Assessing the Impact of Policy Reforms in Ugandan Microfinance

Craig McIntosh*

June 9, 2002

Abstract

This paper utilizes a new impact analysis methodology to measure the effects of two innovations introduced by Uganda's largest microfinance practitioner in April 2000. In one innovation, a voluntary health-insurance package was bundled in with loans, and in the other greater flexibility was extended to groups over the terms of their loan. Since both programs were voluntary, and a spatial treatment-control strategy was used for each program, we are provided with two control groups: the non-choosers and the not offered. By using both controls, we can simultaneously examine the impact of the programs and the substantial spatial shocks which occurred during the test. Both programs are found to increase client volume, and the results also suggest that that the frequent repayment required by most microfinance programs may not be needed in order to prevent delinquency. Additionally, the results provide circumstantial evidence that a significant portion of savings in microfinance institutions are precautionary savings intended to cushion against health-related shocks.

*Department of Agricultural and Resource Economics, 317 Giannini Hall, University of California, Berkeley, CA 94720. Phone: (510)643-5414. Fax: (510)643-8911. Email: mcintosh@are.berkeley.edu. Thanks to Guido Imbens, Alain de Janvry, Ethan Ligon, Elisabeth Sadoulet, Michael Ward, and seminar participants at U.C. Berkeley

1 Introduction

The past decade has witnessed explosive growth in the use of microfinance as a tool for getting credit into the hands of the world's poor. Until very recently, the terms of lending contracts have been almost exclusively supply-driven, and borrowers were forced to make due with one-size-fits-all loan programs. The past few years, however, have seen saturation of many of the most promising lending markets, and as microfinance institutions increasingly square off directly against one another, they have been induced to provide their clients with more customized, varied products. FINCA International is one of the oldest and most geographically diversified microfinance lenders in the world. Part of the key to their success has been their 'Village Banking' lending contract which is consistent across countries, and was well-suited to the early, expansionary phase of microfinance markets. Uganda is typical of many lending markets in that there is rapidly expanding competition, and the major cities of Kampala and Jinja now see at least a dozen major institutions vying for clients. In this context, it has become vital for FINCA Uganda to introduce new lending products that allow them to respond to borrower needs, and so to retain clients who now have a wide variety of alternatives. This paper examines the impact of two experiments in demand-driven products for Ugandan borrowers (a health insurance product and 'Flexibility' determining loan terms), and in so doing hopes to contribute to best practice in the field.

The tests were conducted by dividing the country into three parts; each program was tested in one region and the third area served as the control for both programs. This methodology made the test managable to implement, but also exposes standard difference-in-difference (DID) impact estimates to bias from any imperfectly observed spatial phenomena that were coincident with the test. While one might attempt to mitigate this bias by including as many control variables as possible, one cannot directly correct for spatial effects because the identification of the treatment relies precisely on unexplained spatial heterogeneity. Because administrative units were used as the treatment/control units, we might expect this problem to be even worse, since differences in managerial quality or data collection can only be controlled for by unit-level fixed effects that would be completely collinear with the impact dummy in a DID regression.

The methodology of this paper takes advantage of the fact that the treatments were voluntary, and thus the non-choosers of the program provide a second control population. Specifically, under an assumption about the comparability of choosers and non-choosers, this second population allows us to establish the following counterfactual: what spatial effects would we expect participants in the program to have experienced in the absence of the treatment? By establishing this counterfactual for the location of every agent who chose the treatment, and subtracting it off of the dependant variable, we can recover an unbiased treatment estimate even in the presence of unexplained regional effects. We show how to test for the presence of such spatial effects, and then implement a spatial matching technique to remove them if they are found to be present. Thus we are able simultaneously to investigate the outcome of several substantial shocks that occured during the test (an ebola epidemic and the runup to a presidential election) and to estimate the impact of the treatments independant of those shocks. In addition, by observing these two phenomena separately, we are able to identify a kind of impact that is invisible to standard techniques, namely the ability of a treatment to insulate agents from shocks.

2 Review of the Literature

Much of the theoretical interest in microfinance has focused on the properties of the joint-liability contracts through which borrowers without collateral are able to leverage credit as a group. Stiglitz (1990) wrote a seminal article which showed that although joint-liability contracts shift risk from neutral lenders to risk-averse borrowers, such contracts can still be welfare-improving. Ghatak (1999) shows that these contracts will induce assortative matching, whereby high- and low-quality clients will endogenously select into groups with similar types of clients. Besley & Coate (1995) show that joint liability can address both adverse selection, through mobilizing local information, and moral hazard, by making punishment regimes more stringent than could be achieved in the absence of joint liability. The tradeoffs between direct monitoring and the indirect monitoring accomplished by joint-liability has been another focus of the theoretical literature. Stiglitz (1998) introduced the first model of this tradeoff, and show that if there are scale effects in lending (due to fixed costs) that subsidies to lenders may actually raise the costs of borrowing in the marketplace. Conning (1999) models the lending problem in a principal-agent framework whereby the amount that can be spent on monitoring is directly related to the returns from a loan, which induces higher interest rates and higher employee costs as an equilibrium when lending to poorer clients.

Of the empirical issues which the literature has attempted to address, perhaps the least satisfactory answers have been those to the simplest question: does microfinance work? This branch of the microfinance literature has been plagued by the many subtle, unobservable types of selection bias which are present when potential clients make their borrowing decisions. Thus, even if we can identify a village which is perfectly comparable to a treatment village, whom within the village should be used as a control? Perhaps the most serious effort to identify this treatment effect is to be found in a paper by Pitt & Khandker (1998), who use the ineligibility of those with large landholdings to identify a treatment effect among those who qualified. This paper estimates that household expenditures rise by 18 percent of the principal when women borrow, and by 11 percent when men borrow. Evidence of the lack of fungibility within the household comes from the significant effect of credit in the hands of women upon a wide variety of household outcomes, and the absence of significant effects when men borrow analagous amounts. The results of the paper are based, however, on the unpalatable assumption that landholding is exogenous.

A related empirical question involves the degree to which microfinance institutions are, and should be, subsidized in their operations. Morduch (1999) shows that the 'sustainability' of the Grameen Bank in Bangladesh is much further off than claimed; a detailed analysis of the Bank's balance sheets reveal a range of implicit subsidies to the institution. Morduch (1997) and Morduch (2000) make further attempts to puncture the common idea that microfinance presents a win-win situation in which there are no trade-offs between the goals of helping the poor and those of attaining financial sustainability. At present, the pressure from donors to achieve rapid sustainability is growing, and so the current state of the market involves escalating competition, decreasing subsidies, a slow drift towards sustainability in most institutions, and an increasing focus on individual lending and client-driven product innovations. Rhyne & Christen (1999) present a good summary of these trends, using data gleaned from most of the world's major microfinance marketplaces.

This paper, then attempts to add to the literature by conducting a systematic assessment of two important new products in microfinance. The Flexibility in determining loan terms can be thought of as a step in the direction of the customization of lending products. The Health Insurance product is an experiment in using microfinance institutions as conduits for a richer variety of financial services to the poor, hopefully generating synergies along the way between the use of credit as a productive tool and the use of credit as a form of insurance. Because we compare one type of microcredit to another, our control groups are well-defined, and so we sidestep the hurdle present in assessing the impact of credit versus no credit.

3 The Innovations

FINCA Uganda is the oldest microfinance institution in the country and one of the largest and best-established in Africa. Their standard lending product utilizes a group-lending methodology, wherein all members of 'Village Banks' are jointly liable for each others' loans. It lends almost exclusively to women. There is no formal screening of new clients, so membership in groups is constrained only by the selection imposed by current clients on members of their community for whom they will accept liability. Loans begin at 50 dollars, and subsequent increases are based on fixed multiples of clients' savings determined by the client's grade, which in turn is based on repayment and attendance of weekly meetings. The standard loan has a 16-week cycle, and clients pay 4 percent per month flat interest (87 percent effective). Each client is covered by a life-insurance policy, whose premiums are included in the interest payments. FINCA Uganda now has almost 25,000 clients in almost 1,000 Village Banks, spread over most of the conflict-free parts of Uganda.

At the beginning of 2000, FINCA Uganda instituted two new policies. The first, the so-called 'Flexibility Program' was designed to give groups substantial say over the terms of their loans. Under Flexibility, eligible groups could elect (by a unanimous vote) to change the length of the lending cycle or the frequency of repayment of their loans. The majority of the interest among clients was in the Biweekly repayment program; since observations were very limited in the other components of Flexibility, we have discarded groups that received them and this analysis focuses on the impact of Biweekly payment. While it is relatively obvious that making fewer payments would be preferable to clients, there is a widespread perception amongst both practioners and clients of microfinance that frequent repayment is key to the high rates of repayment. Thus, a primary concern of those groups which switch to biweekly repayment is whether the reliability of repayment drops. Weekly repayment also places a very tight cash-flow constraint on client businesses; from anecdotal evidence, the amount which can be repaid in the worst typical week often determines what clients are willing to borrow. Thus, we are also interested in seeing if biweekly repayment causes loan volume to increase or precautionary savings to drop.

The second new policy offered a Health Insurance package to Village Banking clients and their families. As an attempt to control adverse selection, more than sixty percent of the clients in any village banking group were required to enroll in the program for it to be offered to the group. The package costs roughly 15 dollars per cycle and covers the client plus four dependants and a husband for any routine medical treatments. Uganda is an environment characterized by high mortality, and an extremely large fraction of FINCA clients were caring for children orphaned by AIDS or other diseases. More endemic diseases such as malaria and dysentery are very common, and in addition many older Ugandans suffer from hypertension brought on by a starch-heavy diet. Consequently, even in areas with poor health care, medical costs can constitute a major burden for poor families.

Crucial to understanding the impact of the Insurance program is understanding the shortcomings of its implementation. The first major problem is that, since the country lacks a unique national ID system, it is very difficult to verify identity. Thus, there is a strong suspicion that children who did not belong to FINCA clients were receiving care under the plan. Secondly, the participating hospitals were reimbursed by DFID (the underwriter of the program) regardless of the costs that they incurred, and so had every incentive to provide a very high level of service to FINCA clients. The combination of these two factors led to severe cost overruns; the price of the program was raised subsequent to the end of this test but was still not expected fully to cover costs. This is an important fact, as it means that participation in the insurance program was more than risk avoidance; it actually lowered the expected costs of medical treatment for clients that participated. Since FINCA does not check how the money it lends is used, it is entirely likely that some portion of borrowing and savings behavior is engaged in intertemporal transfers to cushion to the household against disease-related shocks. So we are interested not only in seeing how borrowing behavior changes when that source of risk is removed, but in seeing if the insurance program has the ability to cushion participants against shocks that are picked up in neighboring agents that were not covered by insurance.

The period during which the test took place contained several events which we might expect to be non-randomly distributed across space. An ebola outbreak occurred in October of 2000 and the disease was contained to the northern part of the country; the only town which reported cases in which FINCA operates is Masindi. An unexpectedly close presidential election led to insecurity and some minor rioting in the capital, but had little impact on business in other parts of the country. Both of these events are genuine 'shocks', however the ongoing political unrest and the extreme ethnic and climatological variation across the study area are certainly typical of applied contexts in Africa and the developing world in general. The areas included in this experiment comprise three major ethnic subfamilies and around 20 distinct languages; rainfall, transportation quality, arbitrage, and competition from other micro-finance institutions all vary widely as well. While all of these contributors to spatial heterogeneity are, in theory, observable, it is difficult to see how we should proceed in trying to 'dummy out' these shocks, or else these shocks are identical with the treatment area and so cannot be removed through dummies. The standard double-difference approach would involve making the untestable assumption that these shocks had an identical effect on outcomes in the treatment and control areas. We show that this assumption is not warranted in our data; and while we cannot determine within-sample whether these spatial effects arise due to shocks, non-linearities, or omitted variables, the spatial residuals are not related in any obvious way to the shocks that we were able to observe. Thus, it would appear that such spatial effects are not only a result of shocks particular to this dataset, but also are due to underlying specification problems that are likely to be found in many applied contexts in the developing world.

Perhaps the most severe spatial problem involves the Insurance program, which was implemented only by two hospitals, one in the capital and the other in the third-largest town in Uganda. This means that the treatment area for this program essentially constitutes the urban portion of two of the country's three largest cities, and the control is the remainder of the country. It is not at all unreasonable to think that urban regions in Africa have a different rate of growth from the surrounding rural areas; if this is the case then a DID approach to such a treatment/control strategy will be biased. Through the use of spatial matching, however, we are able to separately estimate the underlying growth rates of each part of the country, and thus recover a treatment effect which is based on the correct counterfactual.

We attempt to estimate impact on the lending institution only; as for individual impacts we take a position of rational expectations and revealed preference to argue that any innovation which was accepted by a group made that group better off on average. The question of the impact on the institution is essentially a cost-benefit analysis. The Biweekly program substantially reduces costs of lending to the institution, and so as long as it is not found to damage the institution in other respects, we deem it beneficial. The Insurance program, while underwritten by another agency, does impose indirect administrative costs on FINCA, and so we must consider the bar to be higher; meaning that only if it demonstrably improves outcomes for FINCA as a lender should it be considered a successful program purely from the perspective of a microfinance agency. We make no claims to have included the benefits of the program on client health outcomes in this analysis.

4 Data

The unit of analysis is the 'Village Banking' group. The data is taken from the accounts of FINCA Uganda, and from surveys conducted at the individual and the group level during the test. In addition, FINCA credit officers were asked to identify the location of all of their groups on maps, and this data has been converted into GIS format in order to allow for spatial analysis.

FINCA tracks the entire population of 1000 groups from its six regional branches. Arua branch, the newest, is far off in the politically unstable northwest of the country, and has been dropped from the analysis due to lack of comparability. Because the use of a DID regression requires pre-test data, only groups which were in existence for at least one cycle prior to the beginning of the test can be included. These restrictions reduce the number of usable groups to around 450. Data was missing from roughly 10 percent of the surveys from the remaining parts of the country, and so these observations have been dropped. Groups that are located inside the treatment area for one program but that did not receive the treatment (e.g. those offered the treatment who rejected it) are used as a part of the control group for the other treatment. Groups that received one treatment are eliminated from the sample used to test the other treatment, which reduces the overall number of observations to roughly 350 for each program.

Since the analysis requires partitioning the data into groups of choosers and non-choosers, it is vital that we correctly establish counterfactuals over choice; namely which among the groups that was not offered the treatment would have chosen it? The chooser/non-chooser status of groups in control areas was established through a mock election on each of the innovations which was conducted as a part of the group survey. The majority voting rules were applied in the same way that they had been in the treatment (e.g. 60 percent required for health insurance, and unanimity for biweekly repayment). Following Rosenbaum (1982), the four groups that are relevant for each analysis can be categorized as follows:

	Choosers	Non-choosers
Treatment	Offered and Accepted	Offered and Unaccepted
Control	Unoffered and Accepted	Unoffered and Unaccepted

A DID analysis will compare choosers in the treatment to choosers in the control, and throw away the data on the non-choosers. We utilize the spatial information present in the second control group, the non-choosers, to establish spatial counterfactuals at the location of each chooser.

The following outcome variables are used in the analysis:

1. Dropout (percent that took a loan last cycle and do not return)
2. New clients (percent of clients starting this cycle that are new)
3. Grades (attendance of weekly meetings and repayment performance)
4. Average loans in a village bank
5. Average savings in a village bank

The control variables used are:

1. Loan cycle number (e.g. loans taken by this VB)
2. Ethnic homogeneity of the group
3. Borrower's perception of their local business climate
4. A dummy equal to one if the VB is in a rural area
5. The average number of children in clients' households
6. The average number of non-working adults in clients' households
7. The share of clients in a VB that own their own homes
8. Does the VB conduct other informal internal savings and lending
9. Did the group pre-exist in some form prior to formation of the VB

5 Econometric Specification

The methodology outlined here is based upon McIntosh (2002). That paper introduces a flexible Gaussian smoother which nests the approach of this paper and the standard DID into one regression model, and performs experiments on simulated data. It is found that matching to more than one agent does not sufficiently reduce variance to compensate for the bias introduced. These results are consistent with Dehejia & Wahba (1998) and others who have experimented with matching to multiple agents. Thus, this paper proceeds with a matching approach that conducts no local smoothing and instead treats only the closest agent as the indicator of 'local' conditions.

Agents are indexed by i, and each agent is located in some space s_i. The dependent variable, which represents changes in outcomes, is denoted Y_{si}. These changes are explained by some function of observable variables X_{si}, a vector of unobserved variables Z_{si}, and by location in space s_i. A dummy T_{si} indicates whether or not the treatment was offered, and a dummy ω_i indicates whether it was accepted. We define the treatment criterion as a rule that maps $\{X, Z, s\} \in \Re^N \mapsto \{0, 1\}$. We denote this mapping by τ, so $T_{si} = \tau(X_{si}, Z_{si}, s_i)$. The treatment criterion selects a subset of agents for whom $\omega_i = 1$ and subjects them to a treatment effect $t_{si}(X_{si}, Z_{si}, s_i)$. In our case, since physical space defines the treatment criterion, we can write $T_s = \tau(s_i)$, and we assume that there are no spillover effects, so $t_{si} \neq 0 \Longrightarrow T_s = \omega_i = 1$.

In the absence of a treatment effect, the function which explains changes in outcomes can be written as:

$$Y_{si} = f(X_{si}, s_i, Z_{si}) + \epsilon_{si}.$$

When we explain Y with a linear regression, we have

$$Y_{si} = \beta X_{si} + \phi(X_{si}, s_i, Z_{si}) + \epsilon_{si},$$

where $\phi(X_{si}, s_i, Z_{si})$ is the entire systematic component of outcomes that is orthogonal to the linear specification on observables. Without loss of generality, we can think of the treatment effect experienced during a test as an additive term which is some function of observables and unobservables. So, outcomes among choosers are actually explained by

$$Y_{si} = \beta X_{si} + \phi(X_{si}, s_i, Z_{si}) + t_{si}(X_{si}, Z_{si}, s_i) + \epsilon_{si},$$

but we run

$$Y_{si} = \beta X_{si} + \delta T_s + \mu_{si} \ \forall \ i \ s.t. \ \omega_i = 1$$

and so $\hat{\delta} = \bar{\delta} + P_{\tau(s)}(\phi)$, where $\bar{\delta}$ is $E(t_{si} \mid T_s = \omega_i = 1)$, the true average treatment effect and $P_{\tau(s)}(\phi)$ is the projection of the unexplained systematic component into the space spanned by the treatment dummy. Thus, we will recover the correct ATE only under the condition that $\phi \perp \tau(s)$, or that the unexplained systematic component of outcomes is equal in the treatment and the control regions. So any shocks, omitted variables that affect rates of change, or non-linear relationships that differ between treatment and control will bias the standard DID approach. The presence of these shocks cannot be tested for within the sample of choosers jointly with an impact analysis since they rest on the same identification.

The DID approach proceeds under an assumption about the comparability of the treatment and control regions, which we wish to replace with an assumption about the comparability of two populations. Letting Y_{si}^{C} represent outcomes among those who chose the treatment, and Y_{si}^{NC} among those who did not choose the treatment, we make the following assumption:

$$\phi_{si}^{C} \mid s_i = \phi_{si}^{NC} \mid s_i.$$

This says that any part of the unexplained effect which is due entirely to location should be the same for two agents located in the same place. In other words, if non-choosers in a certain city in the treatment area had loans that are 20 dollars higher than those elsewhere in the country controlling for other factors, then we should expect the treated groups to have loans that are 20 dollars higher as well, and should ascribe a treatment effect to the program in this region only if outcomes are more than 20 dollars higher. To be able to identify $\phi_{si}^{NC} \mid s_i$, we need to additionally assume that $(t_{si} \mid \omega = 0) = 0$, or that there is no indirect spillover effect of the treatment upon the untreated.

Given these two assumptions, a quick way to test for spatial effects is to run the 'False DID' regression:

$$Y_{si}^{NC} = \beta^{NC} X_{si}^{NC} + \gamma T_{s^{NC}} + \mu_{si}^{NC}$$

Since we assume that there was no treatment effect among this group, and that the spatial effects in the two groups are comparable, then for the DID to be unbiased, the estimate of γ should be zero. If the spatial distribution of C and NC are identical, then we can estimate impact directly by subtracting γ from δ. In general the distributions are not exactly the same, and so we proceed by 'backfitting' the vector of residuals. This is necessary because, in general, $Cov(X, s) \neq 0$, and so the residuals we estimate on the first iteration are 'too small' because clustering in the observable variables have proxied for space and so absorbed some of the spatial effect. Thus, failing to backfit will impose an additional, unnecessary assumption about the spatial distribution of observables being the same between the two populations. We proceed as follows: First, we estimate the regression that explains outcomes among non-choosers, then we smooth the resulting residuals across space. We then subtract this conditional expectation off of the dependent variable, and iterate until the residuals contain no spatial component. The betas which arise from this iteration are then used to predict a set of residuals that have the full degree of spatial variation in them.

Thus, the backfitted residuals are

$$\hat{\mu}_{si}^{NC} = \epsilon_{si} + \phi_{si}^{NC},$$

where ϵ is an i.i.d. error term, and so

$$E(\hat{\mu}_{si}^{NC} \mid s) = \phi_{si}^{NC} \mid s.$$

Outcomes among choosers of the program can be explained by

$$Y_{si}^{C} = \beta X_{si} + t_{si}(X_{si}, s_i, Z_{si}) + \phi_{si}(X_{si}, s_i, Z_{si}) + \epsilon_{si}.$$

We match each chooser i to the closest non-chooser, who will be denoted by i', and subtract off the residual which we estimated for the nearest neighbor.

$$E(Y_{si}^{C} - \hat{\mu}_{si'}^{NC} \mid s) = \beta X_{si} + t_{si}(X_{si}, s_i, Z_{si}) + \phi_{si}(X_{si}, s_i, Z_{si}) - \phi_{si} \mid s_i + \epsilon_{si}.$$

The crucial factor for our purposes is that the term $\phi_{si}(X_{si}, s_i, Z_{si}) - \phi_{si} \mid s_i$ can be written as $\phi_{si} \perp s_i$, and since we know that $T_{si} = \tau(s)$, then $(\phi \perp s) \perp \tau(s)$. In other words, since the treatment is a function of space, if we can make the unobserved elements of the problem orthogonal to space, then we have made them orthogonal to the treatment and so we have recovered an unbiased impact estimate even in the presence of totally unexplained shocks that are coincident with the treatment.

6 Spatial Effects

Under the assumption that the unexplained spatial effects for the two groups will be the same, then we can begin to examine the nature and degree of spatial effects by looking at the distribution of outcomes for the non-chooser group, none of whom received the treatment. We will investigate this issue in three different but related ways. The results of the previous section are based on the fact that $(\phi \mid s) = 0 \; \forall \; s \Rightarrow (P_{\tau(s)}(\phi) = 0)$, but we note that that the reverse is not necessarily true. It is also the case that $(P_{\tau(s)}(\phi) \neq 0) \Rightarrow ((\phi \mid s) \neq 0)$ for some s but again the reverse is not necessarily true. This means that (given the same spatial distribution) finding a non-zero false impact term is a sufficient but not necessary condition for the presence of spatial heterogeneity, and finding spatial heterogeneity is a necessary but not sufficient condition for a non-zero false impact term. This is because we may have spatial effects that balance perfectly between the treatment and control, and so they will not project into the treatment dummy. If we have projection into the treatment dummy, however, there must be spatial heterogeneity present.

We take loan volume among non-choosers of the Biweekly treatment as a case study. First, we conduct a False DID regression, which (under the assumption of common spatial effects and no spillovers) is measuring $P_{\tau(s^{NC})}(\phi)$. We find a significant positive effect, which means that given the locations of the non-choosers, there are spatial effects that are non-randomly distributed across the treatment and control regions.

We proceed to investigate these spatial effects among non-choosers by creating a moving average composed of residuals from the twelve nearest neighbors to every non-chooser. In Fig. 1, we examine loan volume among non-choosers of the biweekly flexibility. The view is from the south, over Lake Victoria; round dots represent the location of VBs, and wherever topographical features are visible it means that the contour in question is negative. As we would expect we see enormous spatial variation in raw loan volumes. Primarily, we see a huge spike in the average loan over the capital city, and the rural areas without exception appear as troughs in this surface.

In Fig. 2, we examine the surface of changes in loan volume. It is interesting that this surface is in many ways the inverse of the previous one. This indicates a negative second derivative in the slope of the function that explains loan growth; meaning that the higher loan volumes are, the lower the increases in loans. Indeed, the correlation coefficient between loan volume and loan growth in this population is -.36. Consequently, we see high points on the surface in the outlying, rapidly expanding rural towns, while loan growth in the capital and in adjacent cities is very slow.

In Fig.3, we examine the surface of residuals from changes in loans once we have controlled for changes using our full vector of observable variables. Our False DID regression measures the average difference between this surface and the zero-plane in the treatment and control regions at the location

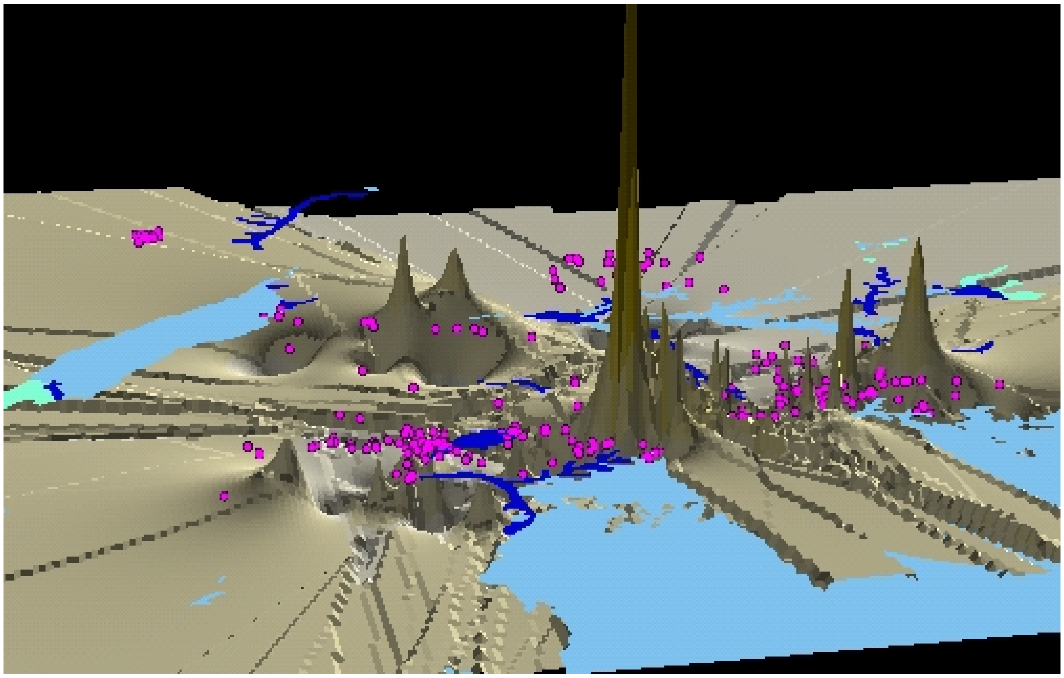

Figure 1: Spatial distribution of Raw Loan Volumes

of each non-chooser. According to the assumptions of the DID estimator, this contour should be completely flat, or at least have the same average deviations in the treatment and control.

So, we have established in two different ways that there are substantial spatial variations in the distributions of changes in loans for the non-choosers. What we are interested in for the purposes of the impact analysis, however, is the expected value of these spatial shocks at the places where the choosers are located. Given the difference in spatial dispersion between the two populations, even under our assumptions we would not in general expect $P_{\tau(s^{NC})}$ to equal $P_{\tau(s^{C})}$. To investigate this issue, we first match each chooser to the residual that is estimated at the location of the closest non-chooser. We then sample with replacement from the empirical marginal distribution of the residuals, randomizing over location, and so we bootstrap datasets that are spatially i.i.d by construction. We then run a spatial smoother (a Gaussian smoother with a bandwidth of .2 standard deviations) over these bootstrapped datasets, and select the envelope that contains 95% of the smoothed surfaces. Finally, we smooth our observed nearest-neighbor residuals with

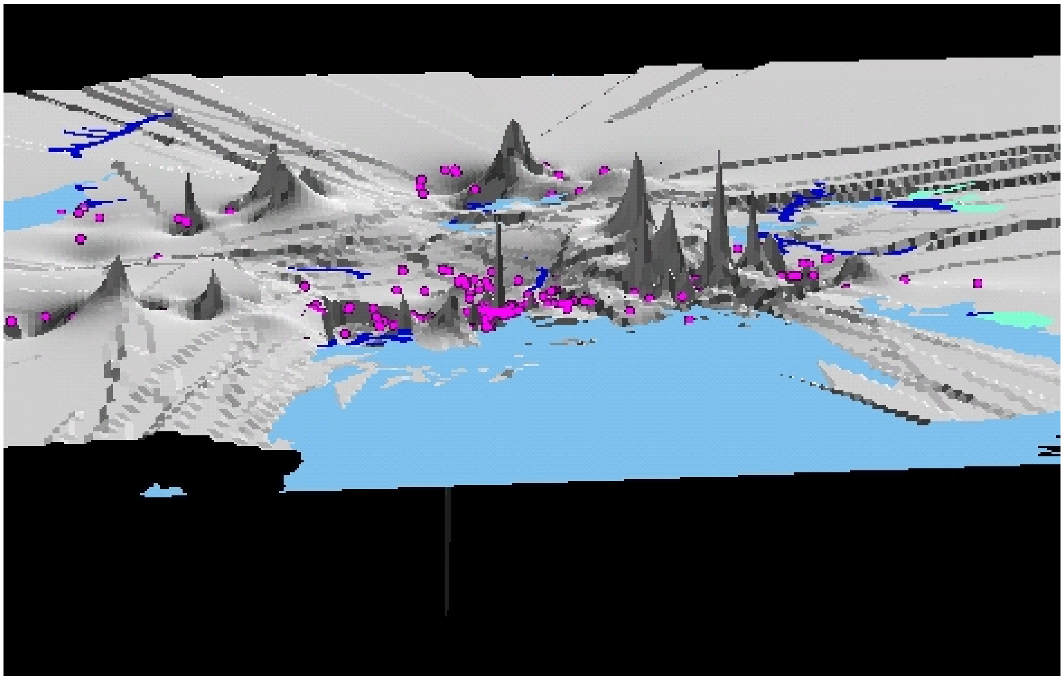

Figure 2: Spatial distribution of Raw Loan Changes

the same smoother, and then compare these pointwise smoothed outcomes with the pointwise smoothed i.i.d. outcomes. Any observations that lie below the confidence region have experienced a significant negative shock, and any observations above have experienced a positive shock. Figure 4 shows the locations of all of the choosers. Those observations which experienced insignificant shocks are denoted with an X, and those with significant shocks are denoted by an arrow pointing in the direction of the shock. It is clear by inspection that the area around and to the west of Kampala, which is the control region, experienced more negative shocks, while the northern and eastern sections of the country, the treatment, experienced a disproportionate number of positive shocks. Specifically, while the treatment and the control contain six groups each that experienced significant negative shocks, the treatment region contains twelve groups that experienced positive shocks, while the control region contains only two. This is consistent with our finding of a significant positive treatment effect in the False DID. We thus proceed confident that we do indeed experience significant spatial heterogeneity in our data, and thus the standard DID estimates will be biased (in this case

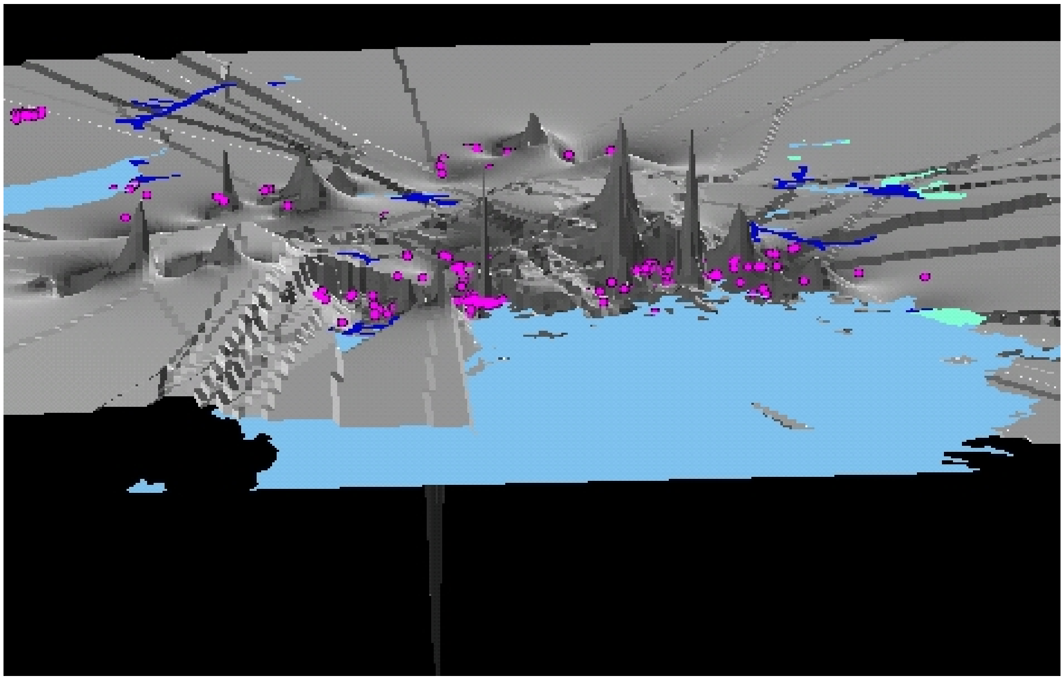

Figure 3: Spatial distribution of Loan Change Residuals

upwards).

7 Bias Correction & Double Counterfactuals

A recent paper by Abadie & Imbens (2001) suggests two techniques that are useful in this application. The first is a bias-correction technique which can be implemented when matching. The basis of the idea is that if we are not picking a perfect match, then we introduce some bias into estimates which is based on the difference in location between the agent and the match. Since the error in this match is observable, we can predict the effect of the error based on the overall spatial distribution of residuals. Letting i index the agent and i' be i's nearest neighbor, we can proceed by estimating

$$\hat{\mu}^N = \lambda s^N + \epsilon$$

where there are the same number of elements in λ as in s. We use the information contained in λ by modifying the estimated shock at each location

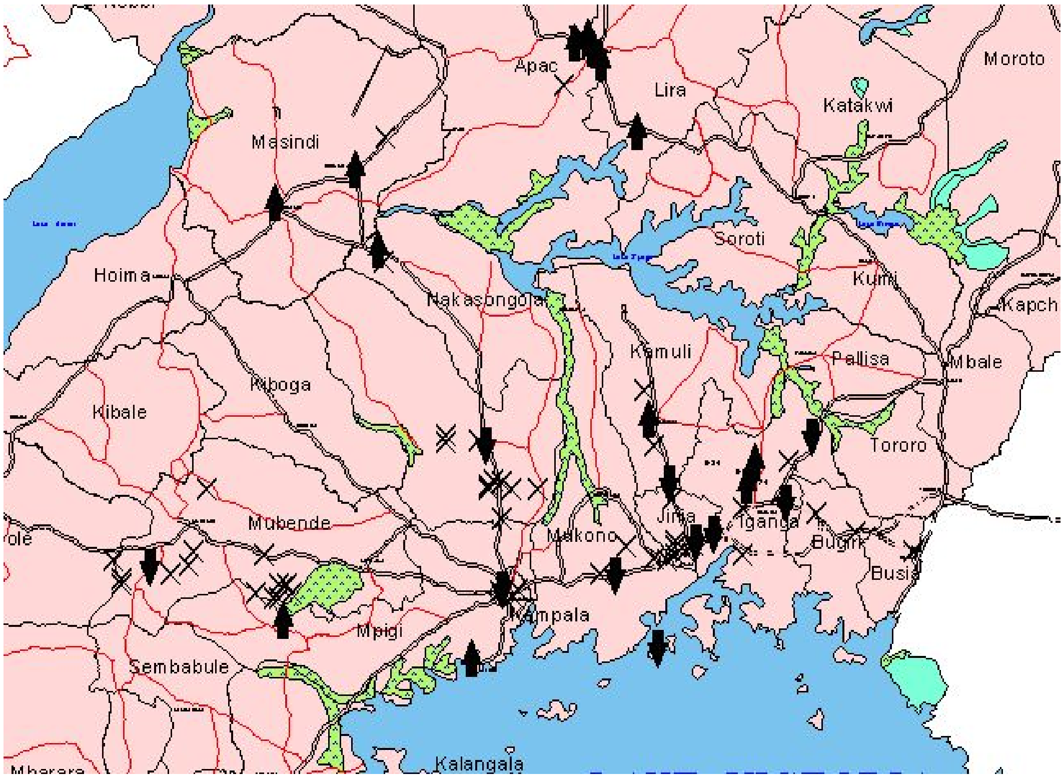

Figure 4: Location of Significant Shocks for Choosers

in the study population by the distance to the nearest neighbor; so

$$\hat{\phi} = \hat{\mu}^N + \hat{\lambda}(s_i - s_{i'}).$$

The distance between i and i' is calculated separately in every element of s, so we use latitude and longitude as separate distance measures, since in general the expected shift in residuals from moving ten miles east will not be the same as that from moving ten miles north. The bias correction has been found in simulations to generate some improvements in estimates, and it does so at no cost in terms of degrees of freedom, and so it is implemented here.

The second technique taken from that paper is to establish a double counterfactual; namely, not only do we estimate what all of the choosers would have looked like in the absence of the treatment, but we also estimate what all of the non-choosers would have looked like had they received the treatment. In practice, this means that not only do we match each of the choosers to the nearest non-chooser and subtract off a residual term, but we match each non-chooser to a chooser and subtract off a residual. Because the treatment effect is purely a function of space, and because we are able to backfit the residuals among both groups, we are able to construct a set of residuals for the choosers that includes the full impact term. Had we not backfit, and were the observables to be correlated with the treatment effect, then the residuals we would estimate from a regression performed on choosers without an impact term would not contain the full treatment effect as some of it would have passed into the betas. Through backfitting, however, we are able to generate the desired vector of residuals for both choosers and non-choosers, and thus the final regression which we use to estimate impact can be written in block matrix form as follows:

$$\begin{vmatrix} Y^C - \hat{\mu}_{i'}^{NC} \\ \\ Y^{NC} - \hat{\mu}_{i'}^{C} \end{vmatrix} = \begin{vmatrix} X^C & 0 & \text{T} \\ \\ 0 & X^{NC} & \text{-T} \end{vmatrix} \begin{vmatrix} \beta^C \\ \beta^{NC} \\ \delta \end{vmatrix} + \quad \epsilon$$

Thus, because we define impact as the difference in differences in changes between choosers and non-choosers in the treatment versus the control, we are able to use the full sample of observations to estimate the treatment effect.

8 Regression Results

It is important to note that the ATE for a voluntary program is an estimate of the impact of extending the choice of the program to the rest of the population, but is not an estimate of the effect of making the program mandatory. The difference arises from the selection effect, and probit regressions which explain the likelihood of participating in each of the treatments indicate that there is a strong selection effect present in both cases. Groups which selected the Insurance are more likely to be uneducated, small, to have older members, to have borrowed for a longer period of time, to have more clients with bad grades, to have had no pre-existing organization, and to have large numbers of adults in the household. There is no clear relationship to wealth of members, however. The Biweekly program, on the other hand, is most desirable to urban women with many children and few adults at home, meaning to those women for whom it is most difficult to be away from home to attend meetings. It was chosen by poorer groups in general, and by those with more young clients. Clearly, the Insurance program is most attractive to individuals with high expected health care costs, and the Biweekly program is most attractive to those with high opportunity cost of the time spent attending meetings. There is no reason to think that the impact of these programs, if made a mandatory part of the lending package for all of FINCA Uganda, would be the same.

We see that, using the DID approach, the programs have few significant effects. However, we also find some very powerful spatial heterogeneity present in the False DID. As expected, there is more spatial heterogeneity for the Insurance program than the Biweekly, because the treatment for the former is less representative of the country as a whole than for the latter. Although the DID shows an increase in loan volume under the Biweekly program which is significant at the 90% level, there is an even stronger effect in the FDID, indicating that the increase is a result of spatial effects rather than the program. Indeed, the nearest-neighbor matches show absolutely no effect of the program. Where we do see significant effects for the Biweek program are in strongly decreased dropout, and grades that are marginally higher. So, while we have not seen the jump in loan volumes that was predicted, we do observe that these groups which have tailored products to the clients have managed to become significantly more attractive to current members, and so have improved retention. The surprise of these results is that, rather than seeing Biweekly repayment worsen grades, it actually improves them. This may be explained by the fact that, since grades are calculated using the percentage of meetings which were attended and where payments were made, that halving the number of meetings increases the chance that each one will be successfully attended. The costs of lending are, obviously, dramatically lower when the credit officers halve the number of meetings they are required to oversee, and so if this transition leaves loan volumes unchanged and actually improves client quality, it suggests that extending Biweekly flexibility to the rest of the country will both increase the sustainability of the institution and help to tailor products towards the needs of groups.

The Insurance program has impacts which might best be termed insulative. In each case where the nearest-neighbor indicates a significant impact, we see no impact in the DID and a significant shock among the non-choosers. We see a huge jump in dropout and a fall in new client enrollment in the FDID, which implies that established, urban groups (e.g., the control) are increasing in size much more slowly than their rural counterparts (the treatment). While we have dummies to explain rural/urban status, cycle number of the group, and other variables which should control for this effect, we have apparently done so imperfectly. This might be a result of 'shocks', such as the upswing in instability in the capital prior to the elections, or the heightened need for credit in the rural north of the country where the ebola outbreak was experienced. On the other hand, it may simply be the result of differential rates of growth in these widely divergent environments which is some non-linear function of our control variables. We cannot test between these two stories within our data, but under our assumptions we need only estimate these spatial effects, not explain them. The upshot is that groups which took the Insurance did not experience the fall in new client enrollment that otherwise plagued urban regions, and so the Insurance had the effect of insulating participating groups against the otherwise lowered attractiveness of urban groups.

Even more interesting is the coefficient on savings for Insurance. We see a much faster rate of savings growth in the treatment than in the control in the FDID, but no such coefficient in the DID. Thus, we conclude that while in general urban groups saved more than rural groups, that those urban groups who had insurance did not. One explanation for this is that it simply reflects the cost of the program. There are several reasons to believe that this is not the case, however. One is that the value of the drop in savings is roughly 20% higher than the cost of the program. The second is that the premiums usually came out of the loan, not out of savings, and so we would expect this effect to be reflected in higher loans, not in lower savings. Consequently, we conclude that what we are observing is the result of decreased risk on households that had been engaging in precautionary savings to cushion themselves from health shocks. While we would expect precautionary savings to be highest among groups that chose this treatment, it is evidence that some non-negligible part of the savings which FINCA clients have are precautionary in nature.

	DID	FDID	NN	BCNN
Biweekly: # obs:	101	247	348	348
New Clients:				
	-5.912	0.7977	-2.7306	-2.8325
	(-1.0533)	(0.2278)	(-0.7583)	(-0.7867)
Dropout:				
	-3.8173	0.2198	-6.9059**	-7.7971**
	(-0.8263)	(0.0811)	(-2.5166)	(-2.8507)
Savings:				
	-1030.8	-2571.1	1202.7	1188.3
	(-1.1573)	(-1.3531)	(0.8493)	(0.8391)
Loans:				
	2605.4*	3355.7**	22.8	33
	(1.6376)	(2.6938)	(0.0197)	(0.0029)
Grades:				
	-0.1227	0.0157	0.1915*	0.1904*
	(-0.8129)	(0.1538)	(1.9345)	(1.9218)
Insurance: # obs:	117	199	316	316
New Clients:				
	0.2263	-4.203*	5.3226**	5.6981**
	(0.0549)	(-1.5523)	(2.0261)	(2.1729)
Dropout:				
	2.5454	8.6821**	3.1809	3.2743
	(0.508)	(2.64)	(0.9052)	(0.9279)
Savings:				
	87.1	4589.3*	-3176.2*	-3177.9*
	(0.0787)	(1.9425)	(-1.8427)	(-1.8439)
Loans:				
	-830.6	-426.8	45.7	28.1
	(-0.4485)	(-0.286)	(0.0337)	(0.0208)
Grades:				
	0.2357	-0.0142	-0.1106	-0.1105
	(1.3368)	(-0.1279)	(-0.9355)	(-0.9332)

(*=90% significance, **=95% significance)

9 Conclusion

FINCA's Village Banks experienced significant spatial effects during the course of this study. We are unable to identify whether these effects are the results of specification errors or of genuine shocks. What is certain is that a standard DID approach to a dataset with this type of heterogeneity across the treatment criterion can produce extremely biased results. Using the non-choosers of the program to eliminate spatial heterogeneity, we find the Biweekly repayment program improving repayment performance and retention of existing clients. Since it also lowers administrative costs, we can unambiguously conclude that extending this program to all of the groups in Uganda should increase the sustainability of the institution.

The Insurance program did not retain existing clients more successfully, but rather attracted new ones at a greater rate than control groups. In addition, it caused savings to decrease. While this indicates decreased need for buffering on the part of borrowers, it may damage the profitability of the institution as FINCA Uganda moves to intermediate savings under the new lending laws of Uganda. This transformation will allow FINCA to generate direct profits by onlending savings, as opposed to the current situation wherein savings are held in accounts at established formal sector banks. Thus, while it is likely that the insurance program, due to its imperfect pricing and enforcement, caused increases in the household welfare of participants, measured from the point of view of the lending institution it has few advantages. A study of individual clients who have joined groups after they received the Insurance product shows that these clients are low-quality borrowers (low grades, small loans, small growth in loans) and may indeed be participating in FINCA only to get access to the Insurance product. Thus while there is some advantage in that the program brings in new clients, those who are attracted seem to be of dubious value seen from the perspective of a lender. Thus, we conclude that the Insurance program has failed to generate synergies within FINCA, and so its extension to other parts of the country is not likely to improve institutional sustainability.

References

Abadie, A. & Imbens, G. (2001), Simple and bias-corrected matching estimators for average treatment effects. Unpublished Working Paper.

Besley, T. & Coate, S. (1995), 'Group lending, repayment incentives, and social collateral', *Journal of Development Economics* **46**, 1–18.

Conning, J. (1999), 'Outreach, sustainability and leverage in monitored and peer-monitored lending', *Journal of Development Economics* **60**, 51–77.

Dehejia, R. & Wahba, S. (1998), 'Propensity score matching methods for non-experimental causal studies', *National Bureau of Economic Research Working Papers* (6829).

Ghatak, M. (1999), 'Group lending, local information, and peer selection', *Journal of Development Economics* **60**, 27–50.

McIntosh, C. (2002), Spatial matching to control for nonrandom treatment assigment.

Morduch, J. (1997), Does microfinance really help the poor? new evidence from flagship programs in bangladesh.

Morduch, J. (1999), 'The role of subsidies in microfinance: evdence from the grameen bank', *Journal of Development Economics* **60**, 229–248.

Morduch, J. (2000), 'The microfinance schism', *World Development* **28**, 617–629.

Pitt, M. & Khandker, S. (1998), 'The impact of group-based credit programs on poor households in bangladesh. does the gender of participants matter?', *Journal of Political Economy* **106**(5), 958–995.

Rhyne, E. & Christen, R. P. (1999), 'Microfinance enters the marketplace'.

Rosenbaum, P. (1982), 'The role of a second control group in an observational study', *Statistical Science* **2**(3), 292–316.

Stiglitz, J. (1990), 'Peer monitoring and credit markets', *World Bank Economic Review* **4**, 351–366.

Stiglitz, K. H. . J. (1998), 'Moneylenders and bankers: Price-increasing subsidies in a monopolistically competitive market', *Journal of Development Economics* **55**, 485–518.