STUDIES

A. Purpose of the Studies

Few of the studies selected for review had methodological development or explication as their primary purpose. Yet, in many of them, considerable effort and innovation was expended on methodology. Studies which used the most rigorous methods to assess the impact of microenterprise programs were selected for inclusion in the review paper.² Primarily, these studies were commissioned by development agencies to evaluate the impact of specific microenterprise credit programs. An interesting component of a few studies was to assess the social impacts of these programs, explicitly their impact on family planning practices.

In most of the studies, there are very clear statements of some of the methodological components, especially in the areas of sampling procedures, variables used, and statistical procedures. Less detailed is the reporting on actual field methods, questionnaire design, recall periods, and locational considerations.

Since most of these studies were explicitly undertaken for the purpose of program evaluation, there is a surprising lack of clarity about which results are being used to determine attainment of program objectives. Is program success measured by:

- \$ the client households' increased overall well-being; or
- \$ clients' business success; or
- \$ the community's improvement; or
- \$ the program's financial viability (e.g., payback/default rates)?

Given the widely acknowledged fungibility of microenterprise credit, the propensity for confusion between the first two aspects noted above is particularly important. A first lesson for the AIMS project is to be quite explicit about the criteria that will be used for evaluation.

Program evaluation methodology has been subject to much debate. From the 1950s and through the early 1970s evaluation work was primarily quantitative. Econometric modeling clearly comes from this perspective. But by the mid to late 1970s, numerous authors made cogent arguments for qualitative or subjective historical case-study approaches. Both approaches have positive and negative aspects, both achieved acceptance, and valid evaluations can be achieved through a variety of approaches.

This desk study will now briefly present an overall review of each study, focussing on its methodology. These reviews will be followed by more in depth comparative analyses of specific methodological components.

²Unfortunately, this did result in a distinct bias towards studies from Asia.

B. Specific Studies

1. Buckley's study of the informal sector in Kenya

Similar to the studies of Pitt and Khandker (1994) and Schuler and Hashemi (1994), Graeme Buckley (forthcoming) focusses on more than one microenterprise program that provides credit to the microenterprise sector in Kenya. The two credit programs were chosen because they were both deemed successful, yet "employed dissimilar approaches and very different lending procedures" (p. 421). The first program, Juhudi-Kibera, is the program of a non-governmental organization (NGO) that targets individual microentrepreneurs, who are "essentially involved in 'survival economies'" (p. 383). It is a group based lending program, which means that a borrower must be part of a group of five microentrepreneurs, in order to receive a loan (which is paid by bank check). The second program, the Informal Sector Program (ISP), is a parastatal that targets "licensed and viable businesses", i.e., more established entrepreneurs. Unlike Juhudi-Kibera, the loan goes directly to the individual entrepreneur in the form of a purchase order. The client populations for the two programs also differ widely. For the first program, the clients all worked in Kibera, a large spontaneous settlement in Nairobi. For the second program, clients surveyed were from four distinct districts in Kenya and were dispersed in towns and rural trading centers. The purpose of Buckley's study, which is part of a larger program evaluation, was to assess the impact of credit on clients' businesses and households. Buckley does not test formal hypotheses, but rather evaluates these programs in two broad areas: institutional efficiency and the socio-economic impact on the business that is financed compared with non-borrower groups.

A highlight of the methodology is the collection of a wide variety of supportive information from sources other than clients and the non-client control group. Structured and unstructured interviews were conducted with lending institution staff, donors, government officials, bankers, and alternative lending institution staff. Secondary data sources were also employed. This information adds richness to a study whose impact would otherwise be constrained by its small sample size.

Both studies used a quasi-experimental design with non-client control groups. The primary data were derived from structured, approximately hour-long questionnaires that were administered to stratified random samples of borrowers servicing their first and second loans. For ISP, these two groups consisted of borrowers who had either completed payment of their first loan or had been paying on their loan for over 12 months, and borrowers who were paying off their second loan. Two control groups were also used: members due to receive a loan, and a non-borrower entrepreneur group.³ This sampling scheme enabled the author to measure the impacts over time. One problem with this comparative study is that different sampling

³The ISP scheme divided new borrowers into two groups: those who were either waiting for their loan or had received their loan in the last three months, and those who had been making payments for more than three but less than twelve months.

stratification procedures were used in each case.

The surveys relied on a before and after method, with the before data based on recall. This yielded 50 variables classed into six categories which elicited background data and attempted to assess credit impacts on both business operations and households. The analysis reports complications due to the fungibility issue. It specifically notes that respondents had problems distinguishing between business and household assets and, in many cases, between business and household expenditures. This difficulty implies that the loans are being used for a variety of purposes and therefore it is problematic to assess the impact of program credit only in terms of its stated purpose.

The findings from the study are mixed. For businesses, there was a positive impact on sales, a limited impact on employment, little impact on assets and a mixed impact on profits. At the household level, increases in expenditure exceeded the mixed effects of increases on income.

Finally, Buckley concludes that the impact is "probably more dependant on the abilities, aptitudes and attitudes of individual borrowers than any particular feature of the respective lending programmes" (p. 418). Therefore, it would seem to be important for the AIMS core impact assessments to control for individual characteristics in order to evaluate the true impact of credit and its possibilities for assisting an increasing portion of the population.

One of the limitations that was recognized and needs to be considered as part of the development of methodology for the AIMS Project is the lack of availability of longitudinal data which the author argues is essential in order to understand the *real* impact of credit. This will be an issue especially in terms of evaluating newer programs and their impacts. Also, a broader interpretation of assets needs to be considered. As noted above, because of fungibility, the respondents have difficulty differentiating between business and household assets and expenditures.

The lesson for programs in general is the importance of product targeting, both in terms of context and product need. Therefore, an assessment of products, i.e., types of loans available, would assist in understanding impact as well. Finally, Buckley's data indicate that repayment rates vary widely by region.

2. Pitt and Khandker's study of credit impacts in Bangladesh

This study combines quasi-experimental design and econometric modeling to assess household and intrahousehold impacts of three credit programs in Bangladesh (including the Grameen Bank). The authors note that previous studies have not provided a comprehensive assessment of impact that takes into account issues of endogeneity and self-selection (see section III. A). A stated methodological intent is "the importance of our attentiveness to endogeneity in evaluating these credit programs." They note that comparing their econometric method with "more naive approaches" shows the importance of their attention to these factors.

One flaw of this and other studies is that they fail to take into account the types of credit

programs being investigated. Pitt and Khandker simply say the three programs are "similar". In fact the three programs are quite different and include not only credit provision, but other support activities that could substantially affect impacts. Differences in type of program and service provision should have been included in the research design.

The research design used a highly stratified random sampling technique in order to test for the impacts of each of three credit programs and to explore gender differences. Surveying 87 villages, the authors used a sampling method that insures comparability in terms of control populations, both of villages and individual households. The justification of their stratification technique is detailed. The technique and subsequent methods are sensitive to selection bias.

A large sample of 1798 households were surveyed three times over a crop cycle year in order to cover seasonal variations. A sub-survey on the nutritional status of children was carried out twice, in both peak and slack periods. Finally, a village survey was undertaken three times to match the cropping seasons.

The survey data provided the empirical basis for parameter estimation in the model which was accomplished by using a Weighted Exogenous Sampling Maximum Likelihood-Limited Information Maximum Likelihood-Fixed Effects (WESML-LIML-FE) statistical procedure. The econometric modeling is based on an efficiency argument. The modeling itself is logically consistent. Another challenge they faced was identifying "valid instrumental values" to use in their equations.

The authors do not explicitly state formal hypotheses, although their results are based on statistical testing. The results allow the authors to list numerous "outcomes of influence":

- \$ school enrollments by gender,
- \$ labor supply of women and men,
- \$ asset holdings of women and men,
- \$ consumption,
- \$ recent fertility,
- \$ contraceptive use, and
- \$ anthropometric status of children.

The authors make the claim that "we find that credit is a significant determinant of most of these outcomes." Furthermore, Pitt and Khandker argue that these positive effects are stronger if credit is provided to women. Finally, the authors note that the measurement of program impacts is influenced by the non-random placement of programs and the self-selection of participants into the programs (two factors that they pay careful attention to [see section III A.])

From a methodological standpoint, one might question whether the econometric model is necessary. While conceptually elegant, the impact results might be more convincing through the rendering of simple statistical results based on the quasi-experimental design, given that readers might rightly question the degree of fulfillment of required assumptions in such a complex model. For example, take Pitt and Khandker's statement:" Our approach is to use an asymptotic

bootstrap estimator of the covariance matrix which is essentially White's (1980) heteroskedasticity-consistent covariance matrix estimator in which the outer-product of the derivatives of the log densities (commonly known as the Berndt-Hall-Hausman or BHHH estimator) is defined so that the log density contains the full set of observations for any household or household member" (p. 25). Not only is such statistical manipulation befuddling to most readers, even schooled statisticians begin to question the myriad of assumptions that the manipulations of asymptotic bootstrapping, derivatives of log densities, etc., require.

In terms of the AIMS assessment, the central question is whether the project would like to develop a predictive econometric model that is generalizable to credit impacts throughout the developing world, or to develop a robust statistical testing procedure, probably using quasi-experimental design techniques, that can be used to evaluate impacts of a wide variety of microenterprise credit programs. It seems clear that this latter approach is more in keeping with the goals of the AIMS Project.

3. Lapar et al.'s study of selectivity bias in the Philippines

This study has methodological importance since it explicitly attempts to address the issue of selectivity bias. Further, it claims to "allow for the separation of the true credit effect from the effect of observable and unobservable characteristics of borrowers and nonborrowers." The model takes an unusual approach by assuming selectivity bias, and then testing for it using a "switching regression" technique.

Are borrowers intrinsically more entrepreneurial? If so, are better outcomes to be expected due to their personal attributes (versus the effect of the credit itself)? The authors use a selectivity model "to disentangle the effects of credit from the effects of observable and latent attributes of the borrowers vis-a-vis non-borrowers" (p. 3). The hypothesis tested is that of a "positive credit effect" (p. 13).

The "switching regression" operates based on a signal of whether the respondent is a borrower or a non-borrower. The argument is that borrowers have systematically different attributes than non-borrowers. The authors have specified a sorting process using a profit model which sorts by credit status. Further, the authors attempt to partition the independent variables into observable and unobservable characteristics. Of interest to the AIMS Project is the survey of borrowers in the Philippines which investigates variables that are heavily business-oriented.

The methodology is laden with procedures that are possibly used to force desirable results. As compared to Pitt and Khandker's more reasoned modeling efforts, there is little explication of why the loan amount is a quadratic expression, or why the second stage variables add "amount of loan received squared," or whether the assumption that "the error vector...is distributed multivariate normal with zero expectations and positive definite covariance matrix..." (p. 6) is reasonable.

Despite the desperate methodology, the result found is useful (if truly believable). The authors claim there is no statistically significant selectivity bias ("borrowers experience no

additional returns from observable characteristics") and that "borrowers experience no additional returns to their unobservable endowments and attributes." A further result is that there is a "23%" positive credit impact on output for the average borrower.

As noted above, the claim that individual characteristics can be separated out in a model is potentially useful; especially in light of the conclusion provided by Buckley (forthcoming). However, a greater understanding of the research methodology would be required before incorporating it into the AIMS impact assessments.

4. Buvinif et al.'s study of credit impact in Ecuador

This study is unusual methodologically because it distinguishes between microproducers and microvendors. Using a quasi-experimental design, the study separates the sample into eight groups. Borrowers and nonborrowers are subdivided by gender and type of activity. Because the study includes two sets of interviews a year apart, the final study separates the sample into sixteen groups. However, the authors note that because of various restrictions, data were not collected on male microvendors - nonborrowers for two different years, leaving two cells empty.

The stratified sampling procedure used a non-randomly selected control group and the stratification of the sample was not proportional to the population. Also, because the authors were concerned about the impact of the holiday season on economic data, interviews were scheduled to avoid the holiday season. Difficulties were encountered that affected the longitudinal analysis: 14.8% of the street vendors were "lost" between surveys. Therefore, what began as a relatively even borrower vs. non-borrower design became biased due to "lost" non-borrowers. The questionnaire contained 63 questions with precoded alternative responses and some open questions. It covers impacts on both the household and the firm.

Formal hypotheses were not stated, although it is clear that the purpose of the study is to "analyze the impact of the credit on these dependent variables in the different groups" (p. 235). The statistical methodology yielded a "barely visible" result of the impact of credit. The authors argue that this could be due to a small sample, the sampling method, and the "inability to measure invisible economies." Statistically, the authors engage in highly questionable behavior in their analysis by eliminating extreme values in order to achieve statistical significance. Further, some of their conclusions are admittedly risky, since they are based on findings that are not statistically significant.

However, regardless of the statistical limitations of the study, the focus on gender is an important element. Several studies emphasize the importance of evaluating program impacts using gender as a variable. Therefore, it would be advisable for AIMS to collect gender disaggregated data. Also, because women have a "twofold responsibility," indicators of impact, such as hours worked and increased efficiency, may need to be interpreted differently based on gender. Although the Buvinif et. al study stratified the sample by gender, this approach has certain limitations. The authors note that the number of women represented in their study were *not* proportional to women in the population as a whole. Finally, the authors make an argument for measuring the impact on employment by looking at job maintenance instead of job creation.

This is an important distinction for AIMS to consider when deciding on which variables to include in its impact assessments.

5. Schuler and Hashemi's study of the impact of credit on women's empowerment and contraceptive use in Bangladesh

This unusual study focuses on the social impacts of the participation in credit programs. It hypothesizes that women credit participants are more empowered, and therefore have higher contraceptive use than nonparticipants. The study also tests a hypothesis that participation in credit programs increases the diffusion⁴ of contraception to nonparticipants in program villages.

The study uses a quasi-experimental method and a two-stage clustered sampling technique. First a random sample of branch offices is taken and then a random sample of participants from the selected branches. This resulted in a sample of 1248 respondents, of whom 1045 were married women under the age of 50. Finally, a control group was selected from other non-program villages based on the size, region and population density of the village. There may be some bias in this part of the procedure since villages in which there was an "intensive presence of credit or family planning programs run by NGOs were not included." This aspect of non-randomness destroys some of the comparability of the study, "stacking the deck" a bit, as it were.

An initial survey was conducted in 1991 and a very successful follow-up survey was carried out 18 months later (96% of respondents were located and reinterviewed).⁵ A second comparison group (n= 315 married women <30) was added at the time of the second survey in order to explore the hypothesis on the diffusion among non-participants in program villages. A complicated index creation activity is part of the research design, in which several variables related to women's empowerment were first identified and then classified into eight "empowerment indicators." Women were classified as "empowered" if they had a positive score on five or more of these indicators.

The study explicitly addressed selection bias which is expected in contraceptive use/credit relationships. The authors admit to being less successful in controlling for village selection bias, noting the clear locational biases of the credit programs themselves in selecting villages for participation.

Logit regression models were used to control for background characteristics of the women in each of the groups. In the regression analysis that uses only program participants, the

⁴*Diffusion* is perhaps best understood as the *spread* of innovations, often over contiguous geographical areas. It forms the basis for such institutions as agricultural extension. A diffusion impact exists, for example, when a household knows an innovation has been adopted successfully by a neighbor, and adopts the innovation itself. Contraceptive use is a classic example of diffusion.

⁵Unfortunately the authors did not share the secret of how such a high response rate in a longitudinal study was achieved.

control variable "used contraception before joining the program" is included.

The study also used ethnographic methods and in-depth case studies to attempt to understand the "change processes" of contraceptive adoption, especially when analyzing the within-village cases. "Women who had departed from prevailing norms and conventions were interviewed to identify the circumstances and the forces behind these behaviors...(such as contraceptive use and non-traditional work for women)" (p. 67).

All of the hypotheses relating empowerment, credit and contraception yielded positive, statistically significant results. This is a straightforward, careful study based on sound methodology.

An interesting note for the AIMS assessments is the attempt by Schuler and Hashemi to quantify the diffusion impacts of the program (the locational "multiplier effect", i.e., the benefits that accrue to other community members who are not participants). This is what Buvini^f et al. (1989) called the "invisible economic benefits" that accrue to the areas in which these programs are located.

6. Nelson and Bolnick's study of credit impacts in Indonesia

Methodologically, this is an important study since the authors spend considerable effort to address the issue of selection bias in their research design. The selection bias issue will be dealt with in detail in a following section. Formal hypotheses are not stated, but the purpose of the Nelson and Bolnick study is to measure the direct and indirect impact of credit (including additionality and productivity, largely within specific enterprises). A secondary objective is to measure customer characteristics which are significantly associated with favorable impacts.

Nelson and Bolnick use an innovation usually found in policy sciences; the "but for" argument. "But for" arguments ask whether changes can be attributed to the intervention itself, or would have happened anyway. In that context, the authors attempt to use their design to test "the extent to which investment can be attributed to credit" (p. 9). An additional technique used is to treat these credit programs as innovations which diffuse; thus, a credit program is "an exogenous institutional innovation that had not yet been diffused throughout the class of eligible recipients" (p. 9). Nelson and Bolnick also make a strong, cogent argument for clarifying the objectives of the research before designing and implementing the survey. They also emphasize the importance of ensuring that the areas to be studied are truly comparable. The authors carefully consider timing effects and note a variety of time factors that can affect the study (gestation lag, secondary effects, and noise). They argue that replication of the study over time increases the robustness of the conclusions. The reliance on recall data is discussed in detail and the authors make suggestions for questionnaire design which facilitate more accurate recollection.

Their sampling design employed a control group that was eligible for a loan in 1980, but did not receive a loan until 1982. This group was compared to the group that was eligible in 1980 and *did* receive their first loan in 1980. This method of control avoids selectivity bias better than do most other quasi-experimental methods. There is still a bias, however, in that borrowers may

well report 'sound' status since they may not wish their creditors to have any concerns about their loans.

The authors use a regression method to estimate the change in the capital assets of borrowers. They structure the equation such that additionality can be estimated from the **b** value of the dummy variable (participation). Other target dependent variables (employment, output, or income) can be substituted into this equation, which uses a vector of baseline and entrepreneurial characteristics of the enterprise as its only other variables. The authors carefully detail some of the issues involved in variable specification, using the dependent variable of employment as an example.

They are forthright about actual problems in the field, such as the unwillingness of enumerators to make use of all available data. For example, in cases where the respondents had transaction records that covered the entire time frame of the survey, they found that the enumerators were unwilling to transcribe this large amount of data. This concern leads back to the importance of survey design to avoid both respondent *and* enumerator fatigue. They also include a full section on field survey implementation, largely addressing the hierarchical relationship between the researcher, supervisors and enumerators. Finally, useful suggestions on questionnaire design are elicited from field experience.

This study contains a wealth of methodological guidance. As noted above, of particular interest is the section on improving the quality of recall data which will need to be used in most cases in order to evaluate the "real" impact of credit. Also, careful attention needs to be paid to their discussion of "data cleaning," a process that significantly increased the cost per interview. The research design could be criticized only because it effectively sidesteps the household/fungibility issue. Nonetheless, it is transparent and very helpful in its discussion of methodological issues. The researchers even attempted to use disparate techniques such as counterfactual and anthropological methodologies as supplements, (although without much success).

7. Dearden and Khan's study of credit impacts on fertility in Bangladesh

This study, while interesting in its topical approach, provides little new methodological analysis. The sole methodological innovation is that data are based on "paid volunteers" visiting households on a monthly basis, so the time series data are substantially stronger than for most studies. Hypotheses were not formally stated, but it is clear that the purpose is to test the impact of credit provision on contraceptive use. They used a simple quasi-experimental design (members vs. non-members), tracking only women who were not using contraception at one point in time. This study ignores the diffusion possibilities within a village that was demonstrated in the study by Schuler and Hashemi (1994). Of interest to the AIMS assessments is the introduction of social class as a relevant and predictive variable. Results of the study show (similar to Schuler and Hashemi) a positive effect of credit participation on empowerment and contraceptive use, which varies by social class. This study, when compared to Schuler and Hashemi, demonstrates the importance of non-village control groups, as well as the importance of having longitudinal data in order to assess impact. The authors recognize that individual characteristics possibly play a role,

i.e., "women who join savings groups are qualitatively different--in their thinking and in their behaviour--from those who do not" (p. 10). Finally, region of residence is also introduced as an important factor, which illustrates the importance of recognizing contextual factors when assessing impact.

8. Zeller et al.'s multicountry study of the impact of credit on food security

Heavily econometric in design, this study focusses on household level impacts of credit provision and the development of methodologies for analyzing financial institutions (although the latter does not pertain to the scope of this study.) It tests whether credit constraints are present, whether there is a demand for credit, what are the determinants of participation, and what the effect of relaxing the credit constraint would be on household utility, consumption and production.

The study states numerous specific hypotheses, of which the following are especially important:

- 1) "...improved access to credit and savings services either augments factor income generated by the production process or reduces costs for smoothing consumption at sufficient levels" (p.15),
- 2) "The initiation and expansion of member-based financial institutions at the community level is hypothesized to increase the volume in loans and savings deposits in rural households without collateral, and change the structure of the market by reducing the share of high-interest informal lending" (p. 31),
- 3) "...the share of zero-interest informal lending by friends and relatives is decreasing with increased access to and diversity of formal financial services and with overall economic development" (p. 31), and
- 4) "...the optimal levels of consumption, production, and investment of a household that faces liquidity constraints are different from, and provide lower utility than, the optimal levels in the absence of credit constraints" (p.41).

The research uses two methodologies: 1) an econometric model, and 2) surveys of household perceptions. Some of the variables are innovative, for example, the econometric model relates credit constraint to weight fluctuations in young children.

One of the main challenges faced by the researchers was comparing the survey data for the nine case countries. Because some of the studies were done much earlier than others, the comparability of data at the household level was limited. Also, they recognized that there was a lack of data on gender differentiation. They noted that the differences in data led to the use of "less preferred methodological approaches."

The study used a "choice-based stratified" sampling procedure. "Choice-based" means

that program participation is clearly a choice made by individual members and thus endogeneity (see Section III.A) comes into play (the client chooses to participate in the program *and* the program chooses to allow the client's participation). In this sampling procedure, population proportions of participants (compared to all possible borrowers) are used as a stratification criterion (versus purely random sampling). It is clear this method was required to fulfill the parameter estimation needs of the econometric model being tested. As such, the surveys did not eliminate much of the selection bias that could have been eliminated using a different procedure (which would have not been as compatible with the econometric modeling needs). The authors note that the case studies use a small sample size (although the actual size is not given) because of a variety of factors: cost, sensitivity of the topic, the necessity of multiple rounds, and the depth of data required.

A common methodological sin is committed early in the research. Direct numerate household information is converted into dummy variables. It is almost never statistically appropriate to convert higher order (interval/ratio) data to nominal data. Not only do the data lose important information, but the statistical techniques which can be used with nominal data are much less powerful than those that can be used with ordinal or interval/ratio data.

The research design is interesting in that it attempts to target constrained borrowers. It identifies *discouraged* borrowers as those who could not offer the required security or thought that the procedure was too complicated and expensive while at the same time they did not rank "no need for credit" or "dislike of credit on religious grounds." It also classifies *rejected borrowers* and *rationed borrowers*. These three types of borrowers are combined into the class of *constrained borrowers*.

A range of interesting variables was covered in the surveys, including anthropometric measures, human capital variables, event variables, determinants of participation variables, and community level variables.

The statistical techniques were nested, staged regressions, using maximum likelihood estimators. The authors attempt to employ a systems approach⁶ with a heavily inclusive package of variables focussing on decision-making.

While the results validate much of the model, problems of sampling procedure, selection

⁶A *systems approach* emphasizes the way in which the interrelated units or objects of a system form a 'set' with common properties. This set displays organization resulting from causal, functional or normative relationships between the specific units. Systems analysts try to analytically and empirically specify the links and feedbacks that comprise a system. Systems approaches tend to be inclusive of all possible inputs and outputs and impacts on a system. Non-academic analogies abound--one can think of plumbing systems, or banking systems, or even global systems. The problem with most systems approaches is that, in an effort to attain inclusivity, rigor is lost. The quasi-experimental attempts to *control* for variables may be seen as effective measures to avoid the complexities of a systems approach.

bias, sample size, and the inability to address the fungibility issue diminish the methodological utility of this study.

9. Pulley's evaluation of the Integrated Rural Development Program (IRDP) in India

This World Bank study is of limited use in considering methodology. While it is clear that there was an extensive survey of clients to assess impacts, the sampling and questionnaire methodology is not explained in detail and not at all transparent. From reading the work, it is likely the author was not involved in the research design of the study, but rather used the available data. Various hypotheses were tested "regarding household characteristics, implementation features, and general economic conditions that contribute to productivity, and repayment objectives of IRDP" (p. 2).

From unexplained Appendix material, it is clear that two rounds of surveys (and a subsequent resurvey) were taken of old and new beneficiaries of programs, primarily focussing on assets. A simple regression model is the analytical core of the work. It tests which factors produce differences in productivity when the dependent variable is the ratio of net income from an asset in year four to the cost of the investment.

Not surprisingly, given the genesis of the study, the results show that there are some serious economic structural problems (capital subsidies and low interest rates) that bedevil the program. Thus the conclusion that "not all poor households are good users of credit or appropriate targets for self-employment approaches to poverty alleviation" (p. v) is an expected outcome of such research.

Unlike many of the other studies, this study contained several variables related to sustained access to credit and involvement of bank officials. For example, the number of visits by bank officials by borrower income level was included as a variable, as well as beneficiaries that obtained subsequent bank loans. For assessing long term impact, sustained access to credit and the movement into the "formal" banking system are important factors. One prominent point the author makes is the need to stress quality rather than quantity of loans. Programs that are more concerned with getting money out the door run the risk of both missing their target population and undermining the long term goals.

10. Hulme et al.'s studies of credit impact in Sri Lanka and Bangladesh

Two studies using similar methodologies of different programs in Sri Lanka and Bangladesh were authored by this research team under the auspices of the Institute for Development Policy and Management at the University of Manchester. In both studies, the team focuses on the impact of program credit on income and productivity, and (in Bangladesh) poverty.

The Sri Lankan study is less detailed, focussing on 151 interviews using a lengthy questionnaire, and using recall data from 12 to 24 months before the interviews. Secondary data were also employed. In this study, there was no attempt at a quasi-experimental design after it was discovered in the early stages of fieldwork that the credit activities in a village tended to

"diffuse downwards" to poorer classes over time. Conversely, in the Bangladesh study, which did use a control group, later borrowers were found to be better endowed than earlier borrowers, partly because concern had grown over repayments in the program. Both results raise concerns for using new borrowers as a control group.

In Bangladesh, a quasi-experimental design was used for a one-time survey. Care was taken to select a representative sample of research sites, seeking "vibrant" and contrasting "depressed" local economies. The research design used length of time since receiving loans as the distinguishing characteristics between three groups.

The intent was to capture both household and enterprise impacts. No formal hypotheses are presented. Respondents were asked to compare their present situation to the situation as it was in the month before they took their last loan.

The analyses are statistically unsophisticated, but probably appropriate given the limited sample and the likelihood of selectivity bias in the sampling procedure. As in the Sri Lankan study, the authors argue that credit has a differential impact depending on the borrowers initial financial position.

One lesson from these studies is the emphasis on context. The author's argue that social, economic, and political factors are important to the success of these types of programs. Therefore any study must be designed to take these external factors into account. Similar to the Mustafa et al. study, they note the importance of looking at the economic impacts of credit in the long term, i.e., impacts, as opposed to more short-term "effects".

11. Mustafa et al.'s impact study of BRAC's Rural Development Program

This study was commissioned "to gain a better understanding" (p. E1) of the impacts of the Bangladesh Rural Advancement Committee's (BRAC's) Rural Development Program (RDP) on the rural poor in light of the goals of the program: poverty alleviation and empowerment. This study had two objectives: (1) to gain a more extensive understanding of the socio-economic impact of RDP, in both quantitative and qualitative terms and (2) to assist BRAC in the development of its ongoing capacity to assess the social and economic impact of RDP, including identifying the most appropriate methodologies to assess different aspects of BRAC's impact.

Important to the study methodology were two factors: "a specific definition of poverty" and an "integrated approach to the assessment of poverty alleviation (p. E1)." The authors used several "key indicators" to assess poverty alleviation including: increased material well-being; reduced vulnerability to seasonality; increased economic security; improvements in women's status; and the development of village organizations and institutions. They make a point to distinguish between the definition of an impact in comparison to an effect. They are concerned with impacts which refer to "sustained structural changes in well-being."

They assessed credit impact using a set of hypotheses under three broad categories: a) village context accounts or economic dynamism of the micro-regions in which, b) the household

context sustains its existence with a certain level of human and material resources, prior to joining c) RDP which creates access to credit and other resources and services. The informal hypotheses attempted to identify "contextual variables that either on their own or in interaction with RDP inputs, are likely to determine levels and direction of impact, including life cycle factors, dynamics in the micro-regional economy, education levels, and the initial endowment of households" (p. 63).

The study used a combination of both qualitative and quantitative methods. First, a household survey was carried out using a pre-coded questionnaire. This survey was conducted on both BRAC and non-BRAC households and was done in two rounds in order to "capture seasonal variations." The stratified random sample included a total of 2250 households: 1500 RDP households and 750 socio-economically comparable non-RDP households. Second, in order to capture contextual factors, they constructed village profiles for 225 villages--150 covered by RDP and 75 without an RDP presence. These profiles were constructed using a "structured form" and based on information from "small groups of key informants." Finally, in order to obtain qualitative as well as quantitative data, they selected 15 village organizations (using random selection from the villages where the household survey was done) for case studies. For these case studies they used both "informal and a limited number of [rapid rural appraisal] techniques."

In their analysis, they differentiate between the impacts of credit on men and women. They further differentiate by using the length of time these households have participated in the BRAC program. They argue that in order to truly determine impacts, longitudinal data must be used.

In their evaluation, Mustafa, et. al., relied heavily on multiple regression and included a number of dummy variables in order to measure factors such as "local economic dynamism." They also use the metaphor of "critical mass" which is required in order to achieve a "measurable impact." The two primary factors include length of membership and amount of credit received.

One of the problems identified in determining impact is the problem of measuring the households' "initial endowment", i.e., the original conditions of the households. Initial endowment was assessed based on recall information and must therefore be treated with care. Because some households had participated in the program for over four years, memory of the initial situation could prove to be faulty. A second problem they encountered was finding that the respondents at times seemed to be providing the answers that they thought the interviewers expected: this occurred in discussions with men regarding the changes in women's status that occurred as a result of the BRAC program. Another problem that was identified emphasizes the importance of having targeted interviews. The interviews for the case studies lasted for two to three days, which led to "apathy" among the respondents. A variety of secondary sources were also used to collect additional information.

Overall, they found that there is "a consistent movement along the path to greater wealth and expenditure, according to loan size and membership age (E3)." They concluded that credit programs did have a marked impact on the poor, especially on those who entered the program with smaller endowments. Also, they found that women utilized the credit more efficiently than

men, thus leading to greater returns on credit and effectiveness at reaching the target population. With regard to contextual factors, they concluded that "there appear [to be] little or no difference between the villages where RDP is in operation and where RDP is not (E2)." They found a positive impact on the majority of the key indicators. The indicator that was least effected was that of institution building, which was acknowledged to be a "lengthy and complex process(E4)."

III. SAMPLE DESIGN AND EXECUTION

The table below is a general description of quasi-experimental methods that can be used to test for credit impact. Indeed, few of the studies utilized all the available options in the table, and most had some element beyond control in their research design.

Table 1. Generic Research Design for Quasi-Experimental Assessment of Credit Impacts.

	BEFORE	AFTER
Control Location	Control Group ⁷	Control Group
Project Location	Control Group	Control Group or New Borrowers Group
	Borrowers Group	Borrowers Group

The need for experimental control has plagued studies of the impact of microenterprise program credit since their inception. Research design attempts to avoid selection bias have remained imperfect, despite heroic efforts. While substantial progress has been made, effective control and avoidance of selection bias remain problematic.

Early studies of program credit impacts tended to use simple before and after surveys, often based on recall. As it became apparent that control was necessary, quasi-experimental designs were used which were of the participant vs. non-participant variety. This was later refined to use time-staggered respondents so that new borrowers (for whom the credit impact could not yet be expected) served as the control group to compare with recurrent or former borrowers. As evolving research demonstrated that the credit impact was not just on the individual and the household, but also had some diffusion effect to the whole community, the concept of control locations was introduced. An argument is easily made that a control group in a community without a credit program may better match the borrowers group in a community with a program.

A. Endogeneity and Selection Bias

Endogeneity occurs when changes in the explanatory (independent) variables are caused in part by the dependent variable. In the study of program credit impacts, for instance, it is possible to make the argument that entrepreneurial behavior causes increased output, as well as make the

⁷The term *control group* is generic and may involve randomized assignment (in experimental design), matched pairs, or non-random assignment.

argument that increased output causes entrepreneurial behavior. Any time the causal link between two variables can go either way, endogeneity exists. Perhaps the best way to deal with endogeneity is to argue that it is due to bias of a variable which can be included and used as a control, as in Zeller et al.'s (1996) study in which the determinants of participation were included in the analysis. Pitt and Khandker (1994) also explicitly control for endogeneity in their work.

Selection bias occurs whenever a sample is not impartially chosen, whether knowingly or unknowingly. If likelihood of success is a criterion for admission to a credit provision program, than measuring a program's success by comparing admittants and non-admittants is highly susceptible to selection bias. Several other varieties of selection bias afflict credit impact studies. Both dependent and independent variables are subject to bias, as is the selection of sample programs and sites.

There is often an admitted selection bias in terms of programs. Buckley (forthcoming) admits his institutions were chosen "because they were examples of innovative, well-organized microenterprise/small business lending programmes that have achieved apparent success" (p.421). This is a recurrent theme in the studies surveyed, and is vigorously argued by Rossi and Freeman (1989). A strong argument could be made for a more random selection of institutions in order to gain insight into credit impacts of less successful (or even failed) programs. But this brings up the initial question of whether the focus should be on the clients (the case in all the studies reviewed) or on the program.

Site selection has a highly predictable bias--accessible locations. Schuler and Hashemi note "Our impression is that a general tendency exists for Grameen Bank, BRAC, and other development programs to avoid the least accessible areas of the country when setting up branch offices" (1994, p. 69), and that proximity to a branch office appears to be the greatest determinant of whether program activities are initiated in a village. Conversely, there are complex locational issues in program targeting. On the one hand, programs attempting to alleviate poverty, objectively, should seek poorer locales. On the other hand, the desire for achieving program success and viability argue for less poor locations.

Most of the works reviewed are not explicit about how locational sampling decisions were made. They also often fail to mention the presence or absence of possibly competing financial institutions. Locational access variables are not included in any of the works surveyed, with the notable exception of Mustafa et al. The latter use locational variables in their consideration of contextual variables, and assign villages a score based on a variety of factors including distance of nearest bank, distance to nearest haat/bazaar, and distance from nearest all-weather road (p. 17). One could argue that, other things being equal, the impact of credit on a borrower would be diminished if that borrower had less access to the overall economy (although this argument could

⁸While many are familiar with the problems of monopoly (one seller dominates the market) or monopsony (one buyer dominates the market), few are familiar with the very common developing country problem of *spatial monopolists* and *spatial monopsonists*. These problems are highly prevalent especially in areas of relative isolation, difficult access or limited markets. A *spatial monopolist* is the only seller in the accessible region. For

be reversed in the case of spatial monopsonists or spatial monopolists⁸).

Few organizations offering credit randomly select their clients. Preferred clients are those predicted to have the most favorable financial outcomes, and are understandably different from the set of all potential clients. Mediating factors, such as entrepreneurial skill, may not always be directly measurable, and thus it may not be possible to eliminate selection bias by using a control variable. A common sampling strategy, as noted above, is to use new borrowers as a control group. The argument is that these borrowers have similar attributes profiles to previous borrowers, and are thus readily comparable. However, Hulme et al.'s (1994) study argues that later borrowers were *poorer* than earlier borrowers and not a good comparison group. As noted earlier, their 1995 study in Bangladesh found that later borrowers were *better off* than earlier borrowers, which again limited the usefulness of new borrowers as a control group.

The preferred strategy for minimizing participant/non-participant selection bias is probably to choose a non-program site that is carefully matched with a program site on a number of community attributes. Further, careful matching of the availability of alternative sources of credit should be made. Then, from a random sample from the non-program village, a statistically equated control framework can be employed. This combines a matching strategy at the community level, with statistically equated controls drawn from a random population in the control location. The key to controlling for selection bias is to find a closely matched control location without the program and to survey enough individuals in that control location to allow for statistically equated controls (or possibly even matching) of entrepreneurial talent.

B. Sampling Issues

Annex 2 documents the complete range of sampling strategies used in the studies surveyed. Models for sampling can best be taken from the studies of Pitt and Khandker (1994) and Nelson and Bolnick (1986). Both studies are very explicit about sample design and some of the rules they used (e.g., the classic $n=30$ in the latter study). However, in spite of their care, in many cases the sample size for Nelson and Bolnick fell below 30, which implies the need for a larger initial sample size.

Determination of the appropriate sample size is a time-honored question, and the standard minimum of $n=30$ in any category remains the limiting norm in statistical testing. It is clear, however, that authors who use this minimum sample size often have difficulty establishing statistical significance. There are considerable gains to be had from increases in sample size,

example, if a villager needs to purchase charcoal, and there is only one dealer locally that sells it, that dealer can make extraordinary profits by raising the price to the point where the villager must make a choice between making a longer journey or purchasing the charcoal locally. A *spatial monopsonist* enjoys similar profit advantages from the same market imperfections caused by isolation, difficult access or limited markets. This is the buyer (often a wholesaler) who is the only purchaser of a good or service in the area. Again, prices are distorted by such spatial market imperfections which are highly prevalent in developing economies.

especially when there are explicit attempts to control for endogeneity and other selection biases through the statistical method. If one compares the studies of Mustafa et. al (1995) and Schuler and Hashemi (1994), which had overall sample sizes of >1000, with the studies of Buckley (forthcoming), Hulme et al. (1994 and 1995), and Nelson and Bolnick (1986), which were constrained to acceptable, but minimal, sample sizes, the effect on results becomes apparent. Each control variable used increases the minimum sample size that is required. In some cases, if resource constraints made it necessary, opportunity samples could be used, although with a sacrifice of control. A sample size of approximately 500 per case should allow for ample use of controls, some drop out of respondents, and interview error loss. Ultimately, however, determining *optimal* sample size is a budgetary question, especially when longitudinal studies are planned.

In econometric methods, the need for accurate parameter estimators leads most authors to use a choice-based sampling strategy, acknowledging for the endogeneity of the participation process in credit programs (see the discussion of this in Zeller et al. above). Most of the studies reviewed use some form of stratified sampling technique, with the stratification criterion being determined as a function of the objectives of the study. For instance, Buvini~~f~~ et al. (1989) stratify based on whether the clients are microproducers or microvendors. Buckley (forthcoming) stratifies based on number of loans and uses two control groups. Pitt and Khandker (1994) discuss in detail some of the logic used in stratification for analyses of credit impacts. Trade-offs can clearly be made between using control variables and sampling stratification.

IV. TEMPORAL ISSUES

As with sampling, the time and financial resources available will be important factors affecting how temporal issues are ultimately resolved. The studies surveyed ranged from use of monthly interviews by paid volunteers (Dearden and Kahn, 1994), which provides a wealth of longitudinal information, to the severely constrained one-time interview using recall for temporal perspective (e.g., Buckley, forthcoming). It should be noted that longitudinal survey data and recall data are comparable and compatible, only differing in terms of reliability.

A. Longitudinal Studies and Recall Methods

Although it is clear that to study impacts, longitudinal data (i.e., where data on pre-loan condition are collected at the time of the first loan) are preferable to recall data on original condition, both may be used effectively in combination. Nelson and Bolnick (1986) provide a solid discussion of time factors in research design. The first issue is gestation lag--how long does it take for a loan to begin noticeably paying off to the client? The authors argue that a period of greater than one year is minimal. The second issue is "secondary" effects, such as the reinvestment of incremental income. Here the temporal issue becomes more muddled. The third issue is noise from other time-related events (e.g., an economic downturn) which the authors argue is unavoidable.

Pitt and Khandker (1994) bring the issue of seasonality into consideration. It is clear that households, especially rural households or village businesses whose clients are rural households, have high seasonal fluctuation in income and expenditures. This reality is seldom considered in other studies and should be explicitly managed in the methodology used for the AIMS impact assessments.

Given likely constraints, longitudinal surveys with an 18 month interval which is sensitive to seasonality and includes recall (especially of the 6 months prior to control for seasonality and also of the 18 month period prior to give longitudinal depth) should serve the AIMS project's objectives. It may be wise, however, to do pilot analyses involving intensive interviewing in order to determine a "time line" of probable credit impacts. It may be possible, after these interviews to better determine an optimal temporal interval to use to assess impacts.

Mustafa et al.'s (1995) study explicitly notes the problems respondents had when asked to recall conditions four years previously. Nelson and Bolnick (1986) make numerous useful suggestions about the "well known problems of recall data" and suggest steps to minimize them. They address issues of questionnaire design and interview techniques, focussing on methods to facilitate more accurate recollection. They include a variety of cross-checks in the questionnaire and inquiries designed to "freshen memories."

V. ANALYTICAL TECHNIQUES

The studies reviewed used a variety of analytical approaches and techniques, often in combination. The dominant analytical approach used was some form of quasi-experimental design. At times these were combined with econometric modeling approaches (e.g., Zeller et al., 1996; Pitt and Khandker 1994).

A. Econometric Modeling

Econometric modeling has the distinct advantage that it is readily generalizable. Because assumptions made in this modeling are rarely case specific, it would seem to be ideal for comparative analyses. Yet, with the advantage of generalizability come attendant drawbacks. The rationale for most technique application in these modeling approaches is to adopt control mechanisms and to obtain parameter estimates. In order to appropriately use these techniques, however, considerable methodological restrictions must be applied. For example, Zeller et al. note that "standard estimation methods that assume random sampling or exogenous stratified sampling yield inconsistent parameter estimates if applied in this context" (1996, p. 62). As with all statistical modeling, rigorous assumptions are required (although not always adhered to), for example, "by assuming that the error vector...is distributed multivariate normal with zero expectations and positive definite covariance matrix..." (Lapar et al., 1995, p. 6).

A further drawback to econometric modeling is its restricted audience. Not all users of the AIMS impact assessments will be schooled in econometrics. The use of production functions, utility, and other econometric standard concepts may be unfamiliar to some, and thus off-putting. While technically more difficult, many of the generalizability advantages of econometric models can be obtained through careful application of quasi-experimental designs which utilize many of the control aspects of econometric modeling, without the assumptions of the models themselves.

B. Quasi-Experimental Techniques

Quasi-experimental designs do not assume certain behaviors in the surveyed population as do many econometric models. Rather, they analyze "uncontrollable" situations in an experiment-like fashion by using "control" populations and by using statistical procedures on control variables. They have become the norm in impact studies of microenterprise credit programs. An interesting method in credit impact analyses is to use new (or soon to be) borrowers as the control group for comparison with an experimental group of existing borrowers who have been program participants for at least a year. (Note however, as explained above, that this method has been called into question.)

A wide variety of multivariate statistical techniques can be used with quasi-experimental design. In the works surveyed, a common flaw was that a particular statistical methodology often "drove" the data analyses. For example, ratio data were rendered into dummy variables, as was mentioned above, so that beta-values were more interpretable.

It is strongly recommended that the data available prescribe the statistical methods to be used. It is usually beneficial to use the most powerful technique available (in the statistical sense of "power"). Once the questionnaire has been designed, the likely form of the variables will be known. Not, however, until the data undergo preliminary analyses will such questions as whether the variables meet distribution assumptions be answerable. Operationally, after the data are gathered, coded, and entered, descriptive statistics will yield the information necessary to determine their distributional properties. It is at this time that the most powerful technique, given data constraints, should be chosen.

Ordinary least squares regression techniques will likely be most useful in analyses of most quasi-experimental data. Two-stage least squares regressions are appropriate for incorporating controls and dealing with endogeneity. The studies surveyed ran the gamut of techniques, from simple cross-tabulations to switching regression, logit and probit models, and a variety of OLS and MLE procedures. Again, however, it is inappropriate to specify a technique until the form of the data analyzed is known and the distribution characteristics have been tested.

Annex 1 gives a comprehensive list of variables that have been used in the studies reviewed. This list is impressive in its diversity, especially in the area of social impact variables. Given the variety of foci of the studies reviewed, the panoply of variables is understandable and expected. The choice of variables should be made after determination of what type of credit impacts are being studied. Is the focus on the individual, household, or enterprise level, or some combination of these? Are the concerns largely economic or do they branch out into social considerations of food security, family planning, or other development-related issues? Are specific control variables (e.g., gender) needed in the design of the study?

Finally, a substantial effort was made in a few of the studies to "clean" and validate data before analysis. This was notable in Buckley (forthcoming) and Nelson and Bolnick (1986) who stated that data cleaning took "three person-years." If the person analyzing the data has a chance to code and enter enough surveys to gain experience with the data, s/he will have much greater ability to choose proper analytical techniques. Anyone with experience in this area quickly gains a sense of data validity after only a brief exposure to a sample of the data. For example, in one of the author's (Gaile's) recent survey work, he came upon a farmer who grew twenty onion trees, according to the enumerator. This clue made Gaile scrutinize that enumerator's work more closely and he realized that none of it could be considered valid. Catching these kind of problems in the field can save much time and money later. Nelson and Bolnick (1986) used a computer program to trap logical errors and check for consistency between responses. Similarly, in a recent survey of market use, Gaile asked three questions: distance of home from market, time to get from home to market (by mode of transport), and the location of the respondent's household on a map. By triangulating these three sources of data, it was possible to determine and eliminate invalid responses.

VI. CONTROL METHODS

Quasi-experimental methods rely on statistically-equated control methods to eliminate alternative explanations for changes. Matched control methods are also appropriate, but are likely to be more cumbersome, more expensive and require a larger sample size.

While most of the studies reviewed used a control group, not all studies used control variables in a multivariate analysis to attempt to eliminate alternative explanations for impacts. Those that did often used a two-stage approach, using a variety of control variables in the first stage. Lapar et al. (1995), for example, used factors affecting an individual's decision to apply for a loan and factors affecting a lender's decision to grant a loan as controls in their first stage profit procedure. Schuler and Hashemi (1994) used and discarded three control variables that were not significant (religion, number of surviving children, and presence of health care facility in the village).

No study has successfully controlled for the fungibility of resources between the household and the assisted enterprise. The interaction between household and business interests may, in fact, be seen as an attractive feature of microenterprise programs, especially when the programs are considered as part of an anti-poverty strategy. Thus, rather than simply controlling for fungibility, it would be useful to study it in depth as part of the research design.

The choice of control variables for the AIMS assessments will take careful consideration. Annex 1 provides some basis for the choices other authors have made. A suite of variables that include demographic characteristics (gender, age), socio-economic characteristics (assets, education) and behavioral characteristics (evidence of entrepreneurial behavior, risk taking or aversion) should be generated from the AIMS impact studies, and should be useful for multiple controls that will be employed during the analysis.

It is relevant to note that the control issue largely arises when analyzing "successful" programs that are lending to "successful" clients. If random programs were selected and random eligible people were analyzed (borrowers or not), the control issue would largely disappear.

Careful use of matched locational controls is highly recommended and is discussed below. Use of matched locational controls avoids many of the problems of selectivity bias.

VII. LOCATIONAL CONSIDERATIONS

Locational considerations are given minimal attention in most credit program impact studies, although it is clear that they play a major role. As Schuler and Hashemi note, most credit programs seek favorable locations with high accessibility levels. Buckley notes that "One of the major determinants of performance for ISP is location" (p. 406). Few studies (Mustafa et al.'s is an exception) explicitly analyze the impact of this location bias on credit program success. Further, changes within a location, such as road improvement, may have substantial impacts on program success.

Using paired non-program locales for control is a recommendation of this study. However, unless the non-participating village or urban site is in a similar geographic location its usefulness as a comparison site is limited. Control locations should be carefully chosen to include matched sites (one with the program, one without) that are comparable in terms of population and agro-ecological zoning, access to roads and larger towns, situation with reference to neighboring market centers, and levels of business activity (including presence or absence of competing institutions). Locational data are often readily obtainable and easily observable, although often neglected during data capture exercises. For example, a survey of all towns in Zimbabwe is available which lists all businesses operating in every market center. It is likely that similar information will be available for other field sites.

The relative location of the client is often unstudied. An entrepreneur who has direct and ready access to neighboring and/or major markets may benefit considerably more from credit than one who is relatively isolated. Yet, as noted previously, credit access may allow an entrepreneur to become a spatial monopolist or spatial monopsonist. In such a situation where the entrepreneur controls the local market (either selling or buying), the community might actually be worse off since prices charged or paid become exploitive, even though the client prospers.

The above situation also points to a promising area of future study in credit impact studies -- an area that is currently largely neglected. While the current foci are on the business, the household or the loan program itself, an important neglected focus is the impact on the community and its hinterland. Three of the surveyed studies (Schuler and Hashemi, Buvini~~f~~ et al., and Mustafa et al.) note definite community impacts that emerged during their field work.

VIII. ALTERNATIVE METHODS, QUESTIONNAIRE CONCERNS AND FIELD ISSUES

Overall consensus on most methodological issues has emerged in the microenterprise credit impact arena. Yet, other possibly relevant issues have largely eluded researchers. For example, to what degree are favoritism, patronage or political issues factors that affect the granting of credit? What are the levels of accountability for both clients and lenders in a program, and how does this affect credit impact? Are there problems of corruption in a program, (does kleptocracy exist?) such that borrowers must pay bribes or kickbacks to receive credit? Are defaults handled in an unbiased manner, or are they relatively painless positive credit impacts for some clients? To what degree is there leakage from the location (e.g., via remittances) such that there may be positive impacts, but they are being felt elsewhere? It is suggested that studies broaden their focus to attempt to address some of these potential complications of credit impact assessment.

Further, in the methodological literature, attention to alternative methods, questionnaire concerns, and field issues is underrepresented, as is briefly addressed below.

A. Alternative Methods

Few researchers have strayed from the methodological norms in microenterprise credit impact studies. Nelson and Bolnick (1986) do try "supplementary methods" (anthropological and counterfactual), which were not effective and which were applied after the methodological norms. It is intriguing to note that their attempt to use a "but for" approach could easily be addressed through counterfactual methodology.⁹

Much is to be gained by engaging in a few in-depth interviews, using indirect observation and measurement, and widening the longitudinal perspective. Pilot studies incorporating a few nonconventional methods may well uncover important considerations that can be included in more conventional analyses.

B. Questionnaire Concerns

Little attention is paid to questionnaire construction in most of the studies reviewed. One author (Buckley) notes that his enumerators translated the questionnaire into the vernacular *in*

⁹Counterfactual analysis, in layperson's terms, attempts to answer the question "What would have happened if other conditions prevailed, or something else happened?" In credit impact analysis, a good example of this is asking loan recipients to speculate on what their situation would be like if *they had not* received the loan. Counterfactual analysis is particularly useful when there is multiple causality (more than one variable is important). It is also appropriate to analyze the "but for" situation--what would have been the results "but for" this intervention?

situ. This seems to be inviting disaster (or at least variability). No authors mention the simple and inexpensive technique of back translation which would avoid this problem. In back translation, one enumerator translates the questionnaire from English into the vernacular. Another enumerator (kept separate from the first) then translates the questionnaire back from the vernacular into English. By simple comparison of the first and last questionnaires, areas of mistranslation are easily exposed.

Few studies discuss *survey fatigue*, yet it is clear that Mustafa et al.'s marathons and Dearden and Khan's monthly visits can lead to respondent weariness. Few studies mention other factors that bias respondents. It seems clear that using borrowers who are about to receive credit for the first time could effectively give a control population similar to those who have already received credit. Yet, it is also clear that these soon-to-be borrowers have a high self-interest in maintaining a strong image and are unlikely to report any conditions that could, if leaked, threaten their loan status. The motives of the respondents should always be considered before carrying out sampling strategy or constructing a questionnaire.

Nelson and Bolnick (1986) provide a very thorough discussion of questionnaire design issues and give a fairly specific list of questionnaire problems. They strongly suggest that questionnaires have a check-off design (i.e., the enumerator has a list of pre-coded responses which are simply checked off) where possible to eliminate problems encountered with handwritten (and occasionally undecipherable) responses. Nonetheless, it is important to retain some open-ended questions on such issues as fungibility and decision-making to allow for qualitative analysis.

C. Field Issues

Little attention is paid to what actually occurs in the field in most of the studies reviewed (again, Nelson and Bolnick are the exception). The possibilities for reducing error in the process of enumeration and supervision are high. Proper selection, training and management of enumerators can greatly improve data accuracy and reliability. The social context of the interview must be understood as thoroughly as possible and biases due to this context should be minimized. If, for instance, the respondent is attempting to please the enumerator, or is afraid to give information that could be construed as negative, the obtained data will be flawed.

It is often possible in the field to gather non-survey information that will add to the study and to data validation. For example, whether a household has a thatched or metal roof is directly observable. So, usually, are such business assets as a cash register, electrical connection, signage, etc. Community level infrastructure variables (electricity, water supply, storm drainage, etc.) may make significant differences to credit impact and take only minutes to record. For example, the credit impact on a carpenter in a village may vary greatly depending on whether he has access to electricity. If so, he may be able to purchase power tools that can enhance his productivity.

IX. CONCLUSIONS

Studies of microenterprise credit impacts have clearly undergone a "norming" procedure during their evolution so that now there is a consensus on most issues. Substantial progress has been made on many issues, including reducing selection bias, improving sampling techniques, increasing appropriate use of control variables and strengthening longitudinal aspects of the analyses. Nonetheless, other aspects such as controlling for the fungibility of resources between the household and businesses have eluded effective solution, although this concern has led to inclusion of a broader list of variables in recent studies.

In many of the studies surveyed, the methodology seemed to "drive" the study. For evaluation to be effective and pass the integrity test, the objectives of the evaluation must be fully specified and defined and these specifications and definitions should then dictate the types of data collected and the subsequent methodology used. Too often, it seems, this is not done.

A. Methodological Recommendations

The following recommendations are directed at the AIMS core impact assessments. Nevertheless, they are likely to be of significance to similar endeavors.

1. Sampling

- a) Sampling should occur with control groups from within program sites and a control population chosen from matched non-program sites.
- b) Sampling should use a quasi-experimental design, with a matched control group strategy to pick control sites (non-program locations) and statistically equated control groups for individual controls.
- c) In program sites, both new (or soon-to-be) borrowers and eligible non-borrowers should be used as control groups.
- d) Sub-samples of multiple borrowers should be included and the value of their loans (as well as the number) recorded.
- e) Minimum sample size for any sub-sample group should be 30. Higher sample sizes increase the likelihood of attaining statistically significant findings. An overall sample size of 500 should allow for effective use of control variables, drop-outs in the longitudinal study, and invalid data issues.

2. Temporal Issues

- a) The longitudinal study with ex ante and ex post data collection should have

at least an 18 month interval, and probably no more than a 24 month interval, between the two rounds of data collection.

- b) The ex ante group should be surveyed using recall methods targeting a 6 and 18 month period prior to the ex ante study, or, as an alternative, the period directly before receiving their loan.
- c) Seasonality should be a consideration in the actual timing of the field surveys.
- d) In-depth interviews in pilot or exploratory studies should be undertaken to attempt to determine actual financial "time lines" of credit impacts, including those of multiple borrowers, and the temporal aspects of the study should be adjusted accordingly.
- e) There may be neglect of the issue of long term credit impacts.

3. Analytical Techniques

- a) A quasi-experimental design is seen as sufficient, without the need for the addition of econometric modeling.
- b) Multivariate techniques will be necessary to control for endogeneity and selection bias issues.
- c) The choice of the precise techniques to be used should be a function of the type of data collected and their distributional characteristics.
- d) The choice of variables included in the analysis should be expanded. Variables of social class and location have been shown to have significance.
- e) Context has frequently been noted as being important, yet context variables are often excluded from analyses. Qualitative techniques may be of utility.
- f) Data cleaning and data validity checking should be built into the research design and plan.

4. Control Methods

- a) Statistically-equated methods are sufficient for controlling for individual selection bias and issues of endogeneity.
- b) Gender is a critical control variable.
- c) Efforts to attempt to control for the fungibility of resources between

household and enterprise must continue to be made. This may require a broader interpretation of assets.

- d) Precise control methods should be a function of the data available. Statistical power should be the critical determinant of the technique used. Questionnaire construction and data manipulation efforts should consider attaining interval data when possible to increase the power of testing.
- e) The focus on successful microenterprise programs, and the tendency of programs toward choice of likely to be successful clients are the major reasons that controls are needed to study impacts. The alternative of studying programs selected at random and eligible borrowers also selected at random could minimize the need for controls and would give a different, but potentially valuable, assessment of possible credit impacts.

5. Other Issues

- a) Several studies note the diffusion of credit impacts to non-participants in program sites. The impact of programs on the community is a noted, but neglected aspect of study.
- b) Alternative methodologies may complement current mainstream methodologies. Some in-depth pilot interviewing should enhance research and questionnaire design and inform temporal considerations. Counterfactual analysis may be used to provide empirical evidence for the "but for" approach.
- c) Back translation should be used in survey construction.
- d) Survey fatigue should be avoided.
- e) Respondent motives should be considered in questionnaire design.
- f) Greater attention paid to field issues will result in greater data accuracy and validity and minimize overall costs.
- g) Gathering of non-survey complementary information (e.g., infrastructure and location) in the field should be built into the research design.

BIBLIOGRAPHY

- Buckley, Graeme, forthcoming. "Financing the Jua Kali Sector in Kenya: the K-REP Juhudi Scheme and Kenya Industrial Estates Informal Sector Program." In *Finance Against Poverty, Volume II: Country Case Studies*, edited by David Hulme and Paul Mosley, 349-427, London: Routledge.
- Buvinif, Mayra, Marguerite Berger, and Cecelia Jaramillo, 1989. "Impact of a Credit Project for Women and Men Microentrepreneurs in Quito, Ecuador," In *Women's Ventures*, edited by M. Berger and M. Buvinif, 224-246. West Hartford CT: Kumarian Press.
- Dearden, Kirk and Nazmul Khan, 1994. "Assessing the Impact of Women's Savings and Credit Programs on Fertility: A Case from Bangladesh." Save the Children/Child Impact Program Monograph No. 2, Westport, CT.
- Hulme, David and Richard Montgomery with Debapriya Bhattacharya, 1994. "Mutual Finance and the Poor: A Study of the Federation of Thrift and Credit Cooperatives in Sri Lanka (SANASA)." *Working Paper No: 11 1994*, Institute for Development Policy and Management, Manchester; Department of Economics, Reading.
- Hulme, Richard Montgomery, and Debapriya Bhattacharya, 1995. "Credit for the Poor in Bangladesh: The BRAC Rural Development Programme and the Government Thana Resource Development Programme." *Working Paper No: 12 1995*, Institute for Development Policy and Management, Manchester; Department of Economics, Reading.
- Hulme, David and Richard Montgomery, forthcoming. "Mutual Finance and the Poor: A Study of the Federation of Thrift and Credit Cooperatives in Sri Lanka (SANASA)." In *Finance Against Poverty, Volume II: Country Case Studies*, edited by David Hulme and Paul Mosley, 349-427, London: Routledge.
- Lapar, Ma. Lucila A., Douglas H. Graham, Richard L. Meyer, and David S. Kraybill, 1995. "Selectivity Bias in Estimating the Effect of Credit on Output: The Case of Rural Nonfarm Enterprises in the Philippines." Rural Finance Program, Department of Agricultural Economics, The Ohio State University. Columbus, OH.
- Lapar, Ma. Lucila A., Douglas H. Graham, Richard L. Meyer, and David S. Kraybill, 1995. "The Effect of Credit on Output: Are There Sectoral Differences?" Rural Finance Program, Department of Agricultural Economics, The Ohio State University. Columbus, OH.
- Mustafa, Shams, Ishrat Ara, Dilruba Banu, Altaf Hossain, Ajmal Kubir, Mohammad Mohsin, and Abu Yusuf. 1995. *Impact Assessment Study of BRAC's Rural Development Programme. Final Report.*
- Nelson, Eric R. and Bruce R. Bolnick, 1986. "Survey Methods for Assessing Small Credit

Programs Evaluating the Economic Impact of KIK/KMKP Credits." Jakarta. (Included were separate pieces by each of the authors relating to the study).

Pitt, Mark M. and Shahidur R. Khandker, 1994. "Household and Intrahousehold Impacts of the Grameen Bank and Similar Targeted Credit Programs in Bangladesh."

Pitt, Mark M. and Shahidur R. Khandker, 1995. "The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter?"

Pulley, Robert V., 1989. *Making the Poor Creditworthy: A Case Study of the Integrated Rural Development Program in India*. World Bank Discussion Paper 58: Washington, DC.

Schuler, Sidney Ruth and Syed M. Hashemi, 1994. "Credit Programs, Women's Empowerment, and Contraceptive Use in rural Bangladesh." *Studies in Family Planning* 25 (2): 65-76.

Zeller, Manfred, Akhter Ahmed, Suresh Babu, Sumiter Broca, Aliou Diagne, and Manohar Sharma, 1996. "Rural Financial Policies for Food Security of the Poor: Methodologies for a Multicountry Research Project." Food Consumption and Nutrition Division, International Food Policy Research Institute: Washington, DC.

ANNEX 1: VARIABLES USED IN STUDIES REVIEWED

VARIABLE*	Hulme	Mustafa	Pulley	Buvinic	Lapar	Pitt	Zelle
GENERAL							
Age	x	x		x	x	x	x
Age of HH Head						x	
Marital Status	x						
Religion							
Social Class ¹							
Income			x				
Occupation of Beneficiary	x	x	x				
EDUCATION							
Education	x	x	x				x
(in years)		x		x	x	x	
HH Aggregate Education Level		x					x
Highest Grade Completed by HH Head						x	
Highest Grade Completed by an Adult Female in HH						x	
Highest Grade Completed by an Adult Male in HH						x	
Children's Schooling (by gender)						x	
HOUSEHOLD							
Head of Household (client)	x						x
No. HH Participants in Program		x					
Household Size	x	x		x	x		
(before and after)	x						
No. Dependents/HH	x			x			
Total No. of Family Dependents	x						
No. of Children under 5 Years				x			
Demographic Dependency Ratio		x					x
No. of Working Age Population/HH (by income source)		x					

*Please Note: Because of differences among the presentations, information on variables is more complete for some cases than in others. It is possible that several studies used more than the variables listed. However, these were the variables that could be determined based on the information included in each of the studies.

¹ Social Class is determined by five factors: type of housing material, quantity of cultivable land owned, household assets, income, and food self-sufficiency.

No. of HH Members with Individual Cash Incomes	x						
Household Outstanding Debt							x
HH Debt to Income Ratio							x
Relative Wealth of HH (selected indicators ²)							
Place of Residence							
Tenancy Status							
ASSETS/WEALTH							
Fixed Assets	x			x	x		
Fixed Capital							
Total Assets	x				x		
Change in Assets	x						
Financial Assets					x		
Asset Growth							
Change in Capital Assets							
Working Capital					x		
Household Assets	x				x		
(first vs. third time borrowers)	x						
Household Assets "before"	x						x
Household Assets "after"	x						
Household Assets -- percentage increase	x						
Enterprise Assets "before"	x						
Enterprise Assets "after"	x						
Enterprise Assets -- percentage increase	x						
HH Cash Holding Practices -- before and after (by gender)	x						
HH Cash Holding Practices -- before and after (by gender)	x						
Respondents Property/Economic Status (selected indicators ³)							
Bank Account					x		
Aver. Date of First Deposit							
Range by Year of First Deposit							

²There are four indicators for household wealth: have tin or tile roof; own radio or TV; own any agricultural land; own cow or buffalo.

³There are six indicators for property/economic status: own homestead, land, houses; own homestead before joining; have own savings; keep cash for purchases; earned independent income last year; make substantial contribution to family support.

Range of Savings	x						x
Aver. Savings for Borrowers with Deposits							
Aver. Savings for All Borrowers							
Purchase Fixed Assets							
HH per capita Income		x					
Per capita Ownership of Wealth		x					
One Week Cash Earning per capita		x					
LAND							
Land Value					x		
Household Land	x	x				x	x
Land Ownership/Start (per capita)		x					
Parents of HH Head Own Land						x	
Brothers of HH Head Own Land						x	
Sisters of HH Head Own Land						x	
Parents of HH Head's Spouse Own Land						x	
Brothers of HH Head's Spouse Own Land						x	
Sisters of HH head's Spouse Own Land						x	
Experience (Dummy)					x		

PROGRAM							
Length of Program Participation	x	x		x			
How Borrower Heard about Loan	x						
Length of Existence of Village Organization	x	x					
No. of Dropouts (reason)		x					
(by wealth)		x					
(length of participation)		x					
Beneficiary/Member Organization	x		x				
Information about Member Organization (selected indicators)	x						
Occupational Skill Training Received (#)		x					
CREDIT/LOAN INFORMATION							
Access to Credit							x
Constrained Access to Credit							x
Credit Status					x		
Credit Limit (potential borrower thinks he could receive)							x
Loan Approvals							
Loan Refused	x						
Loan Amount Requested							x
Loan Value/Amount Borrowed (by gender)	x	x			x	x	x
(per capita)		x				x	
Passbook Issued			x				
Interest Rate	x						
Date Received	x						
Time between Application and Receipt of Loan	x						
Incidental Acquisition Costs	x		x				
No. Loans Received	x	x					
Outstanding Loans	x						
Amount of Outstanding Loans	x						
Types of Credit Used	x						
No. of Credit Sources	x						
Total Repayment Period	x						
Regularity of Payments	x						
Amount Due Each Payment	x						
Income Source for Repayments	x						
Debt Servicing (by season)		x					

Debt Servicing vs. Total Expenditure (by season)		x					
Savings Tied to Loan	x						
Value of Tied Savings	x						
Participation in Credit Camp by Beneficiary			x				
Type of Investment/Use of Loan (by no. of loans)	x	x	x				
Investment in Line with Traditional Family Occupation			x				
Mode of Disbursement							
Source of Loan Repayment	x						
Value Added by Loan							
% Never Applied to Formal Lender for Loan							
% Who Obtained a Loan from a Formal Lender							
% Obtained Finance through ROSCA Membership							
% Who Have Used a Money Lender							
Average Overdues (by income level)			x				
Overdue Loans	x						
Reasons for Overdues	x		x				
Amount of Overdues (by region/state)			x				
Follow up by Bank and Block Officials			x				
Sustained Access to Credit ⁴ (by income level)			x				
PARTICIPATION							
% of # of Loans Made to Small Farmers by Category (e.g., landowners)							x
% of Small Farm HHs in that Category							x
Value of Loans Made to Small Farmers per HH Category (e.g., owners)							x

⁴This variable looks at beneficiaries that subsequently attempted to borrow from the bank as well as those that were successful in obtaining subsequent bank loans. The information was separated into borrowers with and without overdues.

Value of Loans Made to all Owners								x
-----------------------------------	--	--	--	--	--	--	--	---

VILLAGE ATTRIBUTES							
Local Condition ⁵		x					
Aver. Amount of Credit Obtained in the Village							x
Community Wealth							x
Ratio of HHs to Hand Tubewell		x					
Has any Primary School						x	
District Primary School Coverage			x				
Has Rural Health Center						x	
Has Family Planning Center						x	
Access to Family Planning Services							
Is Midwife Available						x	
Price of Staples						x	
Distance to Bank		x				x	
Distance to Nearest Sub-district Town		x					
Distance from Nearest All-weather Road		x					
District Gross Agricultural Production per capita			x				
87/88 Drought			x				
Village Tendency to Default			x				
Adequate Marketing Facilities			x				
INCOME							
Monthly Sales				x			
Gross Sales					x		
Net Income	x		x		x		
(by season)							x
Pre-IRDP Income			x				
Real Income Growth			x				
Previous Years Income					x		
Monthly Net Income	x			x			
Monthly Income "before"	x						
Monthly Income "after"	x						
Change in Monthly Income	x						
Income Change vs. Price of Rural Commodities	x						
Hourly Net Income				x			

⁵Local condition is a composite variable based on weighted scores for four factors in each village: distance of the village from the nearest sub-district town, distance of the nearest haat/bazaar, the number of permanent shops in the nearest haat or bazaar, and the number of shops inside the village.

Size Distribution of Monthly Sales							
Income Reinvested in Enterprise					x		
Income Reinvested as Working Capital					x		
Income vs. Expenses							
Aver. No. of HH Members with Cash Income							
No. of Income Earners HH	x						
One-week-cash-receipt (HH)		x					
Borrowers with Only One Source of Income							
Borrowers Whose Main Source of Income Is the me	x						
Household Sources of Income (#)	x	x					
Real Return on Investment given Year			x				
Net Income Asset given Year/Investment Cost			x				
GENDER							
Gender (of household head)	x	x	x	x	x	x	x
Male/Female Ratio							
No Adult Male in HH						x	
No Adult Female in HH						x	
No Spouse Present in HH						x	
Empowerment ⁶							
Contraceptive Use							
Female Contact with Credit Programs							
Female Contribution to Family Support							
Diversion of Loan for Consumption Expenditures (by Gender)	x						
Person Deciding about Spending Women's Income		x					
Change in Female Hh Status B/C of Loan		x					

⁶Eight variables are used to determine level of empowerment: mobility, economic security, small purchases, large purchases, major decisions, subjection to domination and violence, political/legal awareness, and protest/campaign.

Female Control over Own Income from Investment		x					
Attitude Towards Women's Mobility		x					
Women's Role in Decision Making about Children's Education		x					
Women's Role in Decision Making about Savings and Their Use		x					
Women's Role in Decision Making about Loans and Their Use	x	x					
Women's Role in Decision Making about Selling and Purchasing		x					
Women's Role in Decision Making about Voting		x					
Group Action Towards Common Goals		x					
% Divorced/Widowed Women	x						
Fertility Rate (by participation)							
Women's Non-Land Assets						x	
LABOR							
Aver. No. Hours Worked (before and after)	x			x			
No. Hours Worked/Day					x		
Total No. of Paid Workers (first time vs. third time borrowers)	x			x			
Change in Employment	x						
Value Labor Services							
Labor Supply (by gender)						x	
No. Workers/Start					x		
No. Workers/Currently	x				x		
Employment Growth Rate							
No. Employees (gender)							
No. of Family Workers	x				x		
No. Hired Workers/Start					x		
No. Hired Workers/Currently					x		
Aver. No. of Paid Workers/Business	x			x			
Total No. of Unpaid Workers/Business				x			
Aver. No. of Unpaid Workers/Business	x			x			
Average No. of FT Employees (paid and unpaid)							
Hourly Wage Rate	x				x		
Wage Rate Change	x						
Wage Rate (by gender)	x						

Casual Laborers			X				
Wage Labor Activity	X						X
No. Days/Year Manual Labor is Sold		X					
ENTERPRISE							
Category of Enterprise	X						
Type of Enterprise	X			X	X		
Age of the Enterprise	X				X		
Source of Start Up Funds	X						
Additional Sources of Funds	X						
Ownership Structure	X						
Business Location	X			X	X		
Nature of Premises	X						
Location of Premises	X						
Record Keeping	X						
No. Months Operated					X		
No. Hours Operated	X				X		
Seasonal Operation	X						
Estimated Profit (Annual Growth)							
Current Expenditures (Annual Growth)	X						
Current Sales	X						
Average Monthly Sales (before and after)	X						
Production/Sales (change)							
Increased Competition (after)	X						
CONSUMPTION							
Per Capita Total HH Expenditure (by category)						X	
HH Expenditure (week)		X					
HH Expenditures (by category)	X						X
Growth Rate in HH Expenditure Per Capita							
Food Stock (by season)		X					
Food Security		X					
Weekly Consumption of Rice (gram per capita)		X					
Household Caloric Intake							X
Intensity of Market Search			X				
Source of Ownership of Cattle		X					
IMPACT							

Sustainability According To Investment Type			X				
Sustainability According to Income Group			X				
Sustainability According to Occupation			X				
Sustainability According to Educational Status			X				
IMPACT--KEY INDICATORS							
Density of Living Space		X					
Livestock Ownership		X					
House Structure Value		X					
Total Asset Value		X					
Asset Per Capita		X					
Share of Revenue Earning Assets (% of Total Assets)		X					
Cash Earning (Week/Capita)		X					
Food Expenditure (Week/Capita)		X					
Deficit Months (No. in Last Year)		X					
Non-Program Saving		X					
All Weather Roofing Material		X					
MISCELLANEOUS							
Tractor Ownership							X
Anthropometric Measures							X
Weight Fluctuations in Young Children							X
Sick Days							X
Social Responsibility in Community							X
External Events ⁷							X

⁷These include migration or death of a family member, a bad harvest, and costly social events such as marriage.

ANNEX 2: SAMPLING FRAMES IN REVIEWED STUDIES

	SAMPLING	SAMPLE SIZE	CONTROL GROUP(S)	DATA SOURCES	REPE
Buckley <i>Juhudi- Kibera</i>	Stratified random sample (by number of loans) <i>Type 1 and Type 2:</i> Borrowers servicing their first and second loans.	Total: 160	Yes <i>Type 3:</i> Members who are due to shortly receive a loan <i>Type 4:</i> Randomly selected non-borrower entrepreneurs	Primary Source: Interviews using a structured questionnaire. Unstructured interviews were also used to obtain additional information.	12 m
<i>ISP</i>	Two-tier stratified sample (by repayment length and number of loans) ⁹	Total: 140	Yes <i>Selection:</i> Randomly selected comparable non-ISP entrepreneurs	See above	See a
Lapar, Graham, and Meyer	Two-stage random sample	400 rural non-farm enterprises	No	Survey	
Pitt and Khandker	Choice-based sample design: (1) Random sample of 29 thanas (subdistricts) drawn from 391 thanas; (2) Random selection of 3 villages from each thana ¹⁰ ; (3) 20 households from each village were selected using stratified random sampling.	1798 households (including 1538 target households, and 905 program participants)	Yes <i>Selection:</i> Villages without credit programs. Participant/Non-Participant. Target/Non-Target.	Household survey ⁸ three times over crop cycle year Nutrition sub-survey ⁸ peak and slack season Village Survey	

⁸ Although the recall period is not clearly stated, some of the data presented compare **A**current@ information with information from the previous year.

⁹ Because of program differences (including different lending procedures and repayment cycles), the ISP sample was stratified differently.

¹⁰ The three villages were selected differently for program and non-program villages. Program villages were randomly selected from a list of villages where the credit program had been in operation for a minimum of 3 years. (Further restrictions were placed on village size and number of households per village.) Non-program villages were randomly drawn using the government village census.

¹¹ Target households are defined as households that own less than 50 decimals of land, i.e., the **A**poor@

Berger and Buvinic	Stratified random sample (by gender and type of activity)	Total: 311 156 participants 155 non-participants	Yes <i>Selection:</i> Non-random, matched characteristics ¹²	Interviews (during a non-holiday season) spaced a year apart.	12 m
Schuler and Hashemi	<i>Participants:</i> Random sample, two-stage cluster design <i>Non-participants:</i> Eligible women living in villages <i>not</i> served by either program	Initial Survey: 1248 Follow-up: 1305 ¹³	Yes <i>Selection:</i> Non-participating villages were matched using selected characteristics ¹⁴	Interviews spaced 18 months apart	18 m
Dearden and Khan	All married women ¹⁶ of reproductive age who were not using contraceptives as of January 1, 1987	54,000 households were served by the paid workers (the actual sample size was not given)	Yes <i>Selection:</i> Non-members in the same village	Monthly household interviews ¹⁷	1 mo

¹² Characteristics such as location of firm, gender and type of entrepreneurial activity were used.

¹³ It is unclear from their paper whether or not these numbers include non-participants.

¹⁴ Non-participating villages were matched to participating villages using region, village size, and population density.

¹⁵ Because of program differences (BRACs strategy is to include *all* eligible women in its program villages), non-members from BRAC villages were not included.

¹⁶ These women were from four rural areas in Bangladesh where Save the Children works.

¹⁷ These interviews were done by paid volunteers (i.e., Save the Children employees).

<p>Nelson and Bolnick</p>	<p>Purposive sample frame: Three regions from the three most important islands were selected.</p> <p>29 economic subsectors representing KIK/KMKP lending were selected.</p> <p>Two subpopulations were identified for each of these subsectors:</p> <p>(1) All recipients of KIK/KMKP credit in the first half of 1980</p> <p>(2) All recipients of KIK/KMKP credit in the first half of 1982</p> <p>Circular systematic sampling was used to select 30 credit recipients from each of the subpopulations (which were first grouped by geographical unit).¹⁸</p>	<p>Total: 1740</p> <p>Original sample included 30 small enterprise owners in each of 29 subsectors, and a matching control group.¹⁹</p>	<p>Yes</p> <p><i>Selection:</i> The control group (recipients of credit in the first half of 1982) represent[ed] those who (1) are not denied access to credit, (2) choose to accept credit when available, but (3) did not have credit during the study period.@</p>	<p>Interviews²⁰</p>	<p>4 years (1971-1975)</p>
----------------------------------	--	--	---	--------------------------------	----------------------------

¹⁸This method was also used to develop a list of substitutes.

¹⁹For almost all of the subsectors, the final sample size was less than 30. The various reasons are detailed in the text.

²⁰ Although the questions were not analyzed in this particular study, the authors included *counterfactual* questions to assess perceived performance differences *without* credit.

²¹ Because of the long recall period, the authors note that the questionnaire and interview methods were designed in order to facilitate more accurate recollection.

²² Non-participation was determined by access, i.e., some households would like to be participants, however, their participation was *constrained* for different reasons.

²³ Because the study covered a *sensitive topic*, required *multiple rounds*, and a *depth of data*, the authors note that a small sample size was required. However, except for two cases (one of which seems to contradict the above statement), sample size is not given. According to the authors, the Madagascar study included only 180 households, and the Pakistan study included more than 2000 households.

Zeller et al.	Random selection of households (participants and non-participants) in communities where a credit program is located. ²² Some of the studies included endogenous (or choice-based) stratified sampling (by participation)	Small sample size ²³	No	Extensive household surveys with multiple rounds to capture seasonal (hungry, harvest, post-harvest) variations. Institutional and community level surveys were also conducted.	Yes (not s
Pulley	Two surveys covering over 25,000 households According to Pulley, only the second survey (i.e., round) used a random sample.	<i>Round One</i> (1985-86): Total: 479 240 old beneficiaries 239 new beneficiaries ²⁴ Old beneficiaries with assets still intact (n=197) were resurveyed in 1988 (n=185) ²⁵ <i>Round Two</i> (1987): Total: 467 239 old beneficiaries 193 new beneficiaries (first dose) 35 new beneficiaries (second dose) ²⁶ Old beneficiaries with assets still intact (n=198) were resurveyed in 1988 (n=179)	No	Household surveys	2 year

²² Non-participation was determined by access, i.e., some households would like to be participants, however, their participation was constrained for different reasons.

²³ Because the study covered a sensitive topic, required multiple rounds, and a depth of data, the authors note that a small sample size was required. However, except for two cases (one of which seems to contradict the above statement), sample size is not given. According to the authors, the Madagascar study included only 180 households, and the Pakistan study included more than 2000 households.

²⁴ Old beneficiaries are defined as those that received assets two years prior to the survey date; new beneficiaries are defined as those receiving assets three months prior to the survey date.

²⁵ Beneficiaries that had died, moved, or dropped out of the program were not included in the second sample.

²⁶ This group of households consists of beneficiaries that did not cross the poverty line during the Sixth Plan.

Hulme et al.					
<i>Sri Lanka</i>	Two District Unions ²⁷ : Kurunegala and Moneragala 15 Member Societies: Kurunegala 6 Moneragala 9 151 Individual Members: The membership of the 15 case study societies was used as the sampling frame	Total: 151	No ²⁸	Survey Informal methods such as discussions with members and non-members were also used to obtain qualitative information	1-2 y

²⁷ The two district unions were selected because they provided a contrast in terms of the physical and economic environment as well as the length of participation in the thrift and credit movement.

²⁸ This decision to not include a control group of borrowers about to receive their first loan was made during the early stages of fieldwork after it was noted that the initial members/borrowers were generally of higher income and status than later borrowers.

Hulme et al.					
<i>Bangladesh</i>	<p>Purposive selection of four research areas: two for each program.²⁹</p> <p>Voluntary organizations and kendras within these research areas were classified according to their age (1B years, 3B years, and recent)</p> <p>Members were then stratified according to borrower types (1 loan, 2 loans, and no loans)</p> <p>Final sample was randomly selected from these lists.</p>	<p><i>BRAC</i> Total: 156 63 one loan 33 three loans 60 new members</p> <p><i>TRDEP</i> Total: 160 64 one loan 32 three loans 64 new members</p>	<p>Yes³⁰</p> <p><i>Selection:</i> New members who had not yet received a loan</p>	<p>Survey</p> <p>Informal fieldwork was also done to gather qualitative information in order to establish context.</p>	<p>1B4 1 (the 1 befor respc took loans</p>

²⁹ This was done in consultation with program head office staff in order to select one vibrant and one depressed local economy.

³⁰ The authors noted that the assumption that the borrowers in the control group are in a similar economic situation to those in the other two groups is debatable based on observations that indicated that the economic status of more recent members is better than that of the original members.

³¹ The three male VOs were selected first from different villages; the remaining villages were used as the sampling frame for female VOs. The authors note that this was a purposive attempt to avoid multiple sampling from any village.

³² These factors included distance from the AO, presence of RDP, and the likelihood that RDP would expand to the village in the next six months. The authors note that these villages are usually located on the outskirts of the command area of respective AOs.

<p>Mustafa et al.</p>	<p>Stratified random sampling (by program maturity) was used to select 15 Branch and Area Offices (AOs)</p> <p><i>Village Organizations:</i> The list of VO names in each RDP village was used as the sample frame. Ten VOs from each of the AOs were selected: 3 for men and 7 for women.³¹</p> <p><i>Comparison Villages:</i> Sample frame was developed at each location using local information.³²</p>	<p><i>Households:</i> Total: 2250 1500 RDP³³ 750 non-RDP</p> <p><i>Village Profiles:</i> Total: 225 150 RDP 75 non-RDP</p> <p><i>Case Studies:</i> 15 village organizations</p>	<p>Yes</p> <p><i>Selection:</i> Households that met the criteria used to define RDP target groups: land holding of less than 0.5 acres and the sale of at least 100 days of manual labor in a year. In addition, only households that did <i>not</i> own any purchased cattle were included in the comparison group.</p>	<p>Household Survey[†] two rounds (peak and slack season)</p> <p>Village Profiles [†] constructed using small groups of A key informants[@]</p> <p>Case Studies[†] formal and informal techniques were used to gather quantitative and qualitative information</p>	<p>Yes</p>
<p>Mustafa et al. (continued)</p>	<p><i>Households:</i> Random selection of 10 households from each village using the VO membership list.³⁴</p> <p><i>Households (comparison):</i> 10 households meeting the target group criteria in each comparison village were selected using the NW corner and anti-clockwise traveling method.</p>				

³¹ The three male VOs were selected first from different villages; the remaining villages were used as the sampling frame for female VOs. The authors note that this was a purposive attempt to avoid multiple sampling from any village.[@]

³² These factors included distance from the AO, presence of RDP, and the likelihood that RDP would expand to the village in the next six months. The authors note that [t]hese villages are usually located on the outskirts of the command area of respective AOs.[@]

³³ The authors note that due to practical problems[@] the final sample size was reduced.

³⁴ Substitutes were also selected using this method in order to avoid multiple sampling from a single household.[@]