Notes on Methods Used in a Survey of Rural Clients of Financiera Calpiá in El Salvador

Mark Schreiner,
Claudio Gonzalez-Vega,
Margarita Beneke de Sanfeliu, and
Mauricio A. Shi

August 1998

Microfinance Risk Management
6070 Chippewa St. #1W, St. Louis, MO 63109-3060, U.S.A.
Telephone: (314) 481-9788, http://www.microfinance.com
and
Center for Social Development
Washington University in St. Louis
Campus Box 1196, One Brookings Drive, St. Louis, MO 63130-4899, U.S.A.

Mark Schreiner is a Consultant with Microfinance Risk Managements and a Senior Scholar in the Center for Social Development, George Warren Brown School of Social Work, Washington University in St. Louis. Claudio Gonzalez-Vega is the Director of the Rural Finance Program, Department of Agricultural, Environmental, and Development Economics at The Ohio State University. Margarita Beneke de Sanfeliu and Mauricio A. Shi are in the Department of Economic and Social Studies at FUSADES. This paper was written while Schreiner was at The Ohio State University.
The survey described here would not have been possible without the collaboration of Financiera Calpiá, especially of Aristóteles Esperanza, Silke Müffelmann, Enriqueta de Rodríguez, and Miguel Rodríguez. The authors thank Anabella de Palomo of FUSADES and Rafael Pleitez of the Universidad Centroamericana for their participation in this research. The survey is a component of the research agenda of the BASIS CRSP in Central America and it was sponsored by USAID and the Central Bank of El Salvador. The authors thank Mary Ott and Gino Bettaglio for their support. The survey enumerators were David Augusto Cañas Vides, Roberto Balmore Castro Tobar, Sofía Cuéllar Guzmán, Nelson Adalberto Escolán Jovel, Julia del Carmen García Morán, Ricardo Ernesto Guirola, Edgar Alirio Hernández Delgado, Julio Ernesto López, Moisés de Jesús López Gómez, Silvia Carolina Oliva Campos, Yesenia Marisol Pérez, Douglas Napoleón Romero Melgar, Carlos Enrique Sierra Esquivel, Ernesto Antonio Valle Campos, and Erick Moisés Zaldaña. The enumerators were supervised by Rudy Paniagua, Mauricio Pinto, Mauricio Salazar, and Mauricio Shi. The deepest thanks go to the unnamed rural borrowers of Calpiá who gave their time to answer the questions in the survey.

Table of contents

List of Figures

List of Tables

List of Acronyms

BancoSol	Banco Solidario, S.A. (Bolivia)
BASIS	Broadening Access and Strengthening Input Market Systems
BFA	Banco de Fomento Agropecuario
CAM	Centro de Apoyo a la Microempresa
CARMA	Consortium for Applied Research on Market Access
c.d.f.	Cumulative distribution function
CRSP	Collaborative Research Support Program
FIE	Fomento a Iniciativas Económicas (Bolivia)
FUSADES	Fundación Savadoreña para el Desarrollo Económico y Social
MFO	Microfinance organization
NGO	Non-government organization
OLS	Ordinary Least Squares
OSU	The Ohio State University (Rural Finance Program)
p.d.f.	Probability distribution function
PRODEM	Fundación para la Promoción y del Desarrollo de la Microempresa (Bolivia)
USAID	United States Agency for International Development

Notes on Methods Used in a Survey of Rural Clients of Financiera Calpiá in El Salvador

1. Introduction

This note describes the methods used in a survey of rural clients of Financiera Calpiá in El Salvador.[1] The project is being implemented by members of the Rural Finance Program in the Department of Agricultural, Environmental, and Development Economics at The Ohio State University (OSU) and by members of the Departamento de Estudios Económicos y Sociales at the Fundación Savadoreña para el Desarrollo Económico y Social (FUSADES). The survey is part of the research agenda of the BASIS CRSP in El Salvador (Gonzalez-Vega, 1998).

This is one of two companion notes. This first note has three parts. Part 2 lists some reasons to do research on the rural borrowers of Calpiá. It tells why the survey attempts to answer the questions it poses, and it tells what sort of data are collected to address these questions. Part 3 lists the strong points of the methods used in this survey. Part 4 describes how the sample of rural borrowers of Calpiá was drawn.

The second note looks at some of the traits of the population of rural borrowers of Calpiá and identifies some preliminary issues for research. It also checks how well the sample represents the population as a whole (Schreiner, Gonzalez-Vega, *et al.*,

[1] In Nahuatl, *Calpiá* means a small box to store savings, usually in the ground.

1998). This second note illustrates the methodological issues raised in the present note, and it shows how some interesting results can be obtained even from the limited information available to draw the sample.

2. Why study the rural borrowers of Calpiá?

In the past three decades, few lenders worldwide have been able to make small loans to rural borrowers, to recover the loans, and to pay their own costs with the revenue from the loans (Gonzalez-Vega, Prado Guachalla, and Miller Sanabria, 1997; Yaron, 1994; Adams, 1988 and 1971; Adams, Von Pischke, and Graham, 1984). Calpiá has been an exception. At least when compared with other rural lenders, Calpiá has had uncommon success. This success inspired a quest for knowledge of how it happened, and it led to the survey of the rural borrowers of Calpiá described here.

The success of Calpiá could be due to the traits of its borrowers, the traits of its products, or both. An important aspect of the product is the lending technology used. The *technology* is the set of methods used to attract business, to exchange signals with potential customers and to screen them based on creditworthiness, to disburse loans and collect repayments, to monitor debtors, and to enforce loan repayment.

The work aims to answer two main questions. The first main question is: How do the rural borrowers of Calpiá compare with other rural people? How rural and how poor are they?

For El Salvador, the answer matters because in the next few years both USAID and the World Bank plan to unleash massive projects to promote small loans in rural areas. Calpiá, along with CAM, are the two microfinance organizations (MFOs) in El Salvador with the most success in rural outreach (Sanchez, Cuevas, and Chaves, 1998;

Gonzalez-Vega, 1996; Interdisziplinare Projekt Consult, 1993). The BFA and some financial cooperatives are also important actors in the market for small rural loans (Quintanilla and Rioseco, 1995).

Knowledge of the rural borrowers of Calpiá is needed to predict how much room the market has for more rural lenders. This knowledge could inform choices as to how best to use funds earmarked for rural microfinance in El Salvador. It will map the market niche served by Calpiá and suggest how many other households might be reached with the same lending technology and degree of institutional sustainability.

For the world, the answer matters because Calpiá has already become an example for other lenders who want to reach rural households. If the borrowers of Calpiá live in dense bunches near roads and have access to off-farm jobs, then other lenders worldwide should know this when they attempt to clone the work of Calpiá in a place where people are dispersed, cannot get work off the farm, and do not have quick, cheap access to other markets.

The second main question is: How do loans from Calpiá compare with loans from other sources? The answer matters since Calpiá might do well with small, rural loans because it beats other sources or because there are no other sources. To this end, the survey tries to measure, both for Calpiá and for all other sources, access to loans and the costs borne by a borrower.

The survey helps to answer these questions since it collects some of the data needed to compare the borrowers of Calpiá with other rural people. The survey questionnaire is based on an instrument used in two nationwide surveys of rural households in El Salvador. The first nationwide rural survey covered activities in 1995. It was done in 1996 by FUSADES and the World Bank, and it was funded by USAID. FUSADES (1996) documented this survey, and the data have been analyzed in Sanchez, Cuevas, and Chaves (1998) and in World Bank (1997). The second nationwide rural survey covered 1997. It was done in 1998 by FUSADES and OSU, and it was funded by USAID through the BASIS project. A third wave is planned for the year 2000. Many questionnaire items on the two nationwide surveys were kept for the instrument used with the rural borrowers of Calpiá. The distribution of answers to these matched questions for the sample of rural people as a whole can be compared with the distribution of the answers for the sample of rural borrowers of Calpiá.

2.1 Income

The most important element matched among these instruments is household income in a year. The measurement of income is the bulk of the survey. It includes the following sources:

- Agriculture, whether plants or animals and whether consumed at home or sold;
- Wages for farm or for off-farm work paid in kind or in cash;
- Aid from government, churches, or non-government organizations (NGOs) in kind or in cash;
- Remittances from abroad in cash or in kind;
- Businesses run by the household.

The survey also collects data on some expenses incurred. Among others, the expenses include the following:

- Wages paid for hired hands;
- Expenses incurred by businesses run by the household;
- Purchases of agricultural inputs.

It costs a lot to try to measure income (Deaton, 1997). The survey probably does not capture all sources of income and expenses. This is not a weakness of this survey but rather a weakness of the survey method. The respondents cannot recall all that they earned and spent in the past year without mistakes. Coupled with the fact that some of the questions asked about income were not exactly the same in each of the three surveys, this means that no single measure of net income can be compared across the surveys. Still, it is valid and useful to compare elements of net income, elements of gross income, and partial aggregations of these elements.

2.2 Wealth, health, and human capital

Income is a flow. In a year, the flow of income depends in part on the stocks of wealth, health, and human capital. In turn, these stocks depend on past flows and on past stocks. The flow of income in a year also depends in part on factors external to the household. Some of these change from year to year. This means that a household can be poor in one year and not poor in the next even though in most years in the long term it would be either poor or not poor (Morduch, 1994). For the example of El Salvador, the extreme weather caused by El Niño in 1997 meant that many households,

especially those involved in agriculture, had uncommonly low incomes, regardless of their typical incomes in the long term. Low income in 1997 does not always mean that the household had low income in past years nor that it will have low income in future years. This question is being addressed by another component of the BASIS research agenda in El Salvador.

Wealth, health, and human capital vary less from year to year than does income. Furthermore, these items predict future income well. For example, a household with low levels of education is unlikely to get the good non-farm jobs that mean an escape from poverty in subsistence agriculture (Reardon *et al.*, 1998).

To measure wealth, the survey asks the household about a long list of assets that it might own. The assets range from radios to tractors to financial deposits. The two most important assets are houses and land. The survey measures land owned, land rented, and land sharecropped. The survey records not just the surface area of the land but also its quality and slope. Quality of land matters and may explain much of the oft-found inverse link between farm size and productivity. The quality of the house is a good signal of future income. For example, a household enjoys the use of a concrete floor or a tin roof through time once they are installed. This use is just like income. A safe, clean house also feedbacks to better health and thus more wealth and human capital.

To measure health, the survey asks two groups of questions. The first group records, for all births in the lifetime of fertile women, the death of children before their first birthday. The rate of infant mortality tracks past household health like lines of crust in rock track past earthquakes and volcanoes. The second group of questions asks about spells of illnesses in the past year that kept members of the household in bed, out of work, or home from school. Recent illnesses track current health like a seismograph tracks current earthquakes or volcanoes. Health affects how well a household can earn income, and poor health can lead to low incomes and thus more poor health through time (Glomm and Palumbo, 1993; Gersovitz, 1983).

To measure human capital, the survey collects data on the last grade completed by each member of the household. It also records whether each person knows how to read and write. Children who tend to flunk classes and who drop out of school have been linked to credit constraints faced by the household (Jacoby, 1994).

The survey of the rural borrowers of Calpiá and the two nationwide rural surveys all contain a complete set of questions used to make an index of satisfaction of basic needs from census data from the Ministry of Planning (Ministerio de Coordinación del Desarrollo Económico y Social, *circa* 1994). Most parts of this index were proxies for wealth, health, and human capital. The index can be computed both for the rural borrowers of Calpiá and for the rural households as a whole. Much like income, the index can be compared between the two groups.

2.3 Shocks

Access to deposits and loans can help to smooth the consumption of low-income households that are subject to income and expense shocks (Deaton, 1992a and 1991). Common sources of income shocks are the wide swings in agricultural output from year to year caused by changes in rain and pests. Agricultural prices also change in ways that are both unpredictable and beyond the control of farmers. The rural poor sometimes do start small businesses, but the income from these businesses often comes in spurts and stops, and they often die before their first birthday (Mead and Leidholm, 1998). Common sources of expense shocks are life cycle events such as births, deaths, marriages, and emigrations (Rutherford, 1998).

Donors tout financial services as a tool to promote food security for the rural poor (Zeller *et al.*, 1997). This survey is one in a long line that tries to measure shocks and to link them to the use or non-use of financial services (*e.g.*, Fafchamps, Udry, and Czukas, 1998; Carter, 1997; Udry, 1995; Townsend, 1994; Fafchamps, 1993; Rosenzweig and Wolpin, 1993; Morduch, 1993; Deaton, 1992b; Paxson, 1992). A shock could increase or decrease the demand for loans of a household, and the same shock could also increase or decrease the creditworthiness of the household. The robustness of loan demand and of loan repayment to shocks is of keen interest since the rural portfolio of Calpiá grew through years of good weather. The survey will help to show the effects of El Niño and the bad weather it caused on the rural borrowers of Calpiá.

2.4 Access to markets

The borrowers of Calpiá might be the least rural of all rural households. That is, they may be close to markets, where closeness is measured in distance, time, or cash cost. For example, suppose that the survey shows that at least one member of the households of rural borrowers of Calpiá tends to work for wages in non-farm jobs. On the one hand, that result would strike a blow against the hope of some donors that Calpiá has developed a technology that can recoup the costs of loans to farmers that are repaid just with proceeds from the farm project. On the other hand, it would signal the strength of the technology that Calpiá has purposely designed to draw repayment not just from farm projects but rather from all the cash flows in the household portfolio (Dunn, 1997; Chen and Dunn, 1996). The same results would follow if it were found that rural borrowers of Calpiá are more likely to get remittances each month.

The survey attempts to measure access to markets in a number of ways. For example, it asks whether a member of the household works for wages in a non-farm job. It also asks questions about the distance from the homestead to factories, schools, banks, stores, bus stops, clinics, phone booths, and marketplaces. It measures distance not just in miles but also in the time and the cash needed for transport.

For farms, the survey asks how much produce is sold and how much is consumed at home. It collects the prices of sales and asks for the buyer. It collects expenses on inputs and asks for the seller. The better the link with markets, the higher the

proportion of produce that is sold, the higher the proportion of inputs that is purchased, and the more distant will be the buyer and seller. This allows the producer to operate in larger markets.

2.5 Access to loans

Access to loans depends on at least four things: distance between borrowers and lenders, the total cost of loans to borrowers, the ability of the lender and the borrower to send and to receive signals about creditworthiness, and the credit limit faced by the borrower. Distance proxies information costs and transaction costs.

2.5.1 Distance between borrowers and lenders

Borrowers and lenders can be distant in two ways. First, there may not be lenders at all, at least lenders who make loans with terms that could attract the demand of rural borrowers. Even if there are lenders, borrowers may not know about them or know about the terms of their loan contracts. Big commercial banks in El Salvador have lent mostly just to big rural borrowers. Informal finance has been uncommonly scarce, in part due to the conflict (World Bank, 1997). The survey intends to find out whether Calpiá faces effective competition for rural borrowers. In the case that Calpiá does not face effective competition, then the survey will shed light on whether Calpiá could likely compete well if it had to do so. The survey records knowledge and contact with all sources of formal and informal loans, both in cash and in kind.

Second, borrowers and lenders can be distant in terms of travel length, time, or cost. A battery of questions measures this distance in miles, minutes, and money.

2.5.2 Total cost of loans to borrowers

The total cost of a loan from the point of view of the borrower includes both price and transaction costs (Navajas *et al.*, 1998; Adams, 1993; Meyer and Cuevas, 1990; Gonzalez-Vega and Gonzalez-Garita, 1987). Price includes interest expenses and fees. Prices paid by the borrower are revenues received by the lender. Transaction costs are non-price costs borne by the borrower. Two examples are the opportunity cost of time spent to get a loan and the out-of-pocket expenses for trips linked with a loan. Transaction costs borne by the client are not revenues received by the lender.

Like most surveys of rural finance, this survey measures price as the interest rate and fees of a loan. Unlike most surveys, however, this one also collects all the other times and amounts of cash flows. This includes the disbursement, any grace period with or without interest, and the number and amount of installments. The final price is the internal rate of return, the discount rate at which the cash flows have a present value of zero (Brigham and Gapenski, 1993). The internal rate of return will not be the same as the interest rate. In most cases, it will be much higher.

This survey is also unique in that it gathers the data needed to compute the transaction costs borne by the borrower in each step of the loan process. For example, it asks the borrower for the number of trips needed to apply for a loan. It also asks for

the time and money spent per trip and the amount the applicant would charge if he or she were to do the same work he or she does now for wages. The transaction cost of the loan application is then the out-of-pocket expenses for transport added to the product of the number of trips, the time spent per trip, and the opportunity cost of time lost.

The survey attempts to measure all transaction costs. This includes the cost to get the needed documents, to prepare for visits by a loan officer, to make the application itself, to make trips to check if the loan is ready, to receive the loan disbursement, to receive visits from a loan officer, and to make trips to make repayments.

In the end, the instrument is designed to collect the details needed to plot the amount and time of all the cash flows and cost flows linked to the loan from both price costs and transaction costs. In some ways, the structure of these costs dictates what sort of household will likely demand loans of the type made by Calpiá. For example, the household must have a project or a group of projects whose internal rate of return exceeds that of the cost-and-cash flows. Or, if the cost structure leans toward price costs, then the loans will attract just borrowers with more cash than time.

2.5.3 Signals of creditworthiness

Lenders want to know not just whether a borrower will be able to repay but also whether the borrower will be willing to repay. Potential borrowers need to signal their creditworthiness, and lenders need to use systems to accept and to process some signals

and to ignore other signals. Some signals are observed by the lender and formalized in the process by which the lender as an organization makes choices. Other signals may be observed but are not likely to be formalized. As a complement to the present analysis, the BASIS team is investigating the details and evolution of the lending technology of Calpiá.

2.5.3.1 Formalized signals

Calpiá asks potential borrowers for references. Good references signal a good reputation in the neighborhood. The lender uses the reference to access the local information that has been built up through time about the honesty of the borrower in his or her work-a-day dealings in the neighborhood (Chaves and Gonzalez-Vega, 1996; Gonzalez-Vega, Jiménez, and Quirós, 1996).

The survey uses three ways to examine the likelihood that a borrower can signal a good reputation through good references. First, the survey asks how long a household has lived in one place. More time without a move means more chances to build a good reputation. Second, the survey looks at membership in local groups. Membership leads to frequent contact with people in the neighborhood. These people could become references. Third, the survey checks whether the household keeps written records of its income and expenses. A loan officer can convince a skeptical credit committee of the truth of reported cash flows better with written records than with numbers recalled from memory (Sanchez-Schwartz, 1996).

2.5.3.2 Unformalized signals

People who look the same in most ways on the outside are not always the same in all ways on the inside. All else constant, some people work more, tell the truth more, and do better than others. Perhaps the loan officers of Calpiá have an excellent sixth sense. They may be able to snoop for a few minutes near a homestead and just know whether a person is likely to repay or not. For example, the skill and the drive of the farmer are almost always the most important factors in farm productivity. But smarts and gumption are not easy to measure. They can be observed, but only through time and through a host of details that are not easy to convey to others.

The survey gets at these factors obliquely. It contains questions whose answers may signal that the household is extra smart, extra honest, or extra diligent. For example, leaders in local groups probably resemble the other members of the group in most ways, but their leadership is a clue that they are likely to be more dynamic. Group leaders may also be more prone to rent- seeking and the lender must take this into account. Among farmers with the same size and quality of land, some farmers are more likely to seek out and to get technical assistance and to experiment with new inputs, techniques, or crops. Likewise, some farmers use methods meant to help conserve soil more than other farmers. These clues may hint that the farmers may be smarter or have longer time horizons. A longer time horizon means that good standing with a lender is worth more (Gonzalez-Vega, 1993), and more smarts means a lower

likelihood of project failure. Both increase the likelihood that the farmer will repay a loan.

2.5.4 Amount of access to loans

Access to loans is not all-or-none or take-it-or-leave-it. For example, a borrower can have access to Calpiá and thus be able to get a loan, but this does not mean that the borrower can get a loan for $1 million. The survey attempts to measure degrees of access to loans through credit limits. It asks for the highest amount the household could get from a lender, regardless of the size of the loan the household has now or that the household might want in the future. This is a hypothetical question not based on observed acts. The answers may not reflect exactly the amount that the lender would in fact loan, while some borrowers may overestimate their repayment capacity. To verify the validity of these claims is precisely the challenge for the lender. Still, the concept is central to the measurement of access, and it cannot be ignored in a good survey of rural finance (Zeller, 1994).

3. What are the strong points of this survey?

This survey has two strengths that are more important than all others. First, it was conceived to answer a few pointed questions about rural finance. Second, its design sprung from a research framework based on economic theory. Both strengths mean that the work was planned before it was done and that it was aimed at a few sharp targets.

The framework matters for two reasons. First, it guides the research questions asked. The process that leads to the choice of what questions to ask matters as much as the process that leads to answers to those questions (Smith, 1996). Too often the process that leads to the questions is a black box. Whether reported or not, there is a process. Research questions do not just appear out of thin air. Not to report the process that produced the questions can only obscure the analysis and decrease its chances to be useful outside the small world of the researchers themselves. Second, the framework helps to make sure that the instrument does not omit the most important questions for rural finance. For example, the survey asks about shocks and savings, two topics central to rural finance but forgotten in most instruments.

The survey was designed for just a few, well-defined uses. All the items in the instrument serve a purpose for at least one of the research questions. Streamlined surveys are more likely to be analyzed and to make sense to respondents (McPherson and Parker, 1993). This is not to say that the instrument is short. In fact, it is long—most interviews lasted about 90 minutes. This is also not to say that the

instrument serves just one purpose. In fact, the instrument serves a plethora of purposes. It is funded by more than one source, and each source had its own purposes. Still, the survey does not omit important questions needed to reach these varied purposes, nor does it include items that do not pertain to them.

The methods used for the survey have at least four other strengths.

3.1. Public good

The survey is designed for use by researchers outside of OSU and FUSADES. In principle, most surveys of rural finance are sold to donors as public goods that will lead to more research, better knowledge of the rural poor, and thus better policy to help the rural poor (*e.g.*, Gonzalez-Vega, 1998). In practice, most researchers eschew the use of surveys done by someone else since the records of the methods used and the documents that record the contents of the data files are weak or absent even if the people who collected the data make good on their promise to provide free copies of the data. This survey aims higher than just lip service to the goal to produce a public good that is useful in truth. This goal has at least two implications.

First, it requires careful written records of all steps. This means documents that record the contents of all computer files. In fact, this has meant this document itself. In practice, these documents are not an extra hardship—a good survey project would do them for its own good regardless of whether others might want to use the documents

down the road. Few researchers can remember the details of what they themselves did a few months ago, and most survey projects last for years.

Second, the goal to be useful to third parties requires that all the items in the instrument be scrutinized with the end analysis in mind. Items unlikely to produce data that could be analyzed by researchers not connected with the survey in the field should be cut. To this end, the survey aims to produce formal knowledge that can be stored outside the memory of any one person or organization. Of course, the survey will also produce informal knowledge, especially among the supervisors and the enumerators. Still, questions that could not lead to formal knowledge were omitted.

3.2. Statistical analysis

3.2.1. Parametric econometric regression of conditional means

The survey was designed for two types of statistical analysis. The first type, econometric regression, leads to academic publications. In practice, design for regression has meant extra care to get robust data on those few items that tend to make up the bulk of the independent variables in regressions. These independent variables include:

- Income;
- Structure of income on- and off-farm;
- Amount of remittances;
- Amount and quality of land used;
- Type of roof and floor of house;
- Rooms in the house;
- Number of members in the household;
- Ages and sex of the head and of the other members of the household;
- Years of school for the members of the household;
- Literacy of the members of the household;

- Spells of illnesses;
- Infant death rate;
- Presence of shocks;
- Employment in off-farm jobs;
- Years of residence;
- Years of employment;
- Distance from markets;
- Membership in grassroots groups;
- Use of agricultural extension services;
- Use of written records of revenues and expenses;
- Prices of loans;
- Transaction cost of loans;
- Loans and deposits from formal sources.

The survey also collects data useful as dependent variables. These include:

- Access to formal and informal loans;
- Use of formal and informal loans;
- Amount borrowed in a year;
- Amount of debt held at the end of a year;
- Monthly debt service;
- Level of access as seen through the credit limit;
- Use of deposits.

3.2.2. Non-parametric descriptions of distributions

The second type of statistical analysis, the non-parametric analysis of distributions, leads to work meant to influence policy. *Non-parametric* is a fancy term that just means that the analysis does not assume the data conform to a known probability distribution. It is good not to assume a distribution for at least two reasons. First, in most cases no one knows the true distribution. Second, it would be unwise to assume a probability distribution when the analysis focuses on the shape and location of the distribution itself.

Regression analysis is parametric since it assumes a known distribution for the error term. Non-parametric analysis is descriptive analysis. That an analysis is descriptive does not mean that it is not rigorous. Non-parametric analysis is descriptive just because, unlike econometric regression, it does not assume a cause-and-effect link between independent and dependent variables (Deaton, 1997).

In the past, descriptive analysis fell from grace for two reasons. First, most descriptive analyses were not rigorous. They drew conclusions that were not warranted by the data. Second, most descriptive analyses focused on a single aspect of a distribution, the mean. Both of these weaknesses are problems not with descriptive analysis in principle but rather with its use in practice.

A weakness of non-parametric analysis is that it needs a lot of data to compensate for its lack of identifying assumptions. Thus, without a lot of data, non-parametric description cannot control for as many variables as econometric regression can. Regression can get by with less data since it makes extra assumptions about the functional link between variables and about the distribution of the error term. Regression just estimates the mean of the dependent variable, conditional on the independent variables, but descriptive analysis examines the whole distribution.

Responses to items in the instrument are distributed either as groups or as numbers. Examples of group responses are “yes versus no” or “which lenders made a loan to the household in the past year”. Group responses are analyzed with histograms

or tables which show the number of responses in each category of the group and/or the proportion of the total responses that fall in each category. Like all statistics, these percentages have standard errors. Unlike almost all survey analyses, the analysis of this survey will report the standard errors of its descriptive statistics.

Examples of number responses are the level of income in a year or the area of land farmed. These are analyzed with density and/or cumulative probability distributions as well as with tables of key percentiles. The graphs of these distributions will often be smoothed so that jagged edges caused by sampling variability do not to distract the eye from the most important features (Silverman, 1986).

The analysis of whole distributions goes far beyond the common analysis of means. It is not as simple to look at whole distributions as to look at means, but the distribution tells a lot more about the data. The analysis of this survey will eschew means and focus on the median, a more robust measure of central tendency.

Unlike almost all analyses, the non-parametric analysis of this survey will describe not only estimated non-parametric statistics such as the median or the sample distribution itself but also their sampling variability. Few people realize that the median and the sample distribution, like the mean and like all other statistics, varies in repeated samples. Variability will be computed with the bootstrap and/or the jack-knife (Efron and Tibshirani, 1993).

3.3. Experience with the instrument

Most survey instruments have a lot of problems. That is the nature of their Catch-22: the instrument cannot know the best questions to ask until it has already asked the questions. The instrument for this survey, however, has fewer problems than most. This is not because the instrument was born perfect, but rather because it has evolved through more than three generations. Most of its questions have already been tested in other surveys. For example, the non-finance part of the instrument was based on the ones used in the two nationwide rural surveys done by FUSADES with the World Bank in 1996 and with OSU in 1998. Improvements on this base came from the lessons learned from the field and from the analysis of the data. Furthermore, the new instrument was tested in the field and adjusted before the full-scale survey started.

One example of an improvement was to drop items that always got the same response from all interviewees in the past. Another example was to instruct the enumerator whether or not to read the list of responses with items with such an open list. This helps to ensure that each enumerator handles each item in the same way. Furthermore, the new instrument asks for remittances both in cash and in kind. Likewise, items concerned with the stocks and flows of farm products now use a format of stock-at-start + all inflows – all outflows = stock-at-end. With the past instruments, enumerators in the field had no way to make sure that this string of linked responses balanced. Accurate measurement of the sale and use of farm products matters since it

is a big chunk of income for many rural households. Finally, vast blocks of lengthy items on health and time spent at work were dropped to make room for more items concerned with finance. Most of the items dropped had not been analyzed and so had been collected at high cost for no gain.

To some degree, the improvements do decrease comparability among the three surveys. The judgement is that it is better to gather data thought to be better than to gather data known to be worse. In many cases, the old format is a subset of the new format. In other cases, an analyst can still select parts of the new format to compare with parts of the old format.

The finance part of the instrument is a much-improved child of past finance instruments. The members of this family tree were honed in, among other places, Bolivia (Gonzalez-Vega *et al.*, 1996), Mexico (Sanchez-Schwartz, 1996), South Africa (Ouattara and Graham, 1996), The Gambia (Baydas *et al.*, 1994), the Dominican Republic (Guerrero, 1988; Vasquez, 1986), and Costa Rica (Quiros-Rodriguez, 1991; Gonzalez-Garita, 1986).

The finance part of the instrument has two main strengths. First, it gathers data on the time and amounts of all price costs and all transaction costs from the point of view of the borrower. Second, it runs the whole gamut of finance. *Finance* is an exchange at one time for the promise of a later exchange. Finance is much more than just loans. *Credit* is the option to get a loan if wanted. *Credit* and *access to loans* are

the same. In contrast to the claims of Von Pischke (1991, p. 55), Shipton (1992, p. 27) and Buckley (1997, p. 1092), *credit* and *debt* are not the same. *Debt* is the commitment to repay once a loan has been received. Credit gives people the freedom to choose to get a loan and thus to become indebted when it is in their own best interests. A person can have credit—access to loans—and choose not to borrow. The option to get a loan changes the choices of the household and makes the household better off even if the household never chooses to get a loan (Schreiner, Graham, and Miranda, 1998). Of course, credit can lead to loans and thus to debt, much as employment in a job can lead to work.

Finance encompasses both credit and savings. *Savings* is consumption postponed. Savings is a flow, the change in the stock of wealth. *Deposits* are one form of wealth held as financial assets and entrusted to a third party to be kept safe and perhaps to be invested. Not all savings are accumulated as deposits; the piggy bank of the poor is often a live pig in the backyard, and the few cents they save is the slop they give each day as feed to make the pig fat.

The instrument takes care to measure finance in all of its subtle forms. For example, it asks not just for the loans received from all sources in all of 1997 and for debt from all sources outstanding at the end of 1997; it also asks for all sources from which the respondent believes he or she could borrow if wanted. This measures credit,

loans, and debt. To measure savings and deposits, the instrument asks for deposit accounts and for other assets that maintain at least some of their worth through time.

Finally, the supervisors and the team of enumerators are well versed in the instrument and in rural surveys in El Salvador. Three of the four field supervisors led teams for both of the nationwide rural surveys. The fourth supervisor has years of experience with the Census Bureau. Of the 16 enumerators, 13 were chosen from among the best enumerators that worked in the 1998 nationwide survey. The three new enumerators have worked with FUSADES in other surveys.

All the enumerators were trained, not just in interview technique, but in the questions the instrument is meant to help answer and in the theoretical framework that gives the reasons why each item is asked. This helps to ensure that the enumerator catch responses to items that are not answers to the question intended. Each enumerator also got a manual custom-made for this survey (FUSADES, 1998). The manual describes how best to convince a sampled borrower to respond. It also tells ways to answer the quintessential questions that respondents ask about the purpose of the survey and how it will help the respondent without at the same time divulging that the survey will visit just rural borrowers of Calpiá.

3.4. Complete enumeration of the population

The sample for this survey was drawn at random from the complete, known population of rural borrowers of Calpiá. Few surveys can boast of this. Thanks to the

kindness of Calpiá in the supply of data on all their rural borrowers, the data from the survey will represent the rural portfolio of Calpiá in a known, standard way. The standard assumptions of the statistical tests will hold. The next part details how the sample was drawn.

The support of Calpiá in the form of a list of all of their rural borrowers brings with it the responsibility to protect their trade secrets. The names and addresses of the respondents will not be disclosed. Unlike most rural MFOs, Calpiá is private and is not a not-for-profit firm. Calpiá competes with many other lenders, both public and private, including a lender housed in FUSADES. The names and addresses of the rural borrowers in the sample are written on the paper instrument so that the supervisors and enumerators can find the respondents. FUSADES will keep the completed instruments, and no one outside of the OSU-FUSADES team will have access to them. The public-use data files to be released will not contain names or addresses.

4. How was the sample drawn?

The design of the sample had two goals. The first was to draw a random sample that would represent the rural borrowers of Calpiá. The second was to construct the sample so as to replace lost cases in the field without loss of the randomness of the sample.

4.1 A random sample

To draw a random sample required three choices. The first was the sample size: How many cases to sample? The second choice was the scope: How to define *rural*? The third choice was the stratification: How to divide the sample into subsamples?

4.1.1 Size

In principle, sample size depends on the precision required, on the number of subsamples, and on the variability and the size of the population. In practice, the sample size depends on the budget. More cases give more precise results, but more cases cost more.

The sample size n was set at 240 cases. The goal was to ensure enough cases to have a chance for statistically significant results in regressions and to be able to analyze each subsample with non-parametric statistics.

4.1.2 Scope

The definition of *rural* used for the survey was driven by the need to draw a sample that could be compared with the samples from the nationwide rural surveys in

1998 and 1996. This meant the use of the definition of the census of 1992. This is not the definition of *rural* that would be used in principle, nor is it the definition used by Calpiá in their work-a-day practice.

4.1.2.1 Rural in principle

In principle, *rural* means the opposite of urban. The main characteristic of rural areas is dispersed people. Of course, low population density is itself relative. At least some areas in El Salvador have high population densities even though the areas are seen as rural.

Ruralness is also linked with some other features. For example, most rural areas lack a network of streets and do not have a dense presence of public services and government offices or churches. Low population densities make it difficult to dilute the fixed costs of streets or other public services. Without an exact definition of rural in principle, no one knows just how many borrowers of Calpiá are rural.

Rural is almost always linked to the practice of agriculture. In fact, agricultural finance is often equated with rural finance, although the two are not always the same. Most people like to live near other people, however, so they do not choose to live in rural areas unless they hunt or farm since these acts require untrammeled spaces of land.

4.1.2.2 Rural in the practice of Calpiá

Calpiá defines *rural* based on its own needs to run its business. In essence, borrowers are rural if they are assigned to a rural loan officer. A rural loan officer has some expertise in agriculture. As a rule of thumb, rural loan officers work with borrowers who live more than 20 kilometers from a branch office. The rule is flexible, and a borrower classed as rural could live in an urban area more than 20 kilometers from the nearest Calpiá branch. Furthermore, a borrower in a rural area less than 20 kilometers from the nearest branch could be classed as rural.

In the first few months of 1998, Calpiá gave FUSADES a list of what it counts as rural borrowers to help in the first stages of the design of the sample. At that time and by its rule of thumb, Calpiá had 5,611 active rural borrowers and 4,326 inactive rural borrowers. The list did not tell the number of loans disbursed for farm projects. The list included place of work but not place of residence.

4.1.2.3 Rural in the practice of the census of El Salvador

The 1992 census of El Salvador used its own definition of rural. It is based on the basic geo-political organization of the country (Figure 1 on page 31).

El Salvador is divided in 14 *departamentos*, more than 262 *municipios*, hundreds of *cantones*, and thousands of *caseríos*. Each *departamento* is composed of *municipios*, one of which is a capital (*cabecera de departamento*). Each *municipio* is composed of a capital (*cabecera de municipio*) and *cantones*. For one *municipio* in each *departamento*,

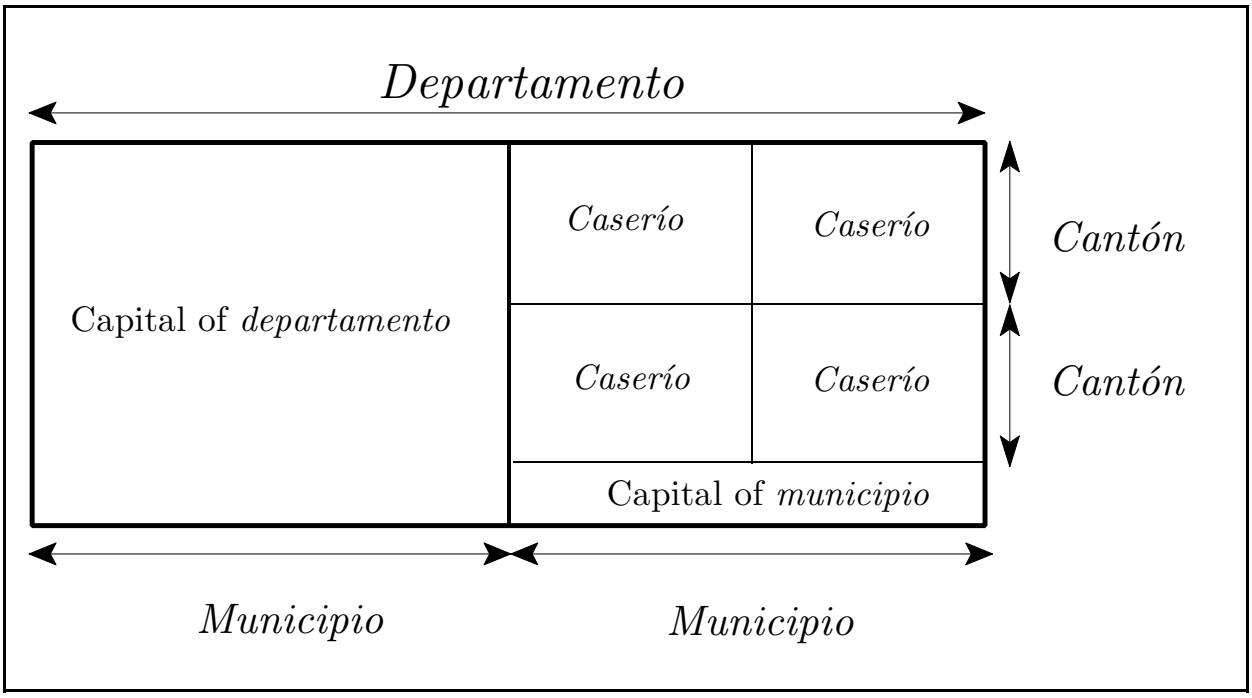

Figure 1: Geo-political divisions in El Salvador

the capital of the *municipio* coincides with the capital of the *departamento*. Each *cantón* is composed of *caseríos* (Figure 1 on page 31).

The census of 1992 of El Salvador defined a household as *rural* if it had its home in a *cantón*. This is the definition used in the 1996 and 1998 nationwide rural surveys, and it is the definition adopted for this survey of the rural borrowers of Calpiá.

There are some exceptions. Urban sprawl has swallowed some *cantones* in El Salvador. Some of these were recognized by the census of 1992 as urban (Table 1 on page 32). For example, *cantón* Santa Elena in *municipio* Antiguo Cuscatlán in *departamento* La Libertad is the area of the office of FUSADES. It is about as urban as it gets in El Salvador. According to the census of 1992, all households in all other

Departamento	*Municipio*	*Cantón*
La Libertad	Antiguo Cuscatlán	Santa Elena
La Libertad	Antiguo Cuscatlán	El Espino
La Libertad	Antiguo Cuscatlán	La Puerta de la Laguna
La Libertad	Colón	Lourdes
San Salvador	Soyapango	All *cantones*
San Salvador	Ayutuxtepeque	All *cantones*
San Salvador	Cuscatancingo	All *cantones*
San Salvador	Mejicanos	All *cantones*
Ahuachuapán	San Francisco Menendez	Cara Sucia

Table 1: *Cantones* counted as urban by the census of 1992

cantones were rural. This was the case in this survey, despite growing urbanization of some *cantones*.

4.1.3 Stratification

The sample design accounts for three different ways in which rural borrowers of Calpiá are likely to differ from each other. First, borrowers active as of April 1998 probably differ from inactive borrowers. Second, new borrowers are not likely the same as repeat borrowers. Third, borrowers whose loans are for crops or livestock probably differ systematically from borrowers whose loans are not for agriculture. Furthermore, borrowers who get loans for crops are not the same as those who get loans for livestock.

4.1.3.1 Active versus inactive

Active borrowers had loans outstanding as of April 1998. These loans could be in good standing or in arrears. In the population of 4,789, borrowers who live in cantones, as discussed below, 2,672 (56 percent) were active and 2,117 (44 percent) were inactive (lines 2 and 3, columns g and h of Table 2 on page 36). The sample aimed to have about 70 percent active borrowers (169 cases) and about 30 percent inactive (72 cases) (lines 2 and 3, columns d and f).

A rural borrower with a loan outstanding in April 1998 is unlikely to be the same as a borrower who got a loan in 1994, paid it off, and never went back to Calpiá a second time. Inactive borrowers are of interest because they are likely to be those who have suffered shocks, those who misjudged the worth and cost of a loan, or those who defected to an alternative source. Like all microlenders, Calpiá may not make a profit until a borrower has taken several loans (Gonzalez-Vega *et al.*, 1998). Thus drop-outs matter a lot.

4.1.3.2 New versus repeat

New borrowers are those whose last loan in 1997 was their first loan. Active borrowers were classed as new or repeat. Thus the classes for new and repeat are the same as the classes for active/new and active/repeat (Table 2 on page 36). In the population, 541 (11 percent) were active and new and 2,131 (45 percent) were active and repeat (lines 7 and 8, columns g and h). The sample aimed for about 25 percent

new borrowers (59 cases) and about 45 percent repeat borrowers (110 cases) (lines 7 and 8, columns d and f).

New borrowers cost a lender a lot. The lender spends to screen the borrower before approval, and after disbursement the lender watches the loan with care. At the same time, revenues from first-time loans are low since the loans tend to be short and small. Once the lender knows the borrower, however, the cost to screen and to monitor decreases. Revenue generally increases since the lender feels safe with longer and bigger loans.[2] Differences between new and repeat borrowers matter since the lender usually wants to avoid borrowers who test with one loan and then do not come back a second time. Development of a client relationship matters a lot in Calpiá's lending technology.

4.1.3.3 Crops, livestock, and non-agriculture

Agricultural borrowers put either crops or livestock (including fishing) as the purpose of the loan on their application. In the population, 2,699 clients (56 percent) borrowed for crops, 985 (21 percent) borrowed for livestock, and 1,105 (23 percent) were non-agricultural borrowers (lines 4, 5 and 6, columns g and h of Table 2 on page 36). The sample originally aimed to have about 70 percent agriculture and about 30 percent non-agriculture borrowers. After the field work, the researchers discovered that Calpiá's definition of agriculture included only crops and that livestock had been included in non-agriculture. To facilitate the analysis, the decision was made to redivide up the

[2] The companion note discusses the behavior of rural loan size at Calpiá (Schreiner, Gonzalez-Vega, *et al.*, 1998).

sub-samples into three classes: crops, livestock, and non-agriculture proper. This in no way influences the randomness of the sample, but the proportion of non-agriculture declined to 15 percent. That is, the sample includes 168 cases of crops (70 percent), 36 cases of livestock (15 percent), and 37 cases of non-agriculture (15 percent) as shown in lines 4, 5 and 6, columns d and f of Table 2.

The purpose of the loan matters since, as a rule, agriculture is subject to more risk than non-agriculture and loans for crops are seen as riskier than loans for livestock. Also, the lender cannot cheaply know whether failed harvests or dead cattle are due to risks such as weather beyond the control of the borrower or due to lack of effort of the borrower. It is more difficult, therefore, to control for moral hazard. Furthermore, agriculture has its own pattern of cash flows. A vendor can get a loan, buy inventory, and make weekly or monthly repayments out of the revenues from steady sales. In contrast, a farmer can get a loan to plant a crop, to kill its weeds, and to make a harvest. Long stretches of time come between all of these acts, and the farmer has no revenue from the project to fund loan repayments until harvest. Unless the farmer has other sources of cash inflows, the repayment of a loan for agriculture must differ from that of a loan for commerce (Gonzalez-Vega, 1997).

	Super-sample	Reg.	Sub.	Total	% Sub.	% sample	Pop.	% Pop.	Sample/Pop.	Pop./Cases
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)
Formula	Data	Data	Data	b+c	100*c/d	100*d/241	Data	100*g/4,789	d/g	1/i
Single strata										
1. Whole sample	321	211	30	241	12	100	4789	100	0.05	19.9
2. Active	225	147	22	169	13	70	2672	56	0.063	15.8
3. Inactive	96	64	8	72	11	30	2117	44	0.034	29.4
4. Crops	148	149	19	168	11	70	2699	56	0.062	16.1
5. Livestock	76	28	8	36	22	15	985	21	0.037	27.4
6. Non-ag	97	34	3	37	8	15	1105	23	0.033	29.9
7. New	79	51	8	59	14	24	541	11	0.109	9.2
8. Repeat	146	96	14	110	13	46	2131	44	0.052	19.4
Two strata										
9. Active/Crops	86	103	15	118	13	49	1061	22	0.111	9
10. Active/Livestock	71	21	5	26	19	11	747	16	0.035	28.7
11. Active/Non-ag	68	23	2	25	8	10	864	18	0.029	34.6
12. Inactive/Crops	62	46	4	50	8	21	1638	34	0.031	32.8
13. Inactive/Livestock	5	7	3	10	30	4	238	5	0.042	23.8
14. Inactive/Non-ag	29	11	1	12	8	5	241	5	0.05	20.1
15. Active/New	79	51	8	59	14	24	541	11	0.109	9.2
16. Active/Repeat	146	96	14	110	13	46	2131	44	0.052	19.4
Three strata										
17. Inactive/Crops	62	46	4	50	8	21	1638	34	0.031	32.8
18. Inactive/Livestock	5	7	3	10	30	4	238	5	0.042	23.8
19. Inactive/Non-ag	29	11	1	12	8	5	241	5	0.05	20.1
20. Active/Crops/New	27	36	5	41	12	17	133	3	0.308	3.2
21. Active/Crops/Repeat	59	67	10	77	13	32	928	19	0.083	12.1
22. Active/Livestock/New	28	2	2	4	50	2	141	3	0.028	35.3
23. Active/Livestock/Repeat	43	19	3	22	14	9	606	13	0.036	27.5
24. Active/Non-ag/New	24	13	1	14	7	6	267	6	0.052	19.1
25. Active/Non-ag/Repeat	44	10	1	11	9	5	597	12	0.018	54.3

Source: Authors' calculations

Source: Computed by the authors.

Table 2: Population, super-sample, sample, and sub-sample frequencies and weights

4.2 A strict system to pick substitutes for lost cases

The care used to draw the sample did not prevent problems in the field. Some rural borrowers of Calpiá refused to answer questions. Others, especially inactive borrowers, were not found. These facts of survey life present two challenges to the design and analysis of the survey. First, the sample must be drawn in such a way as to plan for substitutes for lost cases while still being random. Second, the analysis must account for the fact that the borrowers who do not respond most likely are not the same as the borrowers who do respond.

4.2.1 Substitutes

A total super-sample of 321 cases were drawn with the goal of completing 240 observations. After the field work, 241 interviews had been completed (line 1, columns a and d in Table 2 on page 36). The super-sample included a primary group and a substitute group. Both groups were drawn at random without replacement. Once drawn, each case in the substitute group was assigned a number at random. The substitutes entered the primary group in the order of this number when a case from the primary group was lost. This ensures that the substitutes are not those who are easiest to find but rather are drawn from the population at random.

4.2.2 Sample selection

Sample selection occurs when cases refuse to answer or cannot be found in a non-random way (Greene, 1993). This makes the sample non-random, even if the lost

cases are replaced with substitutes drawn at random. For example, the richest borrowers may tend to refuse more since their time is worth more to them. Likewise, the address of inactive borrowers is more likely to be wrong than the address of active borrowers. Borrowers in arrears may have more fears to answer.

The problem with sample selection is generally that the traits of the lost cases are not known. If the traits of the lost cases were known, then the analysis could relate them with the fact of being lost and then correct for the non-randomness induced by the sample selection. But when the traits of the lost cases are unknown, then there is no way to adjust the analysis.

The dangers of sample selection are attenuated in the case of the rural borrowers of Calpiá. This is due to the fact that a plethora of the traits of lost cases can be gleaned from files maintained by Calpiá. The analysis will be adjusted for any correlation between these traits and lost cases.

4.3 The sample

4.3.1 Rural borrowers

In late April 1998, Calpiá gave OSU a list of all its borrowers, past and present, that lived in *cantones*. This list had 5,558 borrowers. The final population from which the sample was drawn deleted 5 borrowers who tested the instrument, 382 borrowers in *cantones* defined as urban by the census of 1992, and 382 borrowers whose first loan was in 1998.

The sample was drawn at random from the population of 4,789 borrowers that remained. In the super-sample of 321 cases, the primary group has 241 cases and the substitute group has 80 cases (line 1, columns a and d, Table 2 on page 36). About five percent of the population is in the primary group (line 1, column i). This means that each case interviewed represents about 20 rural borrowers, the case itself and 19 others (line 1, column j).

After the field work, 211 interviews were conducted with clients in the original primary group and 30 corresponded to substitutes (line 1, columns b and c). This reflected the persistent efforts of the interviewers to seek the clients in the original primary sample. The substitutes represent only 12 percent of the observations and their share is large for inactive livestock borrowers (lines 1 and 13, column e).

Estimates derived from the sample as a whole should be weighted if they are meant to estimate population parameters. This is due to the fact that the sample proportions of the strata do not match the population proportions (Deaton, 1997). This will be done systematically in this research.

4.3.2 Three types of combinations of the three strata

4.3.2.1 Classes within a single stratum

The sample can be divided into groups by each of the three strata (lines 2 through 8 of Table 2 on page 36). Each division includes all the cases sampled. For example, the active borrowers (169) and the inactive borrowers (72) together sum to

241. Compared with the population frequency, the design oversampled active, new, and borrowers with crops. This can be seen in two ways. First, sampled borrowers in these classes represent fewer non-sampled borrowers (column j). Second, these cases are a greater percentage of the sample total (column f) than of the population total (column h).

Statistics found for one of the classes in these single-level combinations of strata cannot be extrapolated to the same class in the whole portfolio of rural borrowers because of the differences in class proportions in the sample and the population. For example, it would be wrong to claim that the sample median loan for active borrowers is the best estimate of the median loan for all active borrowers in the population. The class of active borrowers in the sample differ in their traits in the classes of agriculture/non-agriculture and new/repeat as compared with the whole population.

New and borrowers for crops bear too much weight in the sample of the active class since they have been oversampled within their own single strata. For example, if new and borrowers for crops had smaller loans than repeat and non-agricultural borrowers, then the sample median for active borrowers would be lower than the true population median.

The best estimate of the population parameters from the sample will give less weight to the active borrowers who are in the oversampled classes of the other strata—new and crops—and more weight to the active borrowers in the undersampled

classes. Most analyses of most surveys ignore the design of the sample. The care taken in the analysis here will help the results to be believed and useful. It is better to make policy choices based on good estimates than on bad estimates.

4.3.2.2 Combinations of two strata

The sample can also be divided into groups by combinations of two of the three strata (lines 9 through 16 of Table 2 on page 36). Not all of these divisions include all the cases sampled just once. For example, active/new (59) and active/repeat (110) exclude all inactive cases (lines 15 and 16, column d). These cases add up not to 241 but to 169. All five classes of actives double-count the actives and omit the inactives. The division of active/new (59), active/repeat (110), inactive/crops (50), inactive/livestock (10), and inactive/non-ag (12) does count each sampled borrower just once. These cases sum to exactly 241. The same holds for active/crops, active/livestock, inactive/crops, inactive/livestock, and inactive/non-ag.

As in the case of classes with a single stratum, statistics derived from classes of two combined strata must be weighted since they contain proportions of the third stratum that do not match the population proportions. For example, the unweighted median of the sample cases in active/crops is not the best estimator of the median of the population of active/crops.

4.3.2.3 Combinations of three strata

The nine classes formed from the combinations of all three strata (lines 17 through 25 of Table 2 on page 36) together count all sampled borrowers just once. The classes of inactive/crops, inactive/livestock, and inactive/non-ag are like combinations of three strata since they ignore the distinction between new and repeat. The most heavily sampled classes reflect the goals of the research. For example, each case interviewed for active/crops/new stands for about 3 borrowers in the population (line 20, column j). Likewise, each case for active/crop/repeat stands for about 12 cases in the population (line 21, column j). Inactive, livestock, and non-ag classes are undersampled.

These nine classes are the primary sampling units. That is, the sample was drawn by first splitting the population into these nine classes and then drawing a random sample from each one. The size of each random sample at the level of triple combinations was set so as to achieve the goals for the sampling proportions chosen for the three strata at the level of single combinations. For example, the design called for 70 percent of the sample to be active. The number of cases in the six active classes (lines 20 through 25, column d) is 169, about 70 percent of 241. Likewise, 30 percent of the sample was to be for other than crops, and the number of cases in the livestock and non-agriculture classes (lines 5 and 6, column d) is 73, about 30 percent of 241. Because livestock had not been separated from non-agriculture in the beginning, in the end some

categories are very small despite the attempt to set up the divisions to get a decent amount in each category.

Unlike the classes of single and double combinations of strata, statistics derived for a class of triple combinations do not need to be weighted to be the best estimate for the population in that class. For example, the median estimated from the sample of the active/crops/new class is the best estimate of the median of that class in the population, given the sample. The classes at the level of triple combinations do not need weights since they do not leave out any strata within their class.

4.3.3 How rural is the sample?

The borrowers in the sample and in the population are not completely rural in all ways. For example, 247 of the 4,789 borrowers who live in a *cantón* nevertheless work in an urban area and yet are classed as rural by Calpiá. Furthermore, the 4,789 borrowers who live in a *cantón* are not the only borrowers of Calpiá with some degree on ruralness. In particular, the sample excludes borrowers who work in *cantones* (perhaps cultivate a farm) but live in urban areas. A lower bound on the number of cases in this class is 858.[3]

[3] This is a lower bound since it comes from a comparison of the early 1998 list of rural borrowers using the definition of Calpiá with the April 1998 list based on borrowers who live in *cantones*. The cross is not complete since Calpiá might have borrowers who work in a *cantón* but who it does not count as rural. In general, however, place of residence seems to be a better criterion of ruralness than place of work.

Of the 9,937 rural borrowers by the definition of Calpiá, 4,436 lived in a *cantón*. This suggests that 4,789–4,436=353 of the borrowers counted as rural by the census are not counted as rural by Calpiá. In contrast, more than half of the clients counted as rural by Calpiá are not counted as rural by the census. This does not mean that Calpiá has tricked the microfinance community worldwide with a rural portfolio inflated with non-rural borrowers. It just means that the definition of rural matters for the measurement of outreach to rural borrowers.

4.4 Conclusion

This note has described the methods used in a survey of rural clients of Financiera Calpiá in El Salvador. A companion note illustrates some of the issues raised here by using the information available to draw the sample to examine some of the traits of the population of rural borrowers of Calpiá and to verify how well the sample represents the population as a whole.

For the reasons discussed in this note, a unique data set is being constructed in El Salvador. This data set will have a higher value as a public good because it is being uniquely documented. Further research will then assess the unique contribution of Financiera Calpiá to the development of rural microfinance in Latin America.

		Calpiá		
		Urban	**Rural**	**Total**
Census	**Urban**	N/A	5501	N/A
	Rural	353	4436	4789
	Total	N/A	9937	N/A

Table 3: Rural and urban borrowers as defined by the census and Calpiá

References

Adams, Dale W. (1993) "Transactions Costs in Decentralized Rural Financial Markets", Economics and Sociology Occasional Paper 2093, The Ohio State University.

_____. (1988) "The Conundrum of Successful Credit Projects in Floundering Rural Financial Markets", *Economic Development and Cultural Change*, Vol. 36, No. 2, pp. 355-368.

_____. (1971) "Agricultural Credit In Latin America: A Critical Review of External Funding Policy", *American Journal of Agricultural Economics*, Vol. 53, No. 2, pp. 163-172.

Adams, Dale W; Graham, Douglas H.; and J.D. Von Pischke. (1984) *Undermining Rural Development With Cheap Credit*, Boulder: Westview Press.

Baydas, Mayada M.; Cuevas, Carlos E.; Graham, Douglas H.; and Richard L. Meyer. (1994) "Capital Structure Determinants among Non-farm Enterprises: The Case of Developing Financial Markets in the Gambia", manuscript, The Ohio State University.

Brigham, Eugene F.; and Louis C. Gapenski. (1993) *Intermediate Financial Management, Fourth Edition*, Fort Worth: Dryden Press.

Buckley, Graeme. (1997) "Microfinance in Africa: Is it Either the Problem or the Solution?", *World Development*, Vol. 25, No. 7, pp. 1081-1093.

Carter, Michael R. (1997) "Environment, Technology, and the Social Articulation of Risk in West African Agriculture", *Economic Development and Cultural Change*, Vol. 45, No. 3, pp. 557-591.

Chaves, Rodrigo A.; and Claudio Gonzalez-Vega. (1996) "The Design of Successful Rural Financial Intermediaries: Evidence From Indonesia", *World Development*, Vol. 24, No. 1, pp. 65-78.

Chen, Martha Alter; and Elizabeth Dunn. (1996) "Household Economic Portfolios", AIMS, Washington, D.C.: USAID.

Deaton, Angus. (1997) *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Baltimore and London: Johns Hopkins University Press.

_____. (1992a) "Household Saving in LDCs: Credit Markets, Insurance and Welfare", *Scandinavian Journal of Economics*, Vol. 94, No. 2, pp. 253-273.

_____. (1992b) "Saving and Income Smoothing in Cote D'Ivoire", *Journal of African Economies*, Vol. 1, pp. 1-24.

_____. (1991) "Savings and Liquidity Constraints", *Econometrica*, Vol. 59, No. 5, pp. 1221-1248.

Dunn, Elizabeth. (1997) "Diversification in the Household Economic Portfolio", AIMS, USAID.

Efron, Bradley; and Robert J. Tibshirani. (1993) *An Introduction to the Bootstrap*, New York: Chapman and Hall.

Fafchamps, Marcel. (1993) "Sequential Labor Decisions Under Uncertainty: An Estimable Household Model of West-African Farmers", *Econometrica*, Vol. 61, No. 5, pp. 1173-1197.

Fafchamps, Marcel; Udry, Christopher; and Katherine Czukas. (1998) "Drought and Saving in West Africa: Are Livestock a Buffer Stock?", *Journal of Development Economics*, Vol. 55, pp. 273-305.

Fundación Savadoreña Para el Desarrollo Económico y Social. (1998) "Manual del Entrevistador. Encuesta Servicios Financieros a Familias Rurales—1998", San Salvador.

_____. (1996) "Encuesta Desarrollo Rural", San Salvador.

Gersovitz, Mark. (1983) "Savings and Nutrition at Low Incomes", *Journal of Political Economy*, Vol. 91, No. 5, pp. 841-855.

Glomm, Gerhard; and Michael G. Palumbo. (1993) "Optimal intertemporal consumption decisions under the threat of starvation", *Journal of Development Economics*, Vol. 42, pp. 271-291.

Gonzalez-Garita, Marco A. (1986) *Farmer Borrowing Costs: The Case of Costa Rica*, unpublished Master's thesis, The Ohio State University, 1986.

Gonzalez-Vega, Claudio. (1998) "Research Planning Framework: Central America Regional Program," Madison, Wisconsin: BASIS.

Gonzalez-Vega, Claudio. (1997) "The Challenges of Transfering the New Lending Technologies to Reach the Poor from the Urban to the Rural Areas: Questions from Theory, Lessons from Microfinance Best Practice," XXIII International Conference of Agricultural Economists, Sacramento, California.

Gonzalez-Vega, Claudio. (1996) *Microfinanzas en El Salvador. Lecciones y Perspectivas*, San Salvador: Fundaungo.

Gonzalez-Vega, Claudio. (1993) "From Policies, to Technologies, to Organizations: The Evolution of the Ohio State University Vision of Rural Financial Markets", Economics and Sociology Occasional Paper No. 2062, The Ohio State University.

Gonzalez-Vega, Claudio and Marco A Gonzalez-Garita. (1987) "Costos de Endeudamiento, Tasas de Interés y Distribución del Crédito en Costa Rica," in Dale W Adams, Claudio Gonzalez-Vega and J.D. Von Pischke (eds.) *Crédito Agrícola y Desarrollo Rural: La Nueva Visión*, San José: Ohio State University.

Gonzalez-Vega, Claudio; Schreiner, Mark; Meyer, Richard L.; Rodriguez-Meza, Jorge; and Sergio Navajas. (1998) *An Ohio State Primer on Microfinance in Bolivia*, manuscript, The Ohio State University.

Gonzalez-Vega, Claudio; Prado Guachalla, Fernando; and Tomás Miller Sanabria. (1997) *El Reto de las Microfinanzas en América Latina: La Visión Actual*, Caracas: Corporación Andina de Fomento.

Gonzalez-Vega, Claudio; Jimenez Rodriguez, Ronulfo; and Rodolfo Quiros Rodriguez. (1996) *Financing Rural Microenterprises: FINCA-Costa Rica*, The Ohio State University, Academía de Centroamérica, and Inter-American Foundation. Published in Spanish as *Financiamiento de la Microempresa Rural: FINCA-Costa Rica*.

Gonzalez-Vega, Claudio; Meyer, Richard L.; Navajas, Sergio; Schreiner, Mark; Rodriguez-Meza, Jorge; and Guillermo F. Monje. (1996) "Microfinance Market Niches and Client Profiles In Bolivia", Economics and Sociology Occasional Paper No. 2346, The Ohio State University.

Greene, William H. (1993) *Econometric Analysis: Second Edition*, New York: MacMillan Publishing Company.

Guerrero, Jose A. (1988) *Determinants of Successful Rural Deposit Mobilization: Banco Agricola in the Dominican Republic*, unpublished Master's thesis, The Ohio State University.

Interdisziplinare Projekt Consult. (1993) "El modelo de bancos comunales como alternativa de financiamiento para micro and pequenos establecimientos rurales en Nicaragua: Analysis de las experiencias de los proyectos FINCA-Honduras y CAM-El Salvador", manuscript.

Jacoby, Hanan G. (1994) "Borrowing Constraints and Progress Through School: Evidence From Perú", *Review of Economics and Statistics*, Vol. 76, No. 1, pp. 151-160.

Jovanovic, Boyan. (1982) "Selection and Evolution in Industry", *Econometrica*, Vol. 50, No. 3, pp 649-670.

McCloskey, Deirdre N.; and Stephen T. Ziliak. (1996) "The Standard Error of Regressions", *Journal of Economic Literature*, Vol. 34, pp. 97-114.

McPherson, Michael A.; and Joan C. Parker. (1993) *A Manual For Conducting Baseline Surveys of Micro- and Small-scale Enterprises*, Gemini Tools for Microenterprise Programs: Field Research Section Paper No. 1, Washington, D.C.: Development Alternatives International.

Mead, Donald C.; and Carl Liedholm. (1998) "The Dynamics of Micro and Small Enterprises in Developing Countries", *World Development*, Vol. 26, No. 1, pp. 61-74.

Meyer, Richard L., and Carlos E. Cuevas. (1990) "Reducing the Transaction Costs of Financial Intermediation: Theory and Innovations", paper prepared for the International Conference on Savings and Credit for Development, Copenhagen, May 28-31.

Ministerio de Coordinación del Desarrollo Económico y Social. (No date) "Priorización de Municipios a partir de Datos Censales", San Salvador: Dirección General de Política Económica y Social, Dirección de Información.

Ministerio de Economía. (1995) "1995 Encuesta de Hogares de Usos Múltiples", San Salvador: Dirección General de Estadística y Censos.

Morduch, Jonathan. (1994) "Poverty and Vulnerability", *American Economic Review*, Vol. 84, No. 2, pp. 221-225.

_____. (1993) "Risk, Production, and Saving: Theory and Evidence From Indian Households", manuscript, Harvard.

Navajas, Sergio; Schreiner, Mark; Meyer, Richard L.; Gonzalez-Vega, Claudio; and Jorge Rodríguez-Meza. (1998) "Microfinance and the Poorest of the Poor: Theory and Evidence From Bolivia", manuscript, The Ohio State University.

Ouattara, Korotoumou; and Douglas H. Graham. (1996) "Rural Finance in the Republic of South Africa: Selected Farm Household Case Studies", report to the Development Bank of Southern Africa.

Paxson, Christina H. (1992) "Using Weather Variability To Estimate the Response of Savings to Transitory Income in Thailand", *American Economic Review*, Vol. 82, No. 1, pp. 15-33.

Quintanilla, José Dimas; and Germán Rioseco. (1995) "El Financiamiento para la Microempresa Rural: El Caso de El Salvador," manuscript, San Jose: RUTA III.

Quirós-Rodríguez, Rodolfo. (1991) *Access to Agricultural Credit in Costa Rica*, unpublished Master's thesis, The Ohio State University.

Reardon, Thomas *et al.* (1998) "Rural Nonfarm Income in Developing Countries. Importance and Policy Implications," in *The State of Food and Agriculture 1998*, Rome: Food and Agriculture Organization.

Rosenweig, M.; and K. Wolpin. (1993) "Credit Market Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-income Countries: Investments in Bullocks in India", *Journal of Political Economy*, Vol. 101, No. 2, pp 223-244.

Rutherford, Stuart. (1998) "The Poor and Their Money: An Essay About Financial Services For Poor People", manuscript.

Sanchez-Schwarz, Susana. (1996) *Assortive Matching of Borrowers and Lenders: Evidence From Rural Mexico*, unpublished Ph.D. dissertation, The Ohio State University.

Sanchez, Susana; Cuevas, Carlos; and Rodrigo Chaves. (1998) "El Salvador: Rural Finance: Performance, Issues, and Options", Report No. 16253-ES, Washington, D.C.: World Bank.

Schreiner, Mark. (1997) *A Framework For the Analysis of the Performance and Sustainability of Subsidized Microfinance Organizations With Application to BancoSol of Bolivia and Grameen Bank of Bangladesh*, unpublished Ph.D. dissertation, The Ohio State University, `www.microfinance.com`.

Schreiner, Mark; Graham, Douglas H.; and Mario Miranda. (1999) "Choices by Poor Households when the Interest Rate for Deposits Differs from the Interest Rate for Loans", pp. 348-355 in G.H. Peters and Joachim Von Braun (eds) *Food Security, Diversification, and Resource Management: Refocusing the Role of Agriculture?*, Aldershot: Ashgate, ISBN 1-84014-04405.

Schreiner, Mark; Claudio Gonzalez-Vega *et al.* (1998) "Notes on Features of the Rural Borrowers of Financiera Calpiá in El Salvador," `www.microfinance.com`.

Silverman, B.W. (1986) *Density Estimation for Statistics and Data Analysis*, New York: Chapman and Hall.

Smith, Eldon. (1996) "Methodology for Agricultural Economists: An Instrumentalist Perspective", *American Journal of Agricultural Economics*, Vol. 78, pp. 906-915.

Townsend, Robert M. (1994) "Risk and Insurance In Village India", *Econometrica*, Vol. 62, No. 3, pp. 539-591.

Udry, Christopher. (1995) "Risk and Saving in Northern Nigeria", *American Economic Review*, Vol. 85, No. 5, pp. 1287-1300.

Vasquez, Archibaldo. (1986) *Determinants of Household Deposit Behavior In the Dominican Republic*, unpublished Master's thesis, The Ohio State University.

Von Pischke, J.D. (1991) *Finance at the Frontier: Debt Capacity and the Role of Credit in the Private Economy*, Washington, D.C.: World Bank.

World Bank. (1997) *El Salvador: Rural Development Study*, Report No. 16253-ES, Washington, D.C.

Yaron, Jacob. (1994) "What Makes Rural Finance Institutions Successful?" *World Bank Research Observer*, Vol. 9, No. 9, pp. 49-70.

Zeller, Manfred. (1994) "Determinants of Credit Rationing: A Study of Informal Lenders and Formal Credit Groups in Madagascar", *World Development*, Vol. 22, No. 12, pp. 1895-1907.

Zeller, Manfred; Schreider, Gertrud; Von Braun, Joachim; and Franz Heidhues. (1997) *Rural Finance for Food Security for the Poor: Implications for Research and Policy*, Food Policy Review No. 4, Washington, D.C.: International Food Policy Research Institute.