

Microfinance and Home Improvement: Using Retrospective Panel Data to Measure Program Effects on Fundamental Events

JEL Classifications: O12, O16, C21

Craig McIntosh, Gonzalo Villaran, and Bruce Wydick

McIntosh: University of California at San Diego
e-mail: ctmcintosh@ucsd.edu

Villaran: Invertir Consulting, Lima, Peru
e-mail: gvillara@invertir.org.pe

Wydick: University of San Francisco
e-mail: nydick@usfca.edu

April 2008

Abstract: Rigorously estimating the effects of development programs is notoriously difficult. In this paper we present a methodology that borrows from "event studies" commonly used in the finance literature to ascertain the impacts of corporate mergers. In our RETRAFECTION methodology a retrospective panel data set is created based on "fundamental" events in the history of surveyed households, events that are discrete, unforgettable, and important to welfare. Based on the relationship between the changes in the estimated probabilities of these events and the timing of the introduction and uptake of a treatment, it is possible to ascertain if the probability of these fundamental events changes significantly after treatment. Our paper presents a series of tests, diagnostics, and corrections to account for potential supply-side and demand-side endogeneity problems in program rollout and adoption, and we caution against causal inference in many types of programs unless specific conditions are satisfied. Our application of the methodology examines housing improvements and purchases of large consumer durables among 1,672 households who gained access to microfinance at different times in Guatemala, India, and Ghana. We carry out estimations using village and country/year-level fixed effects to analyze the timing of these fundamental events relative to the timing of microfinance access and treatment. In our estimations, we find that the probability of a major housing improvement almost doubles from 0.038 to 0.070 in the years subsequent to a household's first microfinance loan, and that the probability of purchasing certain consumer durables increases modestly.

The authors wish to thank Gabriel Dzathor, Alexis Unzeitig, Melanie Steger (Ghana), Eugene Henson, Megan Yash, and Jennifer Worley (India), and Adam Gorski and Karina Vargas (Guatemala) for insightful input into this research and data collection in the field. For valuable comments about this research methodology we would like to thank Alessandra Cassar, Philip Fanchon, Michael Jonas, Dean Karlan, David Levine, David McKenzie, Ted Miguel, Jeff Nugent, Jean-Philippe Plateau, Elizabeth Sadoulet, John Strauss, Xavier Giné, MFI directors Matthew Brown (Ghana), Jim Reiff (India), and Juan-Manuel Ovando (Guatemala), and seminar participants at the World Bank, USAID, the University of Rome, the University of Southern California, the 2007 Pacific Conference for Development Economics at UC Davis, and the 2008 Pacific Conference for Development Economics at UC San Diego. Grant funding from BASIS/USAID, the McCarthy Foundation, and the Jesuit Foundation is gratefully acknowledged.

1. Introduction

There has been much written recently by development economists about the need for rigorous and systematic appraisal of the effectiveness of anti-poverty programs in developing countries (for example, Duflo and Kremer 2003; Armendáriz de Aghion & Morduch, 2005; Savedoff, Levine, and Birdsall 2006; Easterly, 2006). Yet researchers and practitioners seeking to ascertain the true impact of development programs face a daunting task. Accurately appraising the effectiveness of development programs is both time-consuming and costly, especially for small institutions that seek accurate measures of their impact on clients. Moreover, many institutions would like to evaluate the effectiveness of their programs *ex-post* to implementation, creating problems with the establishment of baseline surveys, control groups, and other means of identification. These obstacles have created a demand for new approaches to determining the effectiveness of many types of development programs, including microfinance.

In this paper we present a methodology for ascertaining welfare changes associated with development programs which can be employed *ex-post* to program implementation, and may be applicable in a variety of contexts. Our methodology, a Retrospective Analysis of Fundamental Events Contiguous to Treatment (RETRAFECT), uses a single cross-sectional survey to create a retrospective panel data set based on fundamental events in the history of households. We define fundamental events as those events in a household's history that are discrete, unforgettable, and important to household welfare. Analyzing the timing of these events within a window around the timing of treatment allows for statistical tests over the relationship between the treatment and changes in household welfare. This methodology borrows from "event studies" undertaken in the finance literature, where the effect of events such as mergers and acquisitions are observed on stock prices. An excellent review of this use of event studies in finance is given in MacKinlay (1997). Here, however, instead of examining changes in

equity prices within a time window surrounding a merger, we examine the *probability* of fundamental events within the time window surrounding a treatment.

In presenting the methodology, we stress the fact that identifying the relative timing of events is not equivalent to measuring causal program impacts. However, we suggest a number of diagnostics which allow researchers to test for whether the supply-side rollout of the program, as well as the demand-side uptake of the program, is exogenous to impact variables. To the degree that survey data passes these diagnostics, or in the case that corrections can be made for certain types of endogeneity, it is possible to make a stronger case for causality. However, any time uptake of the treatment is driven by household choice (such as with microfinance) it is important to exercise caution in making causal inferences. In cases where uptake of the treatment affects an entire village population (such as a road or water system), the measured effects of the RETRAFECT methodology are similar to the standard *intention-to-treat effect*, or ITE and can often be interpreted as causal. Because households *choose* to take microfinance loans at any given time, in this context the question that our methodology is most clearly able to answer is "What happens when households take microfinance loans?" rather than "What is the impact of microfinance?".

We apply this methodology to studying the effects of microfinance among borrowers of microfinance institutions (MFIs) in three countries: Ghana, Guatemala, and India. We analyze changes in the probability of major dwelling improvements, such as upgrades of walls, roofs, floors, the installation of electricity and indoor toilets, as well as the purchase of major consumer durables such as stoves, refrigerators, television, bicycles, and cell phones. Throughout our estimations we use a linear probability estimator on panel data that incorporates village and country-year fixed-effects as well as standard errors clustered at the village level.

We find using this methodology that microfinance borrowing precedes significant subsequent increases in the probability of major housing upgrades. The base probability of any housing upgrade across the three-country sample is 3.8% per year. After a household takes its first microfinance loan, this probability increases to 7.0% in countries where dwelling improvements are a strong indicator of

improved household welfare. The effect is especially strong in the first two years subsequent to a household's first loan, where, the probability of a major housing improvement exactly doubles to an average of 7.6% per year. The relationship holds particularly strongly in Guatemala, where the replacement of adobe walls with concrete block walls and the replacement of dirt floors with concrete or tile floors is viewed as a significant upward step for a rural household. We also find some evidence that microfinance borrowing is positively associated with new purchases of televisions and refrigerators across all three countries, although the positive relationship with consumer durable purchases is less strong than with dwelling improvements. These results show that purchases *follow* loans, not that that they are *caused* by them. However, when we control to the fullest extent possible in our data for demand and supply side endogeneity and attrition bias, we continue to find evidence consistent with causal impacts, although they are restricted to improvements in housing.

The RETRAFECT methodology is quite general, and applicable to a wide variety of contexts. Indeed, we argue that this kind of methodology might be applied in a more straightforward way to studying the impact of treatments such as the introduction of fresh water systems, roads, vaccines, or other health interventions for which uptake is instantaneous and all-inclusive within a community. While it offers advantages relative to alternative approaches, however, researchers must exercise caution in implementation, in the use of statistical tests, and in interpretation of results under varying treatments and treatment conditions.

The next section provides a brief review of the impact study literature, and how different impact methodologies have been applied to an analysis of microfinance. Section 3 considers our field research context, methodology, and econometric model. Section 4 presents our results, and Section 5 concludes with suggestions and caveats about the appropriateness of our approach to other contexts.

2. Impact Methodologies and Microfinance

Historically, researchers have used a number of methodologies to ascertain the impact of microfinance. Each offers advantages and disadvantages in terms of cost, accessibility of data, and unbiasedness of impact estimates. One traditional methodology, for example, has been a before-and-after analysis of a treatment group relative to an ostensibly similar sample of individuals outside the treatment. An analysis of the "difference-in-differences" in this context is used to capture the difference in change among impact variables within the treatment group (see, for example, van de Walle, 1999 or Banerjee and Duflo 2004 for an application to the relaxation of credit constraints in Indian firms). While such studies are relatively straightforward to carry out, they require pre-treatment data for both treated and untreated populations, and are thus often time consuming to practitioners.

Matching and propensity score methods attempt to create artificial controls in order to identify treatment effects. Gómez and Santor (2003), for example, use a statistical matching model to identify the effect of group lending relative to individual lending among 1,389 borrowers in a Canadian lending institution. However, analyses performed without a randomly chosen treatment and control groups can lead to impact bias from self-selection into the treatment based on unobservables, such as entrepreneurial drive or a predilection for self-improvement, or even raw IQ. Microfinance borrowers, in particular, are a self-selected group who are likely to possess characteristics that differ from the population norm. For example, entrepreneurial drive is likely to be much stronger among those seeking microfinance loans than a typical subject of a survey, and even a typical entrepreneur. As a result, problems with omitted variable bias are likely to cause an overestimation of treatment effects from microfinance.

Some have tried to skirt these problems by comparing old members of a treatment group with newer members, such as using newly enrolled or "pipeline" borrowers in a microfinance program as a

control group for old borrowers (Coleman 1999).¹ This had been the approach undertaken in some research on microfinance, including some of the early studies of the AIMS (USAID) research project, as well as Copestake et al (2001). But as Karlan (2001) and Karlan and Alexander-Tedeschi (2006) point out, this kind of approach can suffer from ‘attrition bias’ in which the performance of old borrowers may exceed those of new borrowers because of a hidden qualities in old borrowers that have allowed them to remain in the program. Only a subset of new borrowers is likely to share these qualities, and hence the impacts observed by a researcher will be biased by this unobserved difference.

In other instances researchers have used instrumental variables to try to identify impacts. By using a third variable correlated with program access but with no direct effect on the impact variables of interest, the use of instrumental variables can overcome problems of endogeneity to allow for unbiased estimates. Work of this kind often uses the staggered implementation of programs or other participation rules which can be exploited by researchers to analyze program impact.

Wydick (1999), for example, uses the staggered nature of the introduction of lending in different areas to help identify the degree of credit access granted to Guatemalan borrowers in estimating the effects of microfinance on child labor. In this approach credit effects on school enrollment are obtained using the staggered entry of a credit institution into different areas along with gross sales as instrumental variables for quantity of borrowing.

Pitt and Khandker’s (1998) well-known study examines the impact of microfinance among a population of households who were located in areas served by the three largest microfinance institutions in Bangladesh, the Grameen Bank, RD-12, and BRAC. The authors exploit the program participation rules of the microlenders as an identifying instrument, which limit participation to poor households who owned less than 0.5 acres of land. Identification of impact from their study rests on differences in consumption and

1. Coleman (1999) obtains a measure of microfinance impact in 14 villages in Thailand by using a quasi-experimental methodology in which borrowers who would receive microfinance loans in the future act as a control group for borrowers that were actually granted credit access. By including a dummy variable for credit participation by both those that seek credit in the control villages and those with access to credit in the treatment villages, he controls for self-selection issues. Coleman finds the impact of microfinance to be small, yet cautions that the impact may be diluted in his study based on the relatively high degree of wealth and widespread credit access of the borrowers throughout his sample population.

other variables by borrowers marginally on either side of this participation rule. They find that consumption by households increased when loans were granted to women by about 18% of the amount borrowed.

The main difficulty with the use of instrumental variables is logistical; instruments, if they are available, differ from one situation to the next. Furthermore, finding instruments such as Pitt and Khandker's that are strongly correlated with program access in a particular context, but uncorrelated with impact variables, also requires substantial ingenuity, complicating the use of a standardized instrumental variable approach. In the context of microfinance, finding convincing instrumental variables for credit access or actual borrowing has often proved to be a frustrating exercise for researchers (Armendáriz de Aghion and Morduch, 2005). What is more, instruments vary in their strength of correlation with program access; weak instruments yield imprecise estimates of true impact magnitudes. Exclusion restrictions of potential instruments from the main estimation are not always easily satisfied, and are often violated in practice by field officers in the interest of making credit available to desirable borrowers. Pitt and Khandker's study, for example, has been criticized because the land-ownership participation rule that was used for identification of impacts was often fudged by loan officers, allowing some highly motivated participants access to the program who otherwise should have had been excluded (Morduch 1998).

To overcome these problems, the use of randomized field experiments has become increasingly common in ascertaining the impacts of many types of poverty intervention programs (see Duflo, 2006). As yet there are few examples of randomized field experiments in microfinance, but randomized field experiments have become popular because they allow for a maximum degree of exogeneity in treatment and control, allowing researchers to overcome the often-thorny issues of self-selection, endogeneity, and omitted variable bias. One promising methodology is randomization among the pool of "barely rejected" borrowers in institution which use credit scores; this allows for

estimation of the impact of expansion of credit to a previously unserved group of borrowers (Karlan and Zinman, 2008).

Randomized field experiments, however, face their own set of challenges. To create the control group needed for the identification of treatment impact, it is necessary that some who desire access to the treatment (such as health, or education, or microfinance) remain untreated for a specified time so that impact can be measured on an equivalent treatment group relative to the control. This difference in the timing of treatment is usually justified by a constraint on the institution's ability to treat all agents immediately anyway, and in many cases a random lottery can actually be perceived as a fair way to determine the queuing rule. Yet it is essential that the control remain untreated by both the institution and competing institutions for the specified duration of the study.³

The challenge becomes that the longer desirous households are denied a treatment, such as access to microcredit, there is an increasing tendency for untreated units to “bleed” from the study by seeking the treatment (e.g. loans) from other sources. Any synthetic research structure is difficult to maintain for a long period of time, and short-duration studies present several practical problems for the evaluation of interventions such as microfinance, where impacts may accrue over long periods of time. Short-duration studies can be subject to the influence of time-specific economic shocks that are complementary to a treatment, yielding a downward bias in standard errors. Additionally, because short-duration studies represent a snapshot of program impact over a short time frame, they cannot capture important dynamics of treatment impact. Ideally, both practitioners and researchers would like to understand how *long* a given impact takes to become fully realized within a population and when the impact dies out. Movies contain more information than photographs.

³ “Encouragement designs” have offered one important way around the problem of withholding treatment. In an encouragement design, a randomly selected group of people within the study receive a special promotion to undertake a treatment. If the random encouragement is significantly correlated with treatment uptake, it then can serve as a valid instrument for identification.

3. Methodology

The RETRAFECTION methodology tries to address a number of these issues, while creating its own distinct set of challenges. We first carry out a household survey that creates a historical retrospective panel of fundamental events. We define fundamental events as those that are discrete, unforgettable, and important indicators of household welfare. Our fundamental events include major dwelling changes, discrete business investment, and purchases of large consumer durables. We combine this retrospective panel with historical variation in the timing at which different households in our three country sites had access to microfinance and when they actually took up loans.

The ideal sample for a retrospective impact study would include a random sample of the population as well as a random sample of current program participants and program dropouts (most likely one would want to over-sample dropouts and program participants and then weight them by their share in the population). The dropouts allow the researcher to create a random sample of program participants starting after a specified time, and thereby to control for "attrition bias," in which long-term participants may show treatment effects than the average participant (see Karlan, 2001). In one of our country sites (Guatemala) we obtained a special sample of dropout borrowers to test for attrition bias, finding none, but we were not able to collect random samples of non-participants, and so we are unable to measure a standard Intention to Treat Effect in our data.

The cross-sectional survey used for the RETRAFECTION methodology is built around discerning the timing of memorable events in the history of a household. For example, a study on the impact of a pre-natal health program on miscarriage and infant mortality could accurately collect recall data on miscarriages, births, and deaths of children, which are unforgettable events to any parent, but probably not on minor childhood illnesses. Other interventions that attempt to support the weakest members in a society may want to use elder deaths as a fundamental event to households. A study on the impact of roads and infrastructure might want to study the timing of the formation of small businesses.

Historical questions on changes in most non-discrete events, such as revenues and profits in an informal sector enterprise, are inappropriate in the creation of a retrospective panel since their timing and precise quantities may be difficult or impossible for subjects to recall. Thus, unless carefully recorded records exist, our technique can be used effectively only with discrete and psychologically significant dependent variables. The type of event chosen should be a direct or indirect treatment effect, rare and important enough to be memorable, but probable enough such that it can be meaningfully used as a dependent variable. A balance must be struck thus between memorable events (which are more rare), and events which are frequent enough to permit detection of meaningful program effects.

In our study, our survey teams often ascertained the timing of these fundamental events by referencing them off the ages of children and other key events in the life of the household and village.⁴ Impact studies using variables such as housing are advantages because the qualities of a house are manifest to a surveyor visiting the household, and thus it is necessary only to help the household pinpoint the timing of when, for example, a house was upgraded from adobe to concrete. This contrasts with data on, say, infant mortality, for which the event itself must be elicited by the researcher, along with the timing. From our data we create a history for each household consisting of these fundamental events along with the timing of initial credit access and initial borrowing back to the time of occupation of the dwelling unit. The sum of these recreated histories across households forms an (unbalanced) panel data set from which estimations are carried out.

The RETRAFECTION methodology exploits the differential timing of program implementation, using households in the sample who access the program at different times as mutual controls. The counterfactual in our methodology is identified within the sample itself by the probabilities of home improvements among the whole sample of borrowers before credit was introduced into each household's village. Thus one of the advantages of the RETRAFECTION methodology from a practitioner standpoint is the attraction of being able

⁴ Using this methodology, it is helpful for the surveyor to create a chronological record of births of children in the family as well as deaths of parents and grandparents, years which are likely to be fixed in people's minds. The timing of these events can then be used to pinpoint the timing of an event by asking whether certain children or grandparents were alive when, for example, a new house was built.

to form counterfactuals within an institution's own client base. The fact that all in the sample (at least eventually) chose the credit treatment addresses problems of self-selection, because we do not compare recipients to non-recipients. Because we never observe an ‘average’ individual, however, we can make no claims about the average effect of the treatment, and we are restricted to measuring how the *timing* of treatment effects the *timing* of consumption within a sample that will eventually receive credit.

The idiosyncratic influences of the economy over different years are controlled for through country-year fixed-effects. The idiosyncratic differences between villages are dealt with through village-level fixed effects. We use clustered standard errors at the village level so that our *t*-statistics are robust to cross-village heteroskedasticity.

Care must be taken in inferring causality through the RETRAFFECT methodology when the timing of the receipt of treatment is endogenously chosen by households. Subject to supply-side endogeneity tests outlined below, we can begin by estimating the impact of the *availability* of treatment in a community. Because our sample consists entirely of one-time program participants, this estimates treatment effects that are different from the standard “intention to treat effect” (ITE) and “treatment effect on the treated” (TET). Akin to the ITE, it estimates what we refer to as a *intention to treat effect on the eventually treated*, which is defined as the change in the probability of a fundamental event from a household having access to the treatment in its community, among those households who eventually take up the treatment. This is different from the standard ITE, which compares impact across a random sample of households in a community over a given time period from a treatment being available when uptake is less than 100% within the community. More subject to potential endogeneity issues, but still of real policy interest, is the *take-up effect*, which measures changes in expenditures subsequent to the *receipt* of credit, rather than subject to its *availability*. Because households choose the timing of both events, we cannot cleanly ascribe causality to this relationship, but we may still be interested in understanding this sequencing. We again propose a series of diagnostic checks for demand-side endogeneity which examine whether those who took the treatment immediately behave differently than those who waited to take it until after it was available.

As a result, an important challenge with our approach is that data must meet certain specific criteria for exogeneity in order to ascribe causality in the relationship between treatment and impact. Hence, before we can begin to assess causality from the rollout of the program, we must test for obvious kinds of supply-side and demand-side endogeneity. In short, we attempt not to use words such as "impact" or "causality" lightly, and we outline the process of our econometric methodology in the following sequence of steps:

Step 1: Simple Estimation on Treatment Availability and Treatment Uptake

One can obtain a simple idea if a treatment is associated with any significant change on the treated population by carrying out two regressions. These are seen in Table 2 and consist first in regressing the impact variables on a set of village and country/year fixed effects, as well as a dummy variable equal to one after an individual has *taken* a loan. We then use the same setup to regress our impact variables on a dummy variable indicating the *presence* of the program beginning in the year when it was introduced in each geographical area. Unless there is strong evidence of negative supply-side endogeneity, a lack of significance of the treatment uptake dummy would seem to be indicative of insignificant program impacts. However, positive significance of either of these warrants further investigation into the dynamics of the relationship on the impact variables and the nature of the positive relationship.

Step 2: Treatment Window Regression

Event studies use a treatment window within which changes in the dependent variable are analyzed. There is a trade-off in determining the width of the treatment window that includes the number of pre-treatment and post-treatment years used in the final take-up effect estimation. A larger window reveals longer-term effects, but more recent periods contain fewer observations and are estimated with less precision. A smaller window contains a smaller set of more precise observations, but reveals less about the dynamics of impact. We believe it reasonable that most housing effects would occur within a 3-4 year time period after credit access (and our data appear to show a tapering off after

this point), but other types of effects may suggest a shorter or longer window. The treatment window also yields the estimated parameters necessary for a test of take-up effects via an F -test in which we can test if the sum of the lagging (post-treatment) coefficients is significantly greater than the leading (pre-treatment) coefficients. Because households choose the timing with which they take credit, we include windows for both actual credit take-up and credit availability.

Step 3: Testing for supply-side endogeneity in the rollout of a program.

Positive coefficients on post-treatment coefficients in the event window may be indicative of causal impact, but may also stem from non-random program placement (*i.e.* supply-side endogeneity) or treatment uptake timing which is correlated with impact variables (*i.e.* demand-side endogeneity). The first set of diagnostics examines the pre-treatment outcomes across communities in order to see whether there are signs of supply-side endogeneity. Direct questions to field directors regarding the nature of how programs have been rolled out across regions are important, but the independence of the rollout with respect to impact variables should be confirmed statistically. The first test uses only pre-treatment data, calculates village/year average outcomes, and regresses these on country/year fixed effects and a variable which gives the year in which credit was first introduced into the village. This tests whether early recipient villages had different mean pre-treatment outcomes than late recipients. Given the use of fixed effects in the regression, however, this kind of endogeneity alone would not bias coefficient estimates. The second test checks for whether the pre-treatment *trend* in outcomes varies with the order of receipt of the treatment. If, for example, the program was offered first to those communities that were growing quickly anyway, then fixed-effects regression using retrospective panels will be biased. We perform this test by taking first difference of the village average pre-treatment outcomes and regressing these on country/year fixed effects and the year the treatment was introduced. If we find bias of this kind, we can proceed to run the simple treatment regressions in Step 1 which interacts a time trend with the village fixed effects, thereby allowing each community to have a different trend. If we do not find

such bias, then we can proceed using a standard two-way fixed effects regression. The third test uses only pre-treatment data and runs a two-way fixed effects regression including a dummy for the year prior to the receipt of the treatment. If this term is found to be significant, then it indicates that the treatment may have been systematically introduced in response to some kind of shock, and any mean-reversion in outcomes will likely lead to biased estimates of program effects.

Step 4: Testing for Demand-Side Endogeneity in Credit Take-up by Clients

The presence of demand-side endogeneity is more difficult to reject, but we propose a diagnostic test which includes a “no program” dummy variable that indicates the absence of program access. This dummy variable is interacted with leads of credit uptake to form a test of whether borrowers who were unconstrained prior to taking credit behaved differently than those who were constrained (the test is described in more detail in Section 4). If this interaction term is significant, it means that there are differences in the pre-treatment outcome variables between those with access to the treatment and those without, indicating the existence of an Ashenfelter’s Dip (see Ashenfelter, 1978). The sign of this interaction term will indicate, for example, if people take microloans when they are facing positive shocks (perhaps reflecting added economic opportunity) or negative shocks (which would indicate consumption smoothing).

A second diagnostic we employ to check for demand-side endogeneity is to restrict the estimation sample to those who received credit shortly after the program entered the village. If these estimates are significantly different from the unconstrained sample estimates, then it may indicate demand-side endogeneity because it implies that omitted variables may be causing households to take credit that are correlated with credit take-up and impact variables. If these estimates are statistically similar, then demand-side endogeneity is likely to be less of a concern since those freshly exposed to credit exhibit similar effects to those “choosing” credit later. Still, results must be treated with some caution. In general, with any treatment in which take-up is optional for households, it is possible that

the treatment has been taken in order to specifically achieve the outcome which we measure, and so we cannot interpret a test for differences in outcomes after individuals choose a treatment as necessarily causal. The difference may be subtle: Microfinance and other interventions may very well operate as a door which those must walk through who desire a third effect, such as home improvement.

Step 5: Other Robustness Checks

With data on borrowers who have ‘dropped out’ of the program, we can form a test for attrition bias by interacting a dummy for dropout with the credit availability and credit take-up dummy variables in Step 2. If we find statistical significance in this interacted variable, it indicates that those who continue in the program longer display different effects from the program than those who drop out early. In this case an estimation should be carried out that weights the "dropouts" equally to their proportional representation in the program.

We can also partition our sample by country to ascertain if the effects we observe are consistent across different contexts. This is helpful in ascribing external validity to positive associations between program and impact variables. Some researchers may also want to interact the treatment dummy variable with household characteristics to ascertain among which household types the treatment displays the strongest take-up effects.

3. Field Research and Estimation

It is likely that while most microfinance loans are intended for business investment and not housing, increased profits from microloans should result in housing changes.⁶ In Guatemala as in many other

⁶ It is also possible that microloans intended for enterprise capitalization may be diverted into use for dwelling improvements. An anonymous Bolivian MFI estimates that 20 percent of its “microenterprise” loans go for home construction and expansion (Center for Urban Development Studies, Harvard University Graduate School of Design, 2000). Nevertheless, some research has pointed out that investing in dwellings may not necessarily represent a complete diversion of credit, since such improvements may increase the income-generating potential of home-based activities (Ibid.). As a response to this phenomenon, many MFIs have become interested in developing new lines of micro-credit specifically to finance housing (Ferguson, 2004). In Guatemala for example, Génesis Empresarial, a Guatemala City-based MFI, has a small portfolio of borrowers with home improvement loan products that carry average terms of two years.

contexts, housing differs from other goods in that it not only represents an important consumption good, but also a major store of wealth and a measure of prestige. Tax (1953), for example, observes that social status among rural Mayans in Guatemala is often reflected in the quality and size of homes and land. For this reason, improvements in houses and land are typically among the first changes rural households make when family income begins to increase.

Our field research took place in three very different country sites with three different microfinance institutions. In Guatemala we worked with *Fe y Alegria* (trans. *Faith and Joy*), a Jesuit-operated MFI working in a number of villages surrounding the cities of Quetzaltenango and Mazaltenango that lends equally with no preference to religious affiliation. Our survey of 262 households in Guatemala took place in July and August 2005 and December 2007. In India we collaborated with Growing Opportunity Finance, Ltd., an MFI affiliated with Opportunity International, one of the world's largest microlenders, where 414 borrowers were surveyed in and around the city of Chennai (Madras) in the summer of 2006. In Ghana we worked with SASL, another affiliate of Opportunity International, where our team surveyed 996 borrowers in the north-central area of the country in the summer months of 2007. Borrowers were part of borrowing groups in Ghana and India, but part of an individual lending program in Guatemala. Loan amounts in all three countries were in the US\$200-\$500 range, payable over 6-12 months at MFI interest rates approximately two points per month over the formal commercial rate. Credit was specifically earmarked for enterprise activity in Guatemala and India, but in Ghana there was less insistence that money should be used strictly for enterprise investment. There were varying degrees of alternative credit access across these three country sites. In India and Ghana, Opportunity International was one of the first MFIs to begin operating in the survey region, while in Guatemala more credit alternatives were available, especially in later years.

In all three countries, borrowers were engaged in a wide variety of microenterprise activities including tailoring, furniture, and other light manufacturing, retailing, and livestock. In each country we obtained basic borrower information (age, education, gender, type of enterprise) as well as ascertained

the timing of the fundamental household events we use in our study. The starkest difference between the three survey areas was in education, where the average number of years of formal schooling was 8.2 in Ghana, 5.2 in India, but only 2.6 in Guatemala. The average age was very similar in all three countries: 31.1 in India, 33.1 in Ghana, and 33.6 in Guatemala.

The questionnaire was intended to measure changes in our different categories of dwelling improvement: upgrades to walls, roofs, floors, plumbing, electrification and purchases of major household durables. Each borrower was asked about changes in these variables during the history of the household, and the timing of these changes. For example, we asked households how long they had lived in that specified location. If a household had cement walls, we asked them if a different kind of wall structure existed since they had lived in that location. If prior to the cement walls the house had had adobe walls, we asked what year the upgrade had taken place. We tried to pin down the exact year carefully by referencing the relative ages of children at the time of the change and by referencing changes to important local events. In like manner we constructed a time series of changes in each dwelling category since the time the borrower lived in the given location. In total we obtained surveys from 1,672 households for a total of 21,328 observations in the backcast panel. Summary statistics from the three country survey can be seen in Table 1.

Clearly, a substantive concern with this kind of survey method is the problem of inaccuracy in the creation of retrospective panel data set. Our survey method seeks to mitigate this problem by asking subjects only to recall discrete, major changes in the history of their home. Because, for example, the upgrade of floors from dirt to cement poses such a major augmentation in quality of living standards for a family, there was relatively little problem with the recall of such events and their timing by year.

From the survey we then create an unbalanced panel data set. The unbalanced nature of the panel data arises because the model considers the number of years the head of household or borrower has been living in the present site as the defining number of years used in the time series for each household. Our estimations were carried out on data beginning in 1990 if a household had lived in the same house since

then, but some households had resided in a particular locale only after 1990, so for these households we used data beginning in the year they moved into their current residence.

Estimation Technique

For housing upgrades, we estimate the probability of one of our households upgrading from a low quality material to a high quality material in the structure of the house. For walls this is from either adobe to finished adobe, or adobe (finished or not) or wood to cement. For roofs this is from either palm leaves or corrugated iron to either cement or tile. For floor upgrades, the changes we analyze are from dirt to cement, cement to tile, or dirt directly to tile. With changes in toilet, our upgrade is from an outhouse to indoor plumbing. We also look at the year when a house was electrified. We also examine the timing of purchases of stoves, refrigerators, televisions, cell phones, and bicycles as well as discrete investments in machinery and physical business structure, often the initial intended use of the loan.

Probit and logit models are commonly used for estimations in cross-sectional qualitative estimations and sometimes in panel data, but we favor the linear probability model, which has become increasingly used in panel data estimations, since as a linear estimator it produces more robust estimates when implemented in fixed-effects estimations (Chamberlain, 1980). Estimations are conditional, of course, upon a household not previously having made the particular type of dwelling upgrade.

The general fixed-effects model we estimate is the following:

$$y_{ijct} = \nu_j + \alpha_{ct} + \beta X_{ijct} + \sum_{t-\bar{t}=-k}^k \tau_{i,t-\bar{t}} T_{i,t-\bar{t}} + u_{ijct}, \quad (1)$$

where y_{ijct} is a bivariate dependent variable that is equal to 1 if household i in village j and country c realizes a given type of event in year t , and zero otherwise. For the independent variables, ν_j is a village-level fixed effect, α_{ct} is a country-year fixed effect, the X_{ijct} represents a vector of household controls such as education, age, age squared, enterprise type, and initial wealth; u_{ijct} is a mean zero error term.⁷ The fourth

⁷ We omit the control variables in our basic estimations presented here; our data contain few time-varying variables and those which we observe, such as age, were found to be insignificant. Our results are not sensitive to the use of control variables.

term is the estimation on a sequence of treatment dummy variables, $T_{i,t-\bar{t}}$, that comprises a “treatment window” of length w years representing a sequence of lags and leads surrounding year \bar{t} for household i . The treatment dummy variable is equal to 1 if household i first received a microfinance loan $t - \bar{t}$ periods “ago,” and zero otherwise. If $t - \bar{t}$ is negative, it means that household i received credit $t - \bar{t}$ years *forward* from time t . For a symmetric treatment window of width w around the time of treatment, then the summation in the fourth term of the model includes $k = (w - 1)/2$ years of leading treatment dummies, $k = (w - 1)/2$ years of lagged treatment dummies, as well as the contemporaneous dummy for when $t = \bar{t}$, for the year in which the household first received microfinance. For example, consider a treatment window of $w = 5$ for a household i that initially received microfinance in 2001. For the observation of household i in the year 2000, the data in the retrospective panel then contains a vector of treatment dummy variables--0, 1, 0, 0, 0--which correspond to estimated coefficients $\tau_{i,-2}, \tau_{i,-1}, \tau_{i,0}, \tau_{i,+1}, \tau_{i,+2}$. For the observation of household i in the year 2003, the vector of dummy variables would be 0, 0, 0, 0, 1. In our estimations we use a treatment window of $w = 9$ and $k = 4$.

Many upgrades to homes and purchases of consumer durables took place during the surveyed history of our households, but each event is still relatively rare. Table 1 gives base probabilities of our fundamental events. Our dataset only contains records on specific home improvements in Ghana and Guatemala; upgrades to dwellings are not nearly as common in the established urban housing stock in our site in Chennai, India. For dwelling upgrades, the probabilities of upgrades in any given year for households in Ghana and Guatemala, respectively, are 0.004 and 0.025 (new walls), 0.005 and 0.010 (new roof), 0.010 and 0.025 (new floor), 0.010 and 0.029 (electrification), and 0.005 and 0.021 (new toilet), to 0.036 and 0.092 (for any of these dwelling improvements). Thus dwelling upgrades were much more common in Guatemala than in Ghana, but these base probabilities are important in understanding the magnitude of changes in these probabilities after microfinance borrowing. Average probabilities of

purchases of consumer goods by any household in a year for India, Ghana, and Guatemala were 0.022 for stoves, 0.017 for refrigerators, 0.043 for televisions, 0.017 for bicycles, and 0.066 for cell phones.

4. Estimation Results

We will work through a summary of our results based on the steps that outline our methodology. Table 2 shows results from the simple before-and-after regressions using dummy variables for household credit uptake and for credit availability in a household's village. First, we find uptake of credit to be associated with a 0.032 increase in the probability of any housing improvement (over a base probability of 0.038). Thus with microfinance loans increase their annual probability of a major dwelling upgrade from just under four percent per year to about seven percent--a substantial difference, significant at the 99% level. Table 2 shows the increase in probability for televisions to be 0.025 (over a base probability of 0.043, 99% significance), refrigerators to be 0.016 (over a base of 0.017, 99% significance), and bicycles 0.010 (over a base of 0.017, 90% significance). Other consumer goods display a positive sign but without statistical significance. We would expect point estimates to be lower for credit availability in the village (because not all borrowers take it immediately), and indeed this is the case. The increased probability of a housing improvement falls to 0.012 ($t = 1.52$), refrigerators to 0.009 ($t = 1.16$), with the increased probability for other goods close to zero.

We first carry out treatment window estimations that show changes in the probability of discrete physical capital investments within our sample of borrowers within the treatment window. We consider both investments in productive capital equipment such as machinery and structural improvements to business facilities. These outcomes are discrete and memorable, and hence while they do not present the nuanced picture that we would get from, say, profits, they are credible outcomes for a retrospective methodology. Table 3 shows estimations around the treatment window of credit uptake as well as when credit was introduced into a particular household's village. As we would expect, the probability of investment in both productive machinery and structural assets increases significantly by 0.128 in the year

a first microfinance loan is taken over a base probability of 0.071, falling subsequently in the first and second years after credit to probability increases of 0.087 and 0.078, still more than double the base investment probability. Figure 1 plots these changes over time and reinforces the sharp increase in business investment in the year a microfinance loan is taken. Coefficients measuring changes in the probabilities of the different types of investments after credit is introduced into a village are often positive and significant, but smaller.

We now consider changes in the probability of housing improvements. Our treatment window estimations around household credit *uptake* for housing improvements in Table 4A show no significant changes in the years before credit but large increases in the probability for housing improvements beginning in the year credit was received and continuing out to approximately three years after the initial microfinance loan. Figure 2 carries out the interesting exercise of plotting the relative timing of loan receipt, business investments, and housing investments on the same figure (the data only permit us to carry out this analysis in Ghana). We see clearly the immediate jump in business investment and the subsequent increase in housing investment which seems to lag the business investment by roughly one year. Coefficients from the treatment window using both Guatemala and Ghana are plotted with a 95% confidence interval in Figure 3. These probability increases for any housing improvement range between 0.015 and 0.48, peaking in the first year after credit is received and then dying down afterwards. The increases shows that households taking credit roughly double the probability of a housing upgrade in the first three years after the initial microfinance loan is taken from a base probability of 0.039 to an average probability of 0.079 during these years. Increased probabilities for upgrades to walls and floors are especially high where F -tests on significant differences between the four post-treatment probabilities and the four pre-treatment probabilities within the nine-year treatment window show F -tests significant at the 95% level. This makes sense especially in Guatemala, where upgrades from adobe walls and dirt floors to concrete walls and floors are viewed as a sign of significant upward mobility in rural areas.

Table 4B uses a treatment window around the provision of *credit* to a given household's village. In these estimations, the same coefficients are smaller on any housing improvements, but extending farther into the future past the time credit was introduced in the village, and still significant at the 95% level. The corresponding plot of treatment window coefficients with a 95% confidence interval is given in Figure 4. Interestingly, the coefficient on any household improvement remains unchanged at 0.028 relative to the credit uptake estimation window in Table 4A. *F*-tests on significant differences between the post-treatment probabilities and the pre-treatment probabilities show significance at the 95% level for any housing improvement, 90% for wall upgrades, and 85% for roofs.

We also carried out treatment window regressions using the purchase of consumer durables, an outcome which was available in all three of our study countries. The plots of the leads and lags for five kinds of consumer durables (television, refrigerator, stove, bicycle, and cell phone) are given in Figure 5. The probability of purchase of several consumer durables increases in the year a household takes its first microfinance loan, but in general the individual year lags are not significant. The probability of a television purchase increases by about half, by 0.021 from a base probability of 0.043, that of a refrigerator nearly doubles, increasing by 0.014 from a base of 0.017, while that of a stove goes up by about two-thirds, by 0.017 from a base of 0.022. (Interestingly, coefficients vary little between the three countries, although credit is constrained for business investment purposes in both the Guatemalan and Indian MFIs.) For televisions, refrigerators, and bikes, *F*-tests on significant differences between the post-treatment probabilities and the pre-treatment probabilities each show significance at the 95% level. Not surprisingly, cell phone purchases are relatively unresponsive to credit access, as they are likely driven mainly by the installation of cell towers and other infrastructure. Figure 6 plots the leads and lags around the availability of credit in a village, and we see that unlike housing improvements, the purchase of consumer goods does not appear to be associated with the presence of credit in the village, but with actual borrowing.

Our diagnostic checks for supply-side endogeneity (Step 3) are presented in Table 5. Regressing average pre-treatment outcomes (new walls, new roof, new floor, new toilet, new land) on the year in which credit was offered to a village, we find no evidence of supply-side endogeneity in the levels of the pre-treatment outcome based on the year that the credit program was introduced into villages. We also find no evidence of supply-side endogeneity in a pre-treatment trend of more rapid or less rapid changes in home improvements (our second estimation in Table 5). Our last diagnostic is a check for the entrance of the program as a result of previous-period shocks. Again, we find no evidence that the provision of credit to a village is a function of an abnormally greater or lower rate of home improvement in a village the year before credit program entry. Thus we uncover no evidence of endogenous program placement across our three country sample.

Our results for Step 4, the test for demand-side endogeneity, are given in Table 6. First, we create a dummy variable indicating whether or not the credit program had been introduced into the village of household i at time t . We interact this dummy variable with the variables representing the years within the treatment window for each household prior to treatment and include it in the estimation along with the other pre-treatment time dummies. We then estimate the equation

$$y_{ijct} = \nu_j + \alpha_{ct} + \sum_{\bar{t}=t-k}^k \tau_{i,t-\bar{t}} T_{i,t-\bar{t}} + \sum_{\bar{t}=t-k}^{-1} \delta_{i,t-\bar{t}} \tau_{i,t-\bar{t}} T_{i,t-\bar{t}} + u_{ijct}, \quad (2)$$

where there are $n = 2k + 1$ leading and lagged treatment dummy variables in the first summation and k leading treatment dummy variables in the second summation interacted with a dummy equal to 1 if the microfinance program was *unavailable* and zero otherwise, where $d_{i,t-\bar{t}}$ is the interacted dummy representing the absence of a credit program. For parsimony we use $n=5$ and $k=2$, although results are similar if we use longer treatment windows..

Significance of the $\delta_{i,t-\bar{t}}$ coefficients in (2) could reflect demand-side endogeneity in divergent ways. On one hand it is conceivable that microenterprise entrepreneurs might choose to borrow in good economic times, in order to take advantage of economic opportunity. Good economic times could thus

initiate borrowing, but also cause high profits by themselves and thus cause dwelling upgrades. Failing to correct for lack of program access would thus *overestimate* the difference between post-credit and pre-credit treatment variables, *i.e.* demand-side endogeneity would bias the difference between post-credit outcomes and pre-credit outcomes *upwards*. This would bias our F -test on impact toward a propensity for Type I errors (rejecting a null hypothesis that there is no significant change in probability of dwelling upgrades yielded after credit). In the presence of endogenous borrowing based on positive economic opportunity, we would thus expect the $\delta_{i,t-\bar{t}}$'s to be positive. The true change in probability of dwelling upgrades for a pre-credit year would not be $\tau_{i,t-\bar{t}}$, but rather $\tau_{i,t-\bar{t}} + \delta_{i,t-\bar{t}}$.

The opposite relationship is also possible: microentrepreneurs might systematically choose to borrow when prices are low for their particular product (or economic times are hard) in order to smooth negative shocks. Here, failing to correct for lack of pre-credit program access would *underestimate* the difference between post-credit and pre-credit treatment variables, making our F -test biased *downwards* and inclined toward Type II errors, accepting the null of no significant change in probability of dwelling upgrades yielded after credit. With this type of endogeneity, we would expect the $\delta_{i,t-\bar{t}}$'s to be *negative* since with the unavailability of credit, negative shocks would further reduce the probability of dwelling upgrades.

To test for systematic demand-side endogeneity in borrowing decisions, we carry out a test for the significance of the $\delta_{i,t-\bar{t}}$'s in Table 6 in a five-year treatment window on whether they are jointly different than zero. If these interactive dummies are jointly *significant* by an F -test, meaning that significant demand-side endogeneity exists, then our new test for the effect of microfinance borrowing on dwelling upgrades would then become the significance of differences between the post-credit treatment $\tau_{i,t-\bar{t}}$'s and the sum of the pre-credit $\tau_{i,t-\bar{t}} + \delta_{i,t-\bar{t}}$'s within the symmetric treatment window. If the interacted variables are jointly *insignificant*, then the decisions of previously credit-constrained households and previously non-credit constrained households are insignificantly different, and we can use our standard F -

test differences between the sum of the post-credit $\tau_{i,t-\bar{t}}$'s and the sum of the pre-credit $\tau_{i,t-\bar{t}}$'s within the symmetric treatment window.

As seen in Table 6, we find little statistical evidence for the joint significance of these interacted dummy variables. With neither housing improvements nor the purchase of consumer durables are the interacted variables on the raw pre-treatment dummy variables and the microfinance constraint dummy significantly different from zero at the 5% level, and the post-credit coefficients retain similar magnitudes as in Tables 3A and 4A. An additional check for demand-side endogeneity is in Table 2, where we carry out the simple "before-and-after" estimations only on borrowers who received credit in the first year it was offered in their village. Here the timing of their borrowing is less likely to be influenced by demand-side conditions and more by the new availability of credit (for which we find evidence of exogeneity with respect to impact variables). Again we find our coefficient on the increased probability of housing improvements remaining robust at 0.030, though coefficients on some consumer durables decline.

Additional analyses not presented here (output available on the authors' websites) examines the robustness of the findings on housing improvements when we partition the sample by country. Similar treatment effects are found in both Guatemala and Ghana despite the fact that the base probability of housing improvement in Guatemala (.092) is three times as high as that in Ghana. We also tried interacting borrower-level characteristics with treatment dummies to see whether there is heterogeneity in treatment effects. These interaction terms are in general not significant, implying that the effects estimated here are relatively consistent across types of borrowers

The final empirical issue which we need to address is that our impact estimations could be affected by an attrition bias based on the composition of an institution's borrowing portfolio. The phenomenon here could be that old borrowers could represent a group that exhibits different responses to credit than newer borrowers, since some borrowers (for whom the impact of loans could be greater or smaller) may have dropped out of the pool from an old cohort. Our methodology avoids the straightforward attrition

bias described in Karlan (2001) because our identification relies on the specific *timing* of dwelling changes after microfinance borrowing rather than the simple differences in impact variables between old and new borrowers. However, any differences in the timing of purchases relative to credit for dropouts versus ongoing borrowers would still result in inaccurate estimates for the client base as a whole if we were to base our analysis only on current borrowers.

Table 7 presents results from our Guatemalan data on home improvements, indicating that dropout borrowers display higher overall rates of certain housing upgrades, and in addition dropouts are somewhat (though insignificantly) more likely to upgrade walls, roofs, and floors *after* receiving credit than ongoing borrowers. Thus, to ignore dropout borrowers in our Guatemalan sample could result in an underestimate of the increased probability of housing upgrades associated with credit.

Table 8 contains perhaps our strongest case for causality within our data set, using a nine-period treatment window around credit provision to a village in Guatemala, with dropouts weighted by their share in the cohort with whom they first took loans, based on an annual dropout probability of 33%. In this estimation, we continue to see some evidence of an increased probability of housing improvements after credit provision to a village with an average post-credit-provision coefficient of 0.0348 in the four years after an MFI begins operation in a household's village with coefficients statistically significant in the third and fourth years after the initiation of MFI lending.

5. Summary and Implications for Future Research

The methodology we present in this paper involves the creation of a retrospective panel database taken from a surveys conducted at a single point in time. This retrospective panel re-creates a history of fundamental events in the household over time, the timing of these events then being analyzed with respect to the timing of a treatment. By carrying out a regression of these fundamental events within a treatment window, it becomes possible to analyze the subsequent changes in the probabilities of important variables correlated with a development project or program such as microfinance.

We think it may be helpful at this point to note where the RETRAFECTION methodology is most likely to be useful, and highlight important caveats with its implementation. First, to be able to fully attribute relative changes in post-treatment coefficients to program impact, a researcher must identify a program which has been phased in over time in a manner that is unrelated to counterfactual changes in impact variables. This requires that the program rollout not be directly sequenced based on pre-existing changes in the outcome of interest, and that there be no obvious shocks which drive both program placement and shifts in outcomes. Even to estimate impacts on program participants only, a survey should include a random sample of current and former program participants who received the treatment after a given time in the past. If dropouts are under-sampled, observations can then be weighted by the probability of dropout as per a standard attrition correction.

When using retrospective survey questions, use of this methodology should focus on correctly ascertaining the timing of fundamental events. Changes in variables such as profit, revenue, and so forth are difficult for subjects to remember, and are often imprecise by their very nature in informal-sector enterprises. Major diseases, deaths, school enrollments, and major asset purchases are the kinds of variables best used within this framework. In many respects, this may not represent a disadvantage since what researchers (and households in development countries as well) often view as “development” may be closely associated with these kinds of fundamental changes.

Causal inference in a RETRAFECTION study is most clearly obtained when treatment adoption is ubiquitous and instantaneous with access. If the nature of a treatment is that its adoption is very high at the moment it is introduced, the effect measured at the time when the program became available offers the cleanest ex-post measure of causal impact. An example might be a randomly assigned (but mandatory) vaccination program in public schools, or the phase-in of clean water systems to a number of villages over time, where everyone prefers the clean water to what existed before. In these examples theory strongly dictates a clear impact, treatment is instantaneous with access, and if there is exogenous rollout, causality can be more confidently inferred. In this sense, our own application to microfinance

faces significant challenges using the RETRAFECTION methodology because take-up of the treatment is a choice made by households which may be influenced by factors that also affect welfare variables.

It is worth considering carefully what can be inferred from the estimate of the “take-up effect,” defined by events that occur after households *choose* to take credit. Showing that households take credit prior to upgrading houses is not equivalent to showing that credit causes these upgrades, because we cannot verify that the upgrades would not have occurred in the absence of credit. For example, it is conceivable that households could have sought loans from family members instead of from MFIs, and so the counterfactual pattern of home improvement could have been similar had the MFI not located in the village. However, there is still important information present in the relative timing of these events. These estimates provide reliable evidence over what kinds of consumption changes lenders are likely to see in their client base subsequent to the taking of loans. Changes observed subsequent to the *availability* of credit, subject to endogeneity tests outlined here, can be interpreted as more likely to be causal effects on the sample of individuals who take loans, for which we still find large and significant coefficients.

The principal advantages of the RETRAFECTION methodology are that it can be pursued using entirely *ex-post* data from a single survey wave, utilizing the sequencing of a program’s rollout as a natural experiment. It can also be implemented within an institution’s own client base, and can explicitly trace out the dynamics of changes in welfare variables. We outline the circumstances under which statistical significance can be attributed to standard causality, and suggest a sequence of diagnostic checks and possible remedies when different types of endogeneity are present in the data.

We find the strongest relationship between credit and housing improvements when we use the most endogenous measure; namely the probability of improvement after *choosing* to take credit. The probability of housing improvements exactly doubles during the two years subsequent to a household’s first microfinance loan from 0.038 to 0.076; the probability of a business investment nearly triples from 0.081 to 0.209 in the year a household takes credit. Evidence of increased propensity to purchase consumer goods after the receipt of credit is weaker, but coefficients are virtually always positive and in

some cases significant. Exploiting a survey of dropouts performed in Guatemala to calculate effects which are representative of borrowers as a whole, we still find an elevated probability to improve housing subsequent to the availability of credit. Future research should compare these estimates to those obtained in experimental field research where it may be more difficult to study the dynamics of program effects, but where one is more able to make firmer inferences over causal impacts.

Bibliography

- AIMS Team. Clients in Context: The Impacts of Microfinance in Three Countries. AIMS Paper. Management Systems International
- Armendáriz de Aghion, Beatriz and Jonathon Morduch (2005) *The Economics of Microfinance*. Cambridge: MIT Press.
- Ashenfelter, Orley (1978) "Estimating the Effect of Training Programs on Earnings" *The Review of Economics and Statistics*, Vol. 60, No. 1., pp. 47-57.
- Banerjee, Abhijit., and Esther Duflo (2004) "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program." MIT Department of Economics Working Paper.
- Brown, Warren (2003) "Building the Homes of the Poor" ACCION InSight Series No. 4.
- Center for Urban Development Studies (2000) "Housing Microfinance Alternatives, Synthesis and Regional Summary: Asia, Latin America and Sub-Saharan Africa." Harvard University Graduate School of Design.
- Chamberlain, Gary (1980) "Analysis of Covariance with Qualitative Data" *Review of Economic Studies*, Vol. 47, pp.225-238.
- Copstake, J.S., S. Bhalotra and S. Johnson (2001) "Assessing the Impact of Microcredit: A Zambian Case Study", *Journal of Development Studies*, 37(4), 81-100.
- De Soto, Hernando (1989) *The Other Path: The Invisible Revolution in the Third World*. New York: Harper and Row.
- Duflo, Esther (2006) "Field Experiments in Development Economics" Massachusetts Institute of Technology Working Paper.
- Duflo, Esther, and Michael. Kremer (2003) "The Use of Randomization in the Evaluation of Development Effectiveness", World Bank Operations Evaluation Department Working Paper.
- Easterly, William (2006) *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done So Much Ill and So Little Good*. New York: Penguin Press.
- Ferguson, Bruce (2004) *Housing Microfinance: A Guide to Practice*. Kumarian Press.
- Ferguson, Bruce (1999) Micro-finance of housing: a key to housing the low or moderate-income majority? *Environment and Urbanization*, Vol. 11, No. 1.

- Halder, Shantana and A.M.M. Husain (1988) "Identification of the Poorest and the Impact of Credit on Them: The case of BRAC", mimeo, BRAC Research and Evaluation Division, Dhaka.
- Karlan, Dean (2001) "Microfinance Impact Assessments: The Perils of Using New Members as a Control Group" *Journal of Microfinance* (December).
- Karlan, Dean and Jonathan Zinman (2008). "Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts." Yale University Working Paper.
- Khandker Shahidur (1988) *Fighting Poverty with Microcredit*. New York: Oxford University Press.
- MacKinlay, Craig (1997) "Event Studies in Economics and Finance" *Journal of Economic Literature*, Vol. 35, pp.13-39.
- Morduch, Jonathan (1998) "Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh", Working Paper.
- Pitt, Mark and Shahidur Khandker (1998) "The Impact of Group-Based Credit Programs on Poor Households in Bangladesh: Does Gender of Participants Matter?" *Journal of Political Economy*, Vol.106, No.5.
- Savedoff, William, Ruth Levine, and Nancy Birdsall (2006) "Will We Ever Learn? Improving Lives Through Impact Evaluation". Report of the Evaluation Gap Working Group, Center for Global Development.
- Shumann, Richard (2004) "Developing Housing Microfinance Products in Central America" ACCION InSight Paper No. 12.
- Sims, Christopher (1972) "Money, Income, and Causality" *American Economic Review*, Vol. 62, 1972, pp.540-52.
- Tax, Sol (1953) *Penny Capitalism A Guatemalan Indian Economy*. Smithsonian Institution, Institute of Social Anthropology. Publication No. 16.
- Van de Walle, Dominique (1999) "Assessing the Poverty Impact of Rural Road Projects" World Bank mimeo. Washington, D.C.
- World Bank (2002) "Microfinance for Housing: The Mexican Case" Report prepared for the World Bank's Latin America and Caribbean Region Finance and Infrastructure Department.
- Zaman, Hassan (2000) *Assessing the Poverty and Vulnerability Impact of Micro-Credit in Bangladesh: A Case study of BRAC*. Washington DC: World Bank Publications.

Table 1: Summary Statistics:
Table of means; standard deviations in parentheses.

Control Variables:

	1st Year of Credit	Dummy for Male	Age	Years of Education	Dummy for Retail Sector	Dummy for Manufacturing
Ghana:	2005.2 (2.15)	0.105 (0.31)	33.1 (9.66)	8.198 (4.98)	0.908 (0.29)	0.017 (0.13)
Guatemala	2002.0 (2.09)	0.000 .	33.6 (8.83)	2.648 (3.29)	0.741 (0.44)	0.418 (0.49)
India	2003.8 (1.56)	0.000 .	31.1 (9.02)	5.174 (3.81)	0.471 (0.50)	0.022 (0.15)
Whole Sample	2004.4 (2.27)	0.069 (0.25)	32.7 (9.48)	6.944 (4.95)	0.786 (0.41)	0.056 (0.23)

Housing Upgrades:

	Any home improvement	Walls improvement	Roof Improvement	Floor Improvement	Electricity Upgrade	Toilet Upgrade
Ghana:	0.036 (0.19)	0.004 (0.06)	0.005 (0.07)	0.010 (0.10)	0.010 (0.10)	0.005 (0.07)
Guatemala	0.092 (0.29)	0.025 (0.15)	0.010 (0.10)	0.025 (0.16)	0.029 (0.17)	0.021 (0.14)
India	0.020 (0.14)
Whole Sample	0.038 (0.19)	0.007 (0.08)	0.005 (0.07)	0.012 (0.11)	0.013 (0.11)	0.007 (0.08)

Consumer Goods:

Businesses:

	TV	Stove	Fridge	Bicycle	Cell Phone	Any Major Investment
Ghana:	0.034 (0.18)	0.026 (0.16)	0.008 (0.09)	0.006 (0.08)	0.086 (0.28)	0.097 (0.09)
Guatemala	0.074 (0.26)	0.036 (0.19)	0.042 (0.20)	0.054 (0.23)	0.102 (0.30)	. .
India	0.054 (0.23)	0.005 (0.07)	0.030 (0.17)	0.030 (0.17)	0.000 0.00	0.038 (0.04)
Whole Sample	0.043 (0.20)	0.022 (0.15)	0.017 (0.13)	0.017 (0.13)	0.066 (0.25)	0.081 (0.07)

Table of means. Variances in parentheses.

**Table 2: Discrete Before-and-After Dummy Variables:
(Whole Three-Country Sample)**

Dependent variable:

Explanatory variable:	Housing Improvement	Television	Refrigerator	Cell Phone	Stove	Bicycle
Household Credit Uptake:	0.032 (3.45)**	0.025 (3.04)**	0.016 (2.69)**	0.008 (0.79)	0.007 (1.33)	0.010 (1.77)
Constant	0.076 (3.91)**	0.076 (3.91)**	0.018 (1.89)	0.01 (2.72)**	0.001 (0.62)	0.012 (1.26)
Observations	21,309	21,309	21,309	21,309	21,309	21,309
R-squared	0.02	0.03	0.02	0.09	0.02	0.03

Whole Sample:

Dependent variable:

Explanatory variable:	Housing Improvement	Television	Refrigerator	Cell Phone	Stove	Bicycle
Credit Available in Village:	0.012 (1.52)	-0.001 (0.12)	0.009 (1.16)	0.005 (0.74)	-0.002 (0.53)	-0.005 (1.13)
Constant	0.078 (4.20)**	0.002 -0.14	0.007 (2.24)*	0.003 -1.23	-0.015 -1.07	-0.007 -0.92
Observations	21,309	21,328	21,328	21,328	21,328	21,328
R-squared	0.02	0.03	0.02	0.09	0.02	0.03

Sample that takes Credit Immediately:

Dependent variable:

Explanatory variable:	Housing improvement	Television	Refrigerator	Cell Phone	Stove	Bicycle
Credit Available in Village	0.030 (1.84)	-0.005 (0.30)	0.021 (1.33)	0.016 (0.93)	-0.008 (0.67)	0.001 (0.04)
Constant	0.015 (1.45)	0.026 (1.98)	0.027 (2.00)	0.010 (1.08)	0.003 (0.63)	0.000 (0.01)
Observations	4,070	4,070	4,070	4,070	4,070	4,070
R-squared	0.03	0.05	0.04	0.12	0.03	0.04

All regressions use country/year dummies and village-level fixed effects.

SEs clustered at the village level.

Robust *t*-statistics in parentheses

* significant at 5%; ** significant at 1%

Table 3: Business Investment, Using Leads and Lags of the Uptake of Credit by Household and Credit Introduction into Village (Ghana and India)

Explanatory variable:	---- (Credit Uptake by Household)----			----(Credit Introduced into Village)----		
	Busn. Invst	BusStruc Inv.	Busn Machine	Busn. Invst	BusStruc Inv	Busn Mach
4 years before	0.006 (0.82)	0.008 (1.92*)	0.001 (0.14)	0.005 (0.80)	0.000 (0.00)	0.008 (1.27)
3 years before	0.008 (0.79)	0.008 (1.73*)	0.003 (0.40)	0.003 (0.36)	0.003 (0.62)	0.002 (0.27)
2 years before	-0.003 (-0.31)	0.014 (1.97*)	-0.016 (-1.80)	0.008 (0.73)	0.006 (1.14)	0.003 (0.38)
1 years before	0.015 (1.50)	0.012 (1.43)	0.006 (0.64)	-0.003 (-0.32)	-0.001 (-0.10)	-0.004 (-0.46)
Year of receipt	0.128 (6.85**)	0.113 (6.72**)	0.044 (3.77**)	0.018 (1.10)	0.013 (1.16)	0.005 (0.46)
1 year after	0.087 (5.10**)	0.086 (4.80**)	0.017 (1.05)	0.030 (1.48)	0.024 (1.78)	0.000 (0.03)
2 years after	0.078 (4.77**)	0.063 (5.04**)	0.029 (1.78)	0.034 (1.92)	0.028 (2.15*)	0.022 (1.55)
3 years after	0.038 (1.83*)	0.026 (1.45)	0.014 (0.84)	0.027 (1.25)	0.011 (1.12)	0.019 (1.14)
4 years after	0.040 (1.13)	0.045 (1.94*)	0.024 (0.54)	0.028 (1.96*)	0.007 (0.71)	0.025 (1.95*)
Constant	0.008 (2.86**)	0.006 (2.35**)	0.006 (2.35**)	0.007 (2.51**)	0.213 (11.02**)	0.23 (11.82**)
R-squared	0.166	0.120	0.096	0.157	0.103	0.095
Observations	18,984	18,960	18,960	18,984	18,960	18,960

All regressions use country/year dummies and village-level fixed effects.

SEs clustered at the village level.

Robust *t*-statistics in parentheses

* significant at 5%; ** significant at 1%

Table 4A: Housing Improvements, Using Leads and Lags of the Uptake of Credit by Household:
(Ghana and Guatemala Only)

Explanatory variable:	Any Housing Improvement	New Walls	New Roof	New Floor	Electricity	New Toilet
4 years before receipt	-0.008 (0.90)	-0.003 (1.12)	-0.001 (0.28)	0.001 (0.31)	-0.003 (0.56)	0.003 (0.73)
3 years before receipt	0.000 (0.02)	0.000 (0.07)	0.002 (0.97)	0.004 (0.86)	-0.008 (1.50)	-0.005 (1.47)
2 years before receipt	-0.008 (0.70)	-0.003 (0.63)	0.003 (0.86)	-0.003 (0.63)	0.001 (0.25)	-0.002 (0.46)
1 years before receipt	0.004 (0.38)	0.006 (1.40)	-0.001 (0.47)	0.002 (0.46)	0.008 (0.99)	-0.005 (1.19)
Year of receipt	0.044 (3.70)**	0.012 (2.25)*	0.006 (1.20)	0.017 (2.49)*	0.016 (1.63)	0.012 (1.86)
1 year after receipt	0.048 (2.85)**	0.02 (2.63)*	0.011 (1.85)	0.017 (1.75)	0.017 (2.00)	0.002 (0.34)
2 years after receipt	0.028 (1.78)	0.018 (1.81)	0.003 (0.46)	0.028 (2.43)*	0.008 (1.07)	-0.002 (0.35)
3 years after receipt	0.015 (0.94)	0.008 (0.93)	0.000 (0.06)	0.039 (2.39)*	-0.004 (0.70)	0.001 (0.13)
4 years after receipt	-0.002 (0.07)	0.033 (2.17)*	0.004 (0.62)	0.008 (0.65)	0.008 (0.46)	-0.01 (3.38)**
Constant	0.078 (4.06)**	0.014 (1.65)	0.011 (0.97)	0.049 (2.50)*	0.017 (0.96)	-0.026 (1.66)

Observations 15,928 15,459 15,640 15,123 14,973 15,349

R-squared 0.01 0.03 0.01 0.01 0.01 0.01

All regressions use country/year dummies and village-level fixed effects.

SEs clustered at the village level.

Robust *t*-statistics in parentheses

* significant at 5%; ** significant at 1%

Table 4B: Housing Improvements, Using Leads and Lags of Provision of Credit to the Village:
(Ghana and Guatemala Only)

Explanatory variable:	Any Housing Improvement	New Walls	New Roof	New Floor	Electricity	New Toilet
4 years before access	0.006 (0.99)	0.004 (1.31)	0.003 (1.30)	0.004 (1.13)	0.002 (0.43)	0.001 (0.30)
3 years before access	0.010 (1.46)	0.001 (0.32)	0.000 (0.06)	0.010 (2.38)*	-0.002 (0.68)	0.001 (0.39)
2 years before access	0.002 (0.32)	-0.003 (1.58)	-0.002 (1.28)	0.001 (0.30)	-0.004 (1.03)	0.000 (0.03)
1 years before access	-0.002 (0.25)	0.001 (0.45)	-0.001 (0.50)	0.003 (0.80)	-0.004 (1.44)	-0.001 (0.56)
Year of receipt	0.013 (1.08)	0.002 (0.76)	0.006 (1.52)	0.006 (1.25)	0.012 (1.41)	0.000 (0.17)
1 year after access	0.020 (2.19)*	0.008 (2.02)*	0.012 (2.49)*	0.001 (0.19)	-0.004 (0.99)	0.005 (1.17)
2 years after access	0.028 (2.24)*	0.012 (2.21)*	0.008 (2.08)*	0.015 (1.77)	0.01 (1.86)	0.001 (0.31)
3 years after access	0.020 (2.58)*	-0.003 (1.17)	-0.005 (2.20)*	0.007 (1.21)	0.002 (0.36)	0.008 (1.82)
4 years after access	0.011 (0.94)	0.004 (0.84)	-0.002 (0.86)	0.011 (1.42)	0.001 (0.16)	-0.002 (0.56)
Constant	0.006 (0.99)	0.004 (1.31)	0.003 (1.30)	0.004 (1.13)	0.002 (0.43)	0.001 (0.30)
Observations	15,592	15,139	15,327	14,871	14,707	15,052
R-squared	0.02	0.01	0.01	0.01	0.01	0.01

All regressions use country/year dummies and village-level fixed effects.
SEs clustered at the village level.

Robust *t*-statistics in parentheses

* significant at 5%; ** significant at 1%

Table 5: Tests for Supply-Side Endogeneity:

(Ghana, India, and Guatemala)

A. Is the pre-treatment *level* of outcomes a function of the order of rollout?

Dependent variable is year/village average LEVELS of:

Explanatory variable:	Home Improvement	Television	Refrigerator	Stove	Bicycle	Cell Phone
Year in which credit received	-0.0002 (0.18)	0.0020 (1.42)	0.0007 (0.61)	0.0003 (0.53)	-0.0003 (0.40)	-0.0001 (0.16)
Constant	0.5443 (0.21)	-3.8961 (1.42)	-1.3035 (0.61)	-0.587 (0.52)	0.688 (0.40)	0.2536 (0.16)
Observations	726	726	726	726	726	726
R-squared	0.2	0.19	0.11	0.31	0.17	0.23

B. Is the pre-treatment *rate of change* of outcomes a function of the order of rollout?

Dependent variable is year/village average CHANGES in:

Explanatory variable:	Home Improvement	Television	Refrigerator	Stove	Bicycle	Cell Phone
Year in which credit received	0.0007 (0.37)	-0.0002 (0.08)	0.0001 (0.07)	0.0000 (0.03)	-0.0004 (0.33)	0.0000 (0.01)
Constant	-1.3866 (0.37)	0.213 (0.05)	-0.7029 (0.20)	0.0552 (0.03)	0.7551 (0.33)	-0.1778 (0.09)
Observations	655	655	655	655	655	655
R-squared	0.09	0.13	0.11	0.26	0.13	0.22

C. Is the order of rollout being driven by pre-treatment *shocks*?

Dependent variable is individual observation on:

Explanatory variable:	Home improvement	Television	Refrigerator	Stove	Bicycle	Cell Phone
Dummy, year before credit offered	-0.0118 (0.99)	-0.0163 (0.96)	-0.0304 (1.58)	0.0032 (0.46)	0.0252 (1.48)	-0.0102 (0.54)
Constant	0.0797 (2.31)*	0.0743 (2.30)*	-0.0324 (1.57)	0.052 (4.87)**	0.0134 (0.84)	-0.0024 (0.13)
Observations	774	774	774	774	774	774
R-squared	0.12	0.19	0.11	0.27	0.16	0.24

**Table 6: Test for Demand-Side Endogeneity:
Interactions between lack of access and leads of uptake**
Dependent variable:

Explanatory variable:	Home improvement	Television	Refrigerator	Cell Phone	Stove	Bicycle
2 years before receipt	-0.011 (1.11)	-0.013 (1.40)	-0.002 (0.35)	-0.007 (0.92)	-0.006 (1.35)	0.002 (0.24)
1 year before receipt	0.009 (0.91)	-0.007 (0.51)	-0.005 (0.47)	0.008 (0.88)	-0.003 (0.35)	-0.006 (0.83)
Year of receipt	0.028 (3.03)**	0.017 (1.67)	0.013 (1.83)	0.007 (0.65)	0.014 (1.92)	-0.002 (0.49)
1 year after receipt	0.036 (3.37)**	0.011 (0.70)	0.009 (1.18)	0.009 (0.50)	0.001 (0.08)	0.003 (0.63)
2 years after receipt	0.017 (1.73)	-0.002 (0.12)	0.02 (1.81)	0.007 (0.58)	-0.004 (0.50)	0.013 (1.27)
2 years before receipt, no access	0.01 (0.89)	0.019 (1.27)	0.015 (0.85)	0.012 (1.25)	0.02 (1.51)	0.021 (1.01)
1 year before receipt, no access	-0.004 (0.24)	-0.007 (0.33)	-0.016 (0.97)	-0.025 (1.85)	0 (0.02)	0.015 (1.25)
Constant	0.076 (3.89)**	0.002 (0.17)	0.006 (2.57)*	0.003 (1.07)	-0.015 (1.03)	-0.005 (0.72)
<i>P</i> -value on <i>F</i> -test that sum of (no access * leads) interactions is equal to zero:	0.7696	0.6532	0.9618	0.4870	0.1626	0.2056
R-squared	0.02	0.03	0.02	0.09	0.02	0.03

All regressions use country/year dummies and village-level fixed effects, SEs clustered at the village level.

Robust t statistics in parentheses

* significant at 5%; ** significant at 1%

Table 7: Test for Attrition Bias
Falsification Tests, Before-After Credit Dummy Variables Interacted with Dropout Dummy:
(Guatemala Only)

Dependent variable:

Explanatory variables:	Any Housing Improvement	New Walls	New Roof	New Floor	New Electricity	New Toilet
Actual Take-up of Credit:	0.02 (0.59)	0.012 (0.65)	-0.007 (0.50)	-0.002 (0.14)	0.001 (0.07)	-0.001 (0.03)
Dropout dummy	-0.01 (0.43)	0.064 (2.86)*	0.068 (3.30)**	0.032 (2.47)*	0.003 (0.31)	0.009 (1.30)
Dropout * Take-up of Credit	-0.033 (1.34)	0.033 (1.69)	0.028 (1.09)	-0.032 (1.35)	-0.004 (0.15)	0.038 (1.66)
Constant	0.078 (6.43)**	0.003 (0.14)	-0.007 (0.41)	0.052 (5.04)**	0.034 (2.98)*	0.005 (0.52)
Observations	3050	2635	2863	2558	2522	2736
R-squared	0.02	0.03	0.02	0.02	0.01	0.01

Whole Sample:

Dependent variable:

Explanatory variables:	Any Housing Improvement	New Walls	New Roof	New Floor	New Electricity	New Toilet
Credit Available in Village	0.034 (1.19)	0.011 (0.78)	-0.016 (1.13)	-0.011 (0.77)	0.039 (3.58)**	0.013 (0.94)
Dropout dummy	-0.017 (0.59)	0.026 (2.01)	0.027 (2.53)*	0.007 (0.37)	0.012 (1.43)	0.016 (2.08)
Dropout * Credit Available	-0.005 (0.18)	0.073 (1.65)	0.075 (1.72)	0.028 (2.11)	-0.017 (1.02)	0.004 (0.53)
Constant	0.08 (6.92)**	0.012 (0.90)	0.002 (0.17)	0.058 (5.95)**	0.032 (3.11)**	0.004 (0.39)
Observations	3050	2635	2863	2558	2522	2736
R-squared	0.02	0.03	0.03	0.02	0.01	0.01

Table 8: Strongest Test for Causality in Housing Improvements
Using Leads and Lags of Provision of Credit to the Village, weighted by dropout probability
(Guatemala Only)

Explanatory variable:	Any Housing Improvement	New Walls	New Roof	New Floor	New Electricity	New Toilet
4 years before access	-0.008 (0.32)	-0.013 (0.64)	-0.033 (2.13)	0.000 (0.03)	0.000 (0.01)	-0.015 (1.00)
3 years before access	-0.006 (0.22)	-0.05 (1.79)	-0.052 (1.68)	-0.006 (0.28)	-0.015 (1.32)	-0.012 (1.22)
2 years before access	0.018 (0.77)	-0.04 (1.23)	-0.037 (1.21)	-0.009 (0.55)	0.004 (0.23)	0.002 (0.08)
1 years before access	0.024 (1.21)	-0.025 (1.08)	-0.04 (1.59)	0.015 (1.08)	-0.016 (1.55)	0.012 (1.18)
Year of access	0.004 (0.22)	-0.015 (0.54)	-0.021 (0.96)	-0.002 (0.22)	0.036 (2.29)*	-0.009 (0.98)
1 year after access	0.015 (0.78)	-0.021 (1.08)	-0.054 (2.09)	-0.005 (0.45)	-0.02 (2.14)	0.009 (0.54)
2 years after access	-0.001 (0.11)	0.002 (0.07)	-0.05 (1.26)	0.014 (0.83)	0.007 (0.40)	-0.013 (0.67)
3 years after access	0.071 (2.60)*	0.008 (0.55)	-0.052 (1.38)	0.027 (1.22)	-0.001 (0.08)	0.038 (2.52)*
4 years after access	0.054 (2.20)*	0.016 (0.48)	-0.026 (0.74)	0.013 (0.55)	-0.017 (1.84)	0.001 (0.04)
Constant	0.051 (3.07)**	0.01 (0.64)	-0.004 (0.31)	0.052 (4.04)**	0.036 (3.32)**	-0.004 (0.40)
Observations	3,046	2,632	2,860	2,556	2,518	2,732
R-squared	0.05	0.03	0.03	0.02	0.02	0.02

All regressions use country/year dummies and village-level fixed effects.
SEs clustered at the village level.

Robust *t*-statistics in parentheses

* significant at 5%; ** significant at 1%

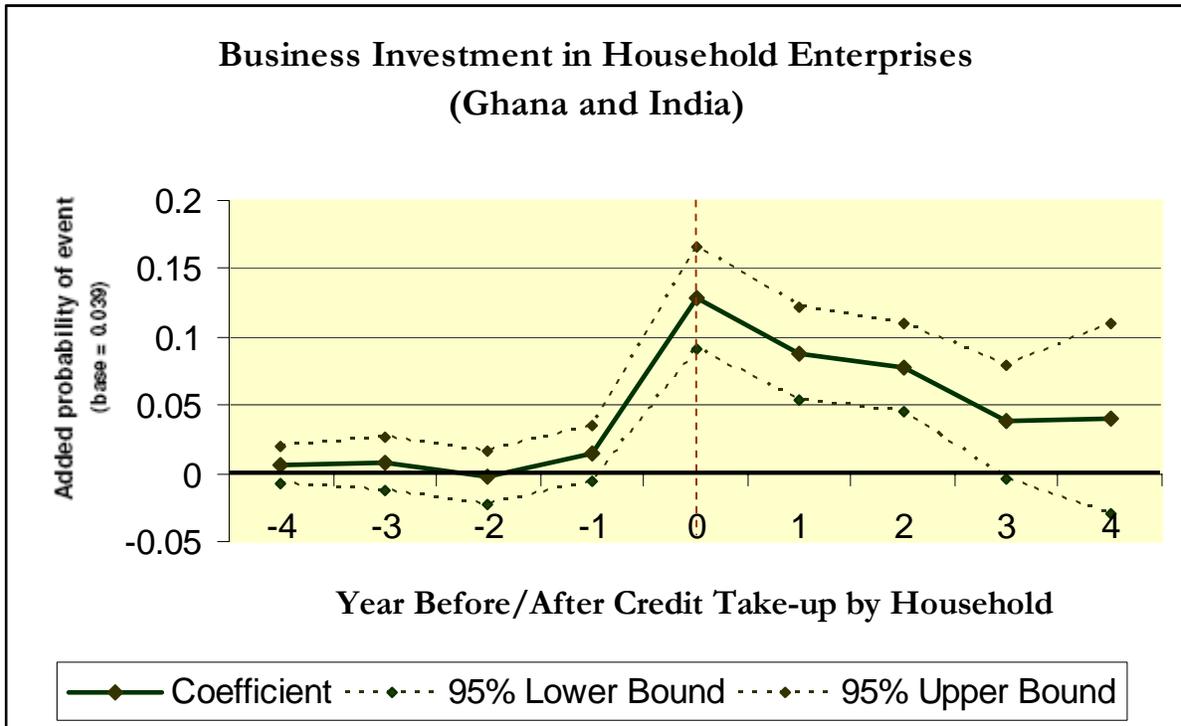


Figure 1

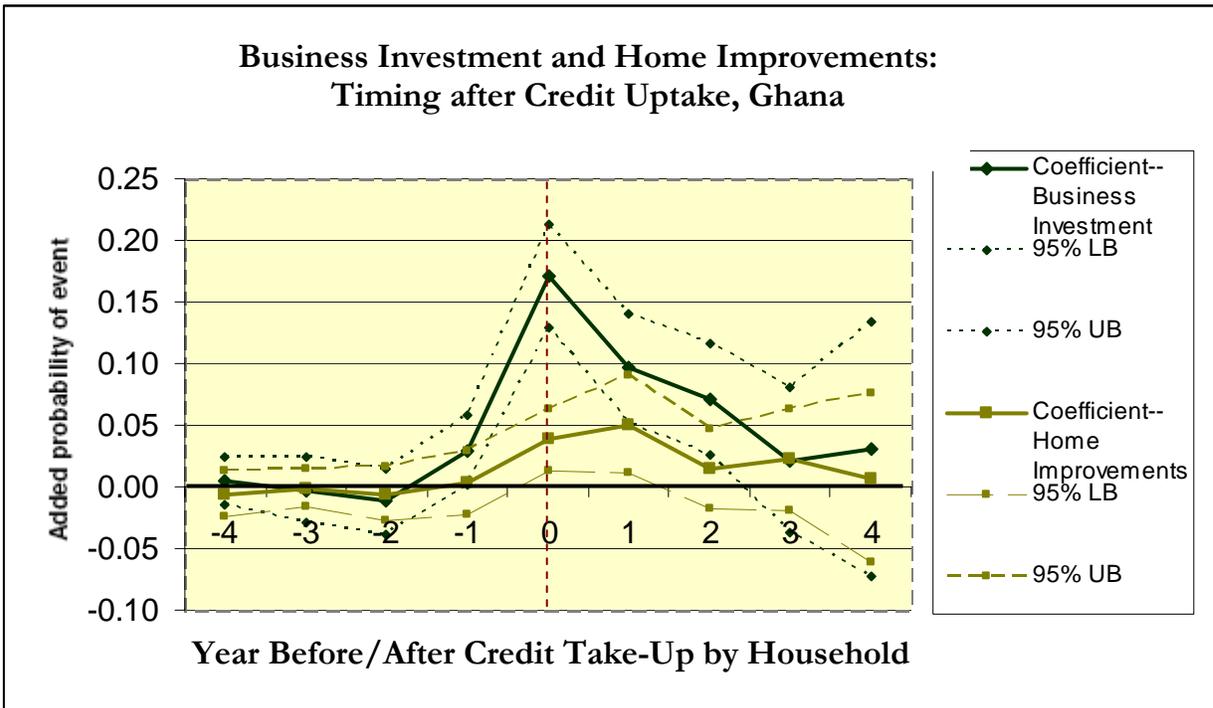


Figure 2

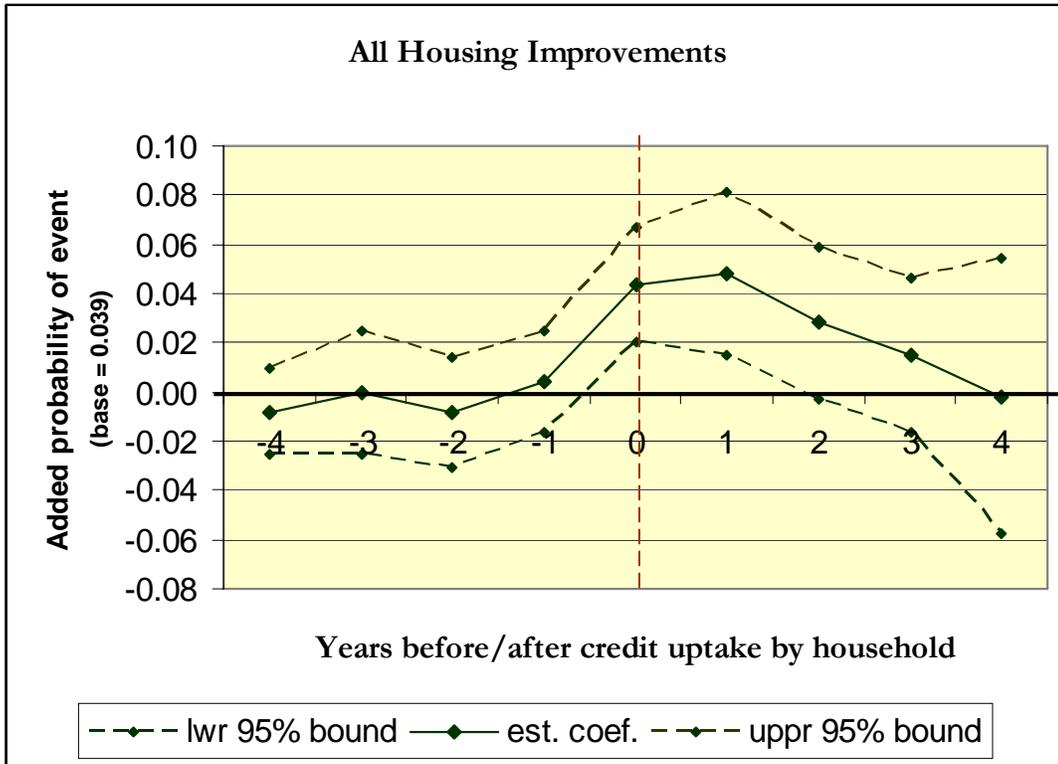


Figure 3

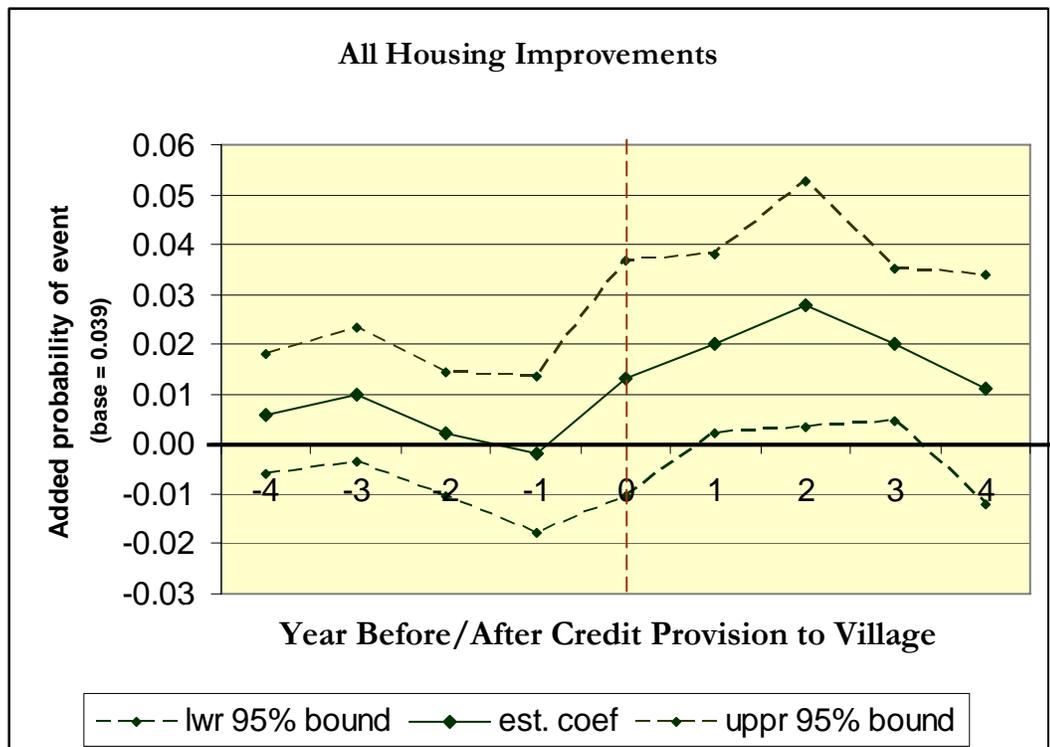


Figure 4

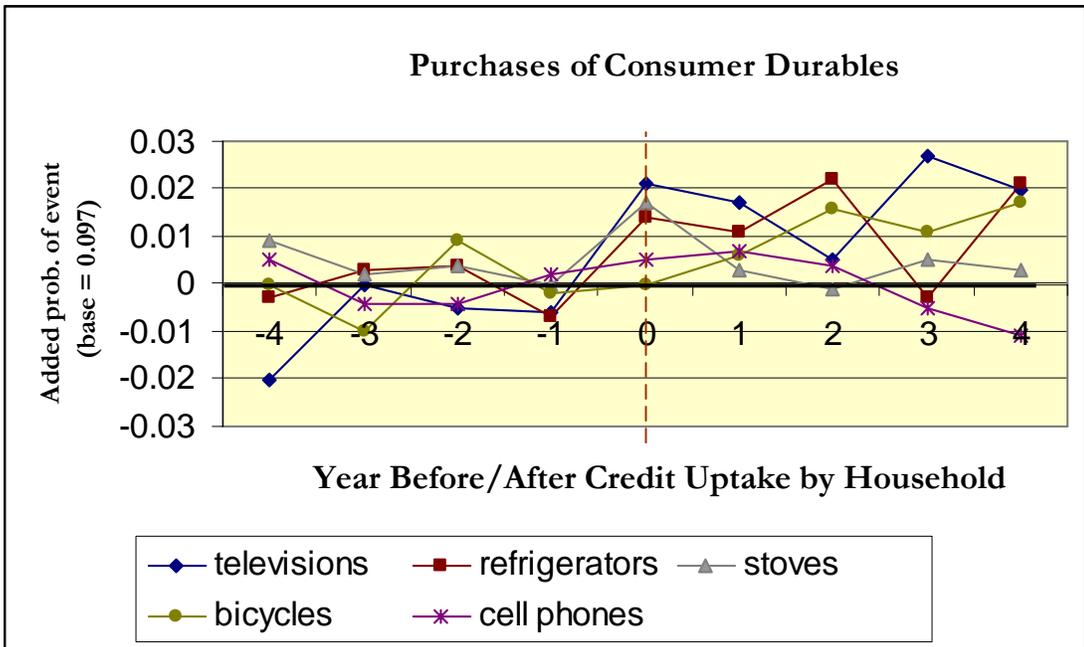


Figure 5

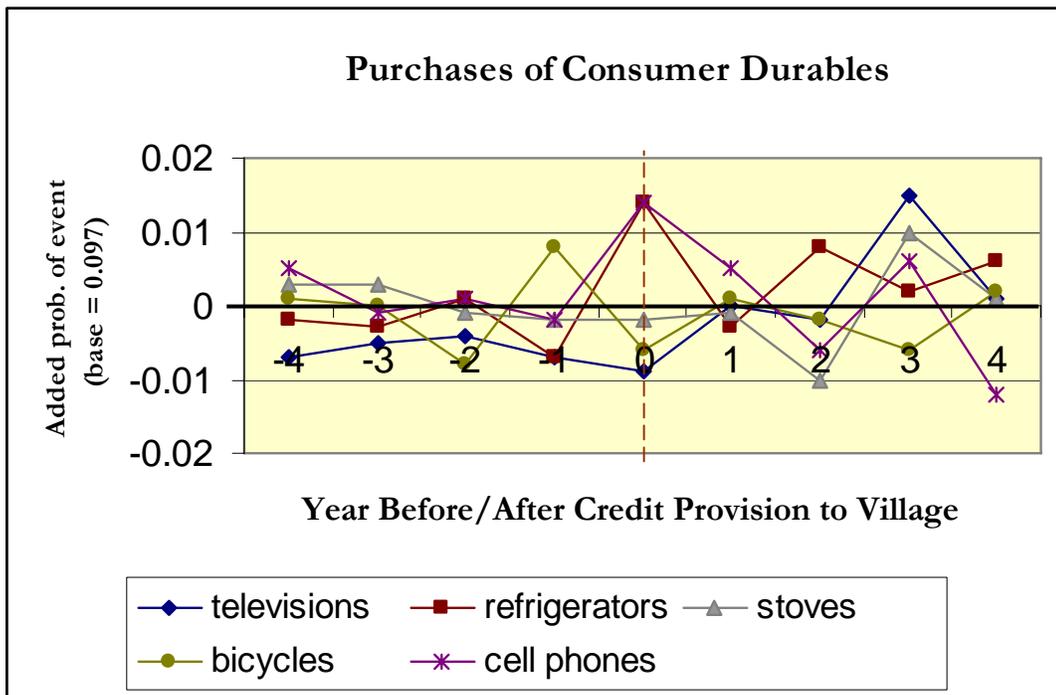


Figure 6