2014

Measuring Microfinance: Analyzing the Conflict between Practitioners and Researchers with Evidence from Nepal

Ram D. Rajbanshi

Meng Huang

Bruce Wydick
University of San Francisco, wydick@lucas.usfca.edu

Measuring Microfinance: Assessing the Conflict between Practitioners and Researchers with Evidence from Nepal

JEL Classifications: O12, O16, C21

Ram D. Rajbanshi[1]
Meng Huang[2]
Bruce Wydick[3]

November 4, 2014

Abstract: What accounts for the discrepancy between the microfinance impact claims of development practitioners and the far smaller impacts found in experimental studies? We demonstrate in a simple theoretical framework why "before-and-after" observations of practitioners overstate microfinance impacts and why estimations in some recent randomized trials understate the average treatment effect on the treated (ATT). Our empirical study uses a unique data set from eastern Nepal to study the impact of microfinance in villages where microfinance did not previously exist. We find that approximately three-fourths of the apparent impact of microfinance observed by practitioners is an illusion driven by correlated unobservable factors.

[1]Rajbanshi: Rural Microfinance Development Centre Ltd. (RMDC), Nepal, e-mail *rajbanshird@yahoo.com* .

[2]Huang: Senior Consultant, Bates White Economic Consulting, LLC, 3580 Carmel Mountain Road, Suite 420, San Diego, CA 92130, e-mail: *meng.huang@bateswhite.com.*

[3]Wydick: Professor of Economics, University of San Francisco, 2130 Fulton St. San Francisco CA 94117-1080, e-mail: *wydick@usfca.edu*.

We would like to thank staff at *Jeevan Bikas Samaj* in Nepal for cooperation and assistance. This work was supported by the U.S. Agency for International Development (USAID), the University of San Francisco Faculty Development Fund, and the University of San Francisco graduate program in International and Development Economics. We thank Goldie Chow and Phillip Ross for research assistance and Chris Barrett, Steve Boucher, Alain de Janvry, Jeremy McGruder, Craig McIntosh, Ted Miguel, Elizabeth Sadoulet, Timothy Van Vugt and seminar participants at the 2012 Pacific Conference for Development Economics and the Development Economics Seminar at UC Berkeley for helpful comments related to this research.

1. INTRODUCTION

Microfinance emerged in the 1970s as an effective strategy to increase credit access among the poor in developing countries who were routinely shunned by formal lenders and left to borrow from informal money lenders at elevated interest rates. The growth of microfinance since this time has been unprecedented. The 2014 Microcredit Summit reports that in 2011 there were 203 million microfinance borrowers in the developing world, among these being 116 million of the world's poor living on less than $1.25 per day.[1]

Yet even with the tremendous resources that have flowed into microlending, there has been substantial disagreement regarding the impact of microfinance. Anecdotal evidence from practitioners in the field tends to overwhelmingly report growth in enterprises, income, and improvements in household welfare after borrowers take microfinance loans. Academic studies are more mixed. While some studies have found substantive impacts of microfinance on household income, consumption, and poverty reduction (Pitt & Khandker, 1998; Gomez and Santor, 2003; Khandker, 2005; Berhane & Gardebroek, 2011; Imai et al., 2012; Field et al., 2013), many studies find only modest or even *no* impact from microfinance (Morduch, 1998; Coleman, 1999 and 2002; Dingcong et al., 2008; Roodman & Morduch, 2009; Karlan & Zinman, 2011; Giné & Mansuri (2011); Angelucci et al., 2012; Attanasio et al., 2011; Augsburg et al., 2012; Banerjee et al., 2013; Crépon et al., 2013).[2]

We demonstrate in a simple theoretical framework that the discrepancy between practitioners and academics stem from two factors. First is the presence of unobserved shocks affecting the *timing* with which a given borrower takes a microfinance loan. These phenomena, which may be related to economic opportunities or changes in self-motivation, are generally unobservable to both practitioners and researchers, but are complementary to microfinance borrowing and enterprise investment. We show that because borrowers have an incentive to take loans at the same time that these opportunities or increases in self-motivation arise, a large portion of the gain in enterprise and household welfare that practitioners observe before and after microfinance loans is illusory, not properly accounting for the counterfactual.

Second, we demonstrate that the impact estimates of many recent experimental and quasi-experimental studies are likely to underestimate the average treatment effects of microfinance. A major shortcoming of even some of the most celebrated microfinance impact studies (e.g. Coleman, 1999; Karlan & Zinman, 2011; Augsburg et al. (2012); Banerjee et al., 2013) is that these studies were implemented in areas where considerable microfinance lending already existed among the treated population when new microfinance was made available in the experimental design. Such studies are forced to obtain impacts on marginal borrowers induced into taking new microfinance loans as part of the implemented research design. While the authors of these papers are generally careful to qualify their results as impacts limited to "compliers," they are often taken by the development community to represent the average impact of microfinance generally, not simply the average impact of microfinance on the last borrowers in an area to take microfinance loans.

Unlike many types of interventions where an estimate on later compliers presents a reasonable approximation to the average treatment effect on the treated (ATT), such as a health intervention where impacts on the human body of a drug is likely to be independent of treatment order, we demonstrate why the impact on late takers of microfinance should be lower than that of the first. The reason is that being an early taker of microfinance is positively correlated with borrower productivity. This productivity in our model is based on hidden economic opportunity that is made up of the sum of two unobservable components: a fixed ability component for each borrower and an idiosyncratic measure of economic opportunity that varies for each potential borrower in each period. Based on the confluence of these two factors, borrowers who take loans in early stages are more likely to possess a higher base level of productivity than those who take up loans in latter stages; hence on average they will realize a higher level of impact from borrowing. Although similarly difficult to quantify, the error from inferring average microfinance impacts across the entire population of borrowers from these marginal borrowers may be no less grave than the well-known biases of borrower self-selection or attrition that researchers have been careful to address in the past.

Recognizing the potential problems with carrying out microfinance impact studies in areas with heavy existing microfinance lending, Crépon, Devoto, Duflo, & Parienté (2013) present results of a randomized microfinance experiment in Morocco carried out in a region in which virtually no microfinance existed before the experiment. Treated villages were subject to the promotion of group-lending based microfinance. Research results from the study show that access to credit resulted in expansion of enterprises and increased profitability, although increases in profits appear to be limited to agricultural enterprises. Interestingly, however, their results show that MFI expansion and increased profitability were realized at a cost: As labor was reallocated to microenterprises, increased profits in these enterprises were offset by a decrease in wages earned in the labor market, so that the net change in income and consumption in these households is insignificant even two years after rollout.

Our study is similar to Crépon et al. (2013) in that it is one of the first efforts to study the impact of microfinance in a rural area previously untouched by microfinance. But in contrast to Crépon et. al., our study uses non-experimental data, employing two different types of estimation. First we use an event study (RETRAFECT) methodology (Retrospective Analysis of Fundamental Events Contiguous to Treatment—see McIntosh et al., 2011) to analyze dynamic changes in the probabilities of fundamental events within an event window over the years surrounding treatment. We also employ more conventional difference-in-difference estimations, which are better at gauging single-parameter impacts.

Our event study methodology borrows from the finance literature, in which it is often used to gauge the impact of announcements of mergers and acquisitions on stock prices. (For an excellent review of event studies, see MacKinlay, 1997.) Instead of analyzing the impact of firm behavior on stock prices, our enumerators compiled a household history of fundamental 1/0 events indicative of "development" that could be accurately pinned to a particular year in the life of the household. The RETRAFECT event-study methodology then examines how the probability of these events changes surrounding a treatment such as the introduction of a microfinance lender in a village or the take-up of microfinance itself.

The data set we use in this analysis is unique in three respects. First, we have a strong understanding about why certain villages received credit before others, where the microlender expanding operations introduced lending into villages off main highways first, and later into nearby villages off these main highways (and not, subject to this expansion decision rule, that the lender initially targeted better off or worse off villages.) Second, like the Crépon et al. study, it is taken from a rural region in which microfinance did not exist before it was introduced by the lender in our study. The third unique aspect of our data is that take-up of microfinance was extremely high in our study area when credit became available: 51% among our random sample of female entrepreneurs in the six villages from which we collect data, much higher than even the 12% take-up rate among the treated villages in the Crépon et al. study. The high take-up rate allows us greater latitude in measuring the impact of microfinance as it has been adopted widely within a given population, and yields considerable power in first-stage instrumental variable estimations.

We use the RETRAFECT event study methodology to estimate dynamic intention-to-treat effects (ITT) controlling for observables while using a region-year fixed effect across paired villages, one of which received credit earlier and a neighboring village that received credit approximately two years later. Using a cross-sectional survey of 703 households in Nepal, enumerators carefully interviewed family members to create a household history over the previous ten years of discrete, fundamental events to the household that we would expect to be associated with growth in enterprises and increases in household welfare: major capital investments in enterprises, major home improvements, and purchases of consumer durables. The product of these 703 household histories is a retrospective panel data set that contains a matrix of over 7,000 dummy variables consisting of these relatively rare (average probability = 0.037) fundamental events.

We believe there are at least two important results from our study that add to this ongoing investigation into the worldwide impact of microfinance. First, we find that subsequent to taking a microfinance loan, microfinance borrowers are significantly more likely to experience an increase in the probability of a broad array of investment and consumption variables. In this respect, the anecdotal evidence cited by practitioners is corroborated by our data, and thus we find that practitioners are neither lying nor likely even exaggerating the positive changes to MFI borrowers that they witness after borrowers take microfinance loans. However, when we analyze changes in the probability of these events within a window surrounding the *introduction* of microfinance into our six villages, allowing us to estimate average treatment effects based on credit availability, we find the *causal* effect of microfinance to be far more modest than the *apparent* impact that a "before-and-after" analysis would suggest. Indeed we estimate that approximately 3/4 of the impact testified to by practitioners after taking microfinance appear to be something akin to an optical illusion, driven by correlated unobservables. But we also find that the approximately one-fourth of this "apparent impact" that remains is non-trivial; our study finds impacts that are lower than those reported anecdotally by practitioners, but higher than those reported by most recent experimental and quasi-experimental studies.

The second striking result from this data is the similarity of the microfinance impact results obtained using the event study methodology with the Crépon et. al (2013) experimental study, the best study of which we are aware that experimentally estimates microfinance impacts in a region previously unserved by microfinance. Although carried out in different countries

(Morroco and Nepal) using different methodologies, the results of both indicate that microfinance has a strong impact on microenterprise expansion and on the size of livestock herds (depending on the type of enterprise), but no significant impact on consumption. These results contrast to some degree with other work on microfinance published in this *journal*, including Imai et al. (2012) who find that countries with higher per capita microlending tend to have lower levels of poverty indices and Becchetti and Castriota's (2011) work on microfinance in post disaster Sri Lanka, in which they find positive impacts on work hours and real income.

The remainder of our paper is organized as follows. Section 2 presents a simple empirical model that illustrates the potential for bias in both informally and formally ascertaining microfinance impact. Section 3 presents survey area, history of MFI, survey methodology and data. Section 4 discusses data analysis and results. Section 5 summarizes and reflects on our results.

2. A SIMPLE MODEL OF MICROFINANCE BORROWING

Consider an economy of *n* household enterprises, each with a parsimonious income function of

$$Y_{it} = \phi_{it} K_{it},$$

where Y_{it} is net income of household *i* in period *t*, K_{it} is the existing capital stock, and ϕ_{it} is an unobservable complement to capital, such as economic opportunity, entrepreneurial ability, or self-motivation.[3] For brevity's sake we will call ϕ_{it} "economic opportunity." In our model ϕ_{it} is a composite of two additive terms, ϕ_i, which is fixed for each *i*, and ϕ_t, which is stochastic and varies over time for each *i*.[4] Thus $\phi_{it} = \phi_i + \phi_t$, each term with respective variance σ_i^2 and σ_t^2. For simplicity of exposition, assume that the distributions of ϕ_i and ϕ_t across the population are such that all realizations of ϕ_{it} lie in the unit interval and are equally likely, *i.e.* $\phi_{it} \sim U[0,1]$. Higher realizations of ϕ_{it}, however, are obviously more (less) likely among households with higher (lower) ϕ_i. The price of output is equal to one so that output is equal to income, which may be spent on consumer goods or added to the next period's capital stock.

Each household enterprise is endowed with K_{it} units of capital, but may borrow to finance one more unit of capital. At the end of the period *t*, a borrowing household returns the one unit of borrowed capital, the principal, along with interest. The borrowing decision depends on the difference borrowing makes to net income; a household borrows if the rate of return on capital equals or exceeds the interest rate.

We assume that prior to the availability of microfinance, the credit market is serviced by moneylenders. Moneylenders have inside information about ϕ_{it} and are fully exploitative first-degree price discriminators, charging an interest rate $r_l = \phi_{it}$. The microfinance interest rate, however, is the same for all borrowers and across all time periods, lying between the extreme values of ϕ_{it}, *i.e.*, $r \in (0, 1)$.

(a) *Apparent Impact of Microfinance*

Microfinance practitioners and other observers of microfinance often make impact assessments based on welfare differences in households and their enterprises before and after taking microfinance loans.[5] Impact is also at times validated by welfare differences between borrowers and non-borrowers. But what we call the "apparent impact" of microfinance is influenced by the timing at which enterprises take microfinance loans. Enterprises only take microfinance loans when opportunity dictates that it is profitable to do so. The average enterprise income that is observed after taking a microfinance loan in period *t* is

$$E(Y_{it}|T=1) = E(Y_{it}|\phi_{it} \geq r) = (K_{it}+1)\phi_f - r \quad (1)$$

where $\phi_f = \frac{1}{2}(1+r)$, and $T = 1$ and $T = 0$ represent the states of treatment and non-treatment (microfinance borrowing), respectively. In other words, after taking a microfinance loan, income is equal to existing capital plus borrowed capital multiplied by opportunity less interest costs. This contrasts with the income that is observed in microenterprises when they do not chose to take a microfinance loan, whose realization of "opportunity" is smaller:

$$E(Y_{it}|T=0) = E(Y_{it}|\phi_{it} < r) = K_{it}\phi_l \quad (2)$$

where $\phi_l = \frac{1}{2}(r)$. The difference in outcomes between borrowing and non-borrowing households (or from before and after a given household enterprise *i* takes a loan), the "apparent impact" of microfinance, is equal to

$$E(Y_{it}|T=1) - E(Y_{it}|T=0) = \frac{1}{2}[K_{it}+1-r]. \quad (3)$$

In the framework of the Rubin (1974) causal model, suppose we were to estimate the average treatment effect of microfinance on microfinance borrowers who would take a microfinance loan if it were available. Let Y_{it}^1 and Y_{it}^0 represent *potential* outcomes for a microenterprise in treated and untreated states, respectively, *i.e.*,

$$Y_{it}^1 = \phi_{it}(K_{it}+1) - r, \quad (4)$$

and

$$Y_{it}^{0} = \phi_{it}\, K_{it}. \tag{5}$$

In (4) the outcome represents net income after borrowing, output less interest costs, while (5) is net income for the non-borrower. Every household enterprise, whether it in fact borrows or not, can realize *either* potential outcome. But only one of the potential outcomes is observed since the treatment is either assigned or not for the same household enterprise, but not both: for borrowers, only Y_{it}^{1} is observed; for non-borrowers, only Y_{it}^{0} is observed. For each household enterprise *i*, the causal effect is $Y_{it}^{1} - Y_{it}^{0}$. For the treated group (*i.e.*, borrowers), the average treatment effect is thus

$$E({Y_{it}}^{1} - {Y_{it}}^{0}|T = 1) = E({Y_{it}}^{1} - {Y_{it}}^{0}|\phi_{it} \geq r) = \phi_f - r = \tfrac{1}{2}(1 - r), \tag{6}$$

so that the effect on the treated is a function of the lower interest rate from the availability of microfinance. The upward bias in "apparent impact" based on differences between borrowing and non-borrowing households is the difference between (3) and (6), or $\frac{1}{2}K_{it}$, which is increasing in the level of existing capital. If we rewrite the apparent effect in (3) in Rubin's framework, we have

$$E(Y_{it}|T = 1) - E(Y_{it}|T = 0) = E\left({Y_{it}}^{1}\middle|T = 1\right) - E\left({Y_{it}}^{0}\middle|T = 0\right). \tag{3'}$$

Comparing (3') and (6), we see the source of the bias:

$$\begin{aligned} &\left[E\left({Y_{it}}^{1}\middle|T = 1\right) - E\left({Y_{it}}^{0}\middle|T = 0\right)\right] - \left[E\left({Y_{it}}^{1}\middle|T = 1\right) - E\left({Y_{it}}^{0}\middle|T = 1\right)\right] \\ &= E\left({Y_{it}}^{0}\middle|T = 1\right) - E\left({Y_{it}}^{0}\middle|T = 0\right) = K_{it}\phi_f - K_{it}\phi_l = \tfrac{1}{2}\, K_{it}. \end{aligned} \tag{7}$$

As can be seen in (7), the bias results from trying to estimate the potential outcome of the treated group were they not treated by the average outcome of the non-treated group. However, the average potential outcome if the household is not treated (Y_{it}^{0}) would be different for the treated and non-treated groups (borrowers and non-borrowers), because the average opportunity of the treated is higher. Opportunity both affects outcomes and induces treatment, and hence leads to bias if it is used as the basis for comparison. This is because households self-select into the treated group if and only if their economic opportunities yield a return greater than the interest rate.

(b) Potential Bias in Randomized Studies in Estimating Average Treatment Effects of Microfinance

Recent studies have sought to measure the impact of microfinance through field experiments. Some of these studies have used experimental designs that have induced new borrowers to take microfinance in areas in which microfinance lending already exists. This means that when researchers estimate either the intention to treat (ITT) or the local average treatment effect (LATE), impacts are obtained on marginal "late-coming" borrowers (compliers) who take up microfinance loans often after credit has been previously available.

While these studies may be able to identify the impact of *credit expansion* in a given area to new borrowers, they provide unbiased estimates of the average treatment effect of microfinance across the population only under a restricted set of conditions that would seem unlikely in most field contexts. One potential context would be in which borrowers are credit constrained in the sense of Stiglitz & Weiss (1981), the assumption behind work such as Karlan & Zinmann (2011). But note that for the LATE to estimate the ATT, even when satisfying the stable-unit-treatment-value assumption (SUTVA), the status of being credit constrained must be fully orthogonal to potential outcomes, which seems unlikely. Another exception in which we would expect LATE estimates to closely approximate the ATT would be a case in which *all* borrowers induced to take loans through the encouragement of the experiment were *information*-constrained. Yet a more likely case is the intermediate one, where there are information constraints on only a fraction of new compliers. The degree of the bias would then depend on the relative fraction of new borrowers who were induced to take loans based on exposure to the information, and the fraction of those taking loans through realizing higher levels of economic opportunity than existed in previous periods. For the latter, if it would have been profitable for the borrowers to take microfinance loans previously, they likely would have done so.

Although the phenomenon has not been recognized routinely by previous work, it is straightforward to show that average treatment effects on later-takers of microfinance should be lower than impacts on early adopters. This holds true in the context of our model except in the special case in which the variance in the fixed individual component of opportunity, σ_i^2, is zero across individuals and that all of the variation in economic opportunity resides in the idiosyncratic component of economic opportunity that varies over time, σ_t^2. In this special (extreme) case, marginal borrowers taking credit in later periods in an established market ought to realize long-term impacts from microfinance that are similar to existing borrowers since they are identical except for differences in the (later) timing of high realizations of the idiosyncratic component of opportunity, ϕ_t.

However, as σ_i^2 becomes larger relative to σ_t^2, then the average productivity of borrowers taking microfinance loans for the first time becomes lower in every sequential period in which microfinance is offered. This is because (assuming $t = 1$ to be the period when credit is introduced) $E[\phi_i \mid \phi_{1t} \leq r\,] < E[\phi_i \mid \phi_{1t} > r\,]$, which directly implies that

$$E[Y_{it} \mid \phi_{1t} \leq r\,] < E[Y_{it} \mid \phi_{1t} > r\,] \text{ for all } t > 1. \quad (8)$$

The intuition is that given $\sigma_i^2 > 0$, borrowers with high ϕ_i are more likely to have a level of unobserved economic opportunity that exceeds the required threshold, $\phi_{it} > r$, upon initial introduction of microfinance; by definition such borrowers are more likely to continue to exhibit high long-term impacts from borrowing. In contrast, borrowers with low ϕ_i are more likely to have realizations of $\phi_{it} \leq r$ when microfinance is first introduced, take up credit only in latter periods when ϕ_t is high, and display lower impacts from it over time. (They are also more likely to drop out of the microfinance borrowing pool, since all borrowers only borrow in periods with sufficiently high ϕ_t that $\phi_{it} > r$.) Consequently, because borrowers with low ϕ_i are less productive on average, microfinance will have a smaller expected impact on borrowers who take their first microfinance loans at later stages.

In summary, the discrepancy between the observations of microfinance practitioners and the estimated microfinance impact of recent experimental studies may result from two factors: (1) Practitioner observations of microfinance borrowers before-and-after borrowing and comparisons between borrowers and non-borrowers systematically overestimate the ATT; and (2) Researcher estimates of microfinance impacts from the uptake of microfinance in later periods after credit introduction are lower than the ATT across all borrowers.

In contrast to most recent studies carried out in areas already served by microfinance lenders, we present microfinance impact estimations from an area previously unserved by microfinance. Thus in these estimations carried out on virgin microfinance areas where credit availability is used as an instrument for take-up, our estimation of the LATE converges to an "instrumented ATT" in the absence of "always takers" in the data (since clearly there can be no takers of microfinance when microfinance is not available). We are able to find evidence of the both upward bias of "before-and-after" borrowing observations and the underestimation of microfinance treatment effects in studies carried out where microfinance has already existed.

3. SURVEY AREA, HISTORY OF MFI, METHODOLOGY, AND DATA

(a) Microfinance Institution and Survey Area

The data for this survey was collected in summer 2010 in the eastern Terai[6] region (Morang and Jhapa districts) of Nepal, in the operating areas of the *Jeevan Bikas Samaj*[7]

(Livelihood Development Society), the first MFI to operate in the region. Several years after it began operations in 2005, other MFIs began to start lending in the region. Established in September 1997 as a local club by a small group of local youths in village *Amahi Bariyati*, a remote village in the South-East region of the country, JBS carried out welfare and social activities for the local people living close around its club office. Initially, JBS had no plan of engaging in microfinance extensively. But in 2003 the club met with the Rural Microfinance Development Centre Ltd. (RMDC), a second-tier microfinance organization, which encouraged the club to implement microfinance program in a wider scale. After having some experience and intensive support of RMDC, JBS established its first branch office in 2005.

JBS uses a group lending methodology similar to the classical Grameen group lending approach.[8] Borrowing groups consist of five adult women of working age; five to ten groups form a center. Groups are formed by eligible women themselves, not by the MFI. Eligibility is determined using the following criteria: A borrower must be (a) female, (b) aged 18-60, (c) residing in the program area, (d) capable of income-generating activities, and (e) has a citizenship or other valid identification card. The center meetings are held every two weeks in the borrowers' neighborhood, where a field staff conducts financial transactions, including collection of loan applications, loan repayments, and member savings.

A loan to a borrower starts from (Nepal Rupees) NRs3,000 ($42) and gradually increases up to NRs60,000 ($833). The loans have a one-year term, and the effective annual interest rate is approximately 25%. The average size of a loan to a borrower in our sample is 13,807(US$192), many loans being used for livestock investment, typically bullocks, or expansion of rural enterprises, often small village shops. In our sample, borrowers took two (one-year) loans on average. Clients are not strictly required to start a business from the loans: the MFI recognizes that money is fungible and clients are left entirely free to choose the best use of money as long as loans are repaid. Along with credit, the MFI offers compulsory and voluntary savings products, but JBS is principally a lending organization, and not directly involved in activities such as business training or financial literacy promotion.

After its initial lending began in 2005, JBS began to expand its microfinance lending operations into new villages that had never before had access to microfinance. The manner in which JBS introduces its microfinance program into new villages is methodical, transparent, and important to our impact-identification strategy.

According to directors, new expansion of JBS lending into villages was not based on local economic conditions, which were viewed to be more-or-less uniform, but rather by a methodical geographical expansion into rural eastern Nepal. JBS begins its operations in an area by opening a major branch office in a new region, first implementing its program in the area located just around its branch office, and then after one or more years of operation, expands over time to provide credit to one or more nearby villages in that region. The first branch office typically is located in a larger village or town that is on a main road. Later, lending is introduced into villages that are located on smaller roads that lie off the initial main road, or farther away on the main road from urban centers until regional coverage is achieved. Although this was the pattern of the JBS rollout, due to uncertainties with regards to funding and other logistical matters, it was certainly not clear to potential borrowers in the secondary villages that JBS microfinance would ever reach their own location. Due to these uncertainties, at the time there was never any solid basis for the belief that credit was eminently arriving into the secondary villages, and anticipation which could have potentially affected borrower behavior.

The six villages in our study were chosen in the following way. We chose three branch areas in which microfinance was introduced into a population for the first time in sequential years, 2005, 2006, and 2007. Microfinance started in a single village in each of these years, but then as JBS extended its credit coverage within each region, other villages near the first village were ultimately included in the program. For each of these three village development committees (VDCs)[9] that received microfinance program in 2005-2007, we chose a second paired village--located within of 5-10 KM of the initial village—that received access to the microfinance program one to two years after the first village in the region gained access to microfinance. We limited our study to these six paired villages because they represented the cleanest example in which microfinance was introduced sequentially, and for the very first time.

The villages include, from Karsiya branch, Dadarbairiya (2005)[10] and Thalaha (2007); from Itabhatta branch, Jyamirgadhi (2006) and Bahundangi (2008); and from Amardaha branch; Amardaha (2007) and Govindpur (2008) (see map in Figure 1). Thus, villages are intentionally grouped into three pairs, one village a first recipient of JBS credit in the early years of operation from 2005 to 2007, and then a second neighboring village that was provided access to JBS credit one to two years later when funding allowed for expansion into the neighboring village. In Figure 1, the villages with the dark circles indicate where credit was introduced first village among the pair (Dadarbairiya, Jyamirgadhi, and Amardaha); the villages with lighter circles indicate the second villages in the pair (Thalaha, Bahundangi, and Govindpur), that gained access to microfinance credit in a subsequent year.

Within these villages, we randomly selected seventy to ninety households in each of the six villages for the survey who met the aforementioned eligibility criteria for JBS. Because microfinance had not been previously available and because there was a great demand for it, take-up among the randomly sampled population had been extremely high; 244 of our 478 randomly selected households that met the lending criteria from the villages had taken credit with JBS. In addition, we intentionally included another random sample of 225 households that had been members of JBS, thus we oversampled borrowers, reweighting our estimations subsequently.

(b) Empirical Methods

We use two approaches to estimating microfinance impacts from the JBS program. First we present results from an event-study methodology, Retrospective Analysis of Fundamental Events Contiguous to Treatment (RETRAFECT). We derive in the Appendix the technical conditions under which RETRAFECT estimates can have a causal interpretation. In brief, causal interpretation of parameters relies on the assumption given in (A7) that $\ddot{Z}_{it} \perp T_i^{t-k,t+k} | \tilde{X}_{it}$, where $\ddot{Z}_{it}$ is a vector of unobservable covariates that must be orthogonal to treatment T_i during the event window $t - k$ to $t + k$ given a vector $\tilde{X}_{it}$ of observed covariates, that includes observed causes and proxies. This assumption is a dynamic form of the basic unconfoundedness assumption originally developed by Rubin (1974, 1977).

In this methodology, enumerators work with subjects to carefully create a household history of fundamental events that have a theoretical basis for responding to development interventions such as microfinance. Following McIntosh *et al.* (2011), these events must be (a) discrete, measured only in terms of ones (event occurred) and zeros (absence of the event); (b) memorable and able to be pinpointed to a particular year; (c) reflect welfare changes of a household; and (d) potentially responsive to the intervention. All survey questions were designed to be compatible with these criteria.

Secondly, we use this data on fundamental events in difference-in-difference estimations for which we organize our data into the three regions, each with one village receiving credit one to two years earlier than its partner village. We first check to verify that rates of investment, home improvement, and durable goods consumption were not statistically different between these pair villages, and then test for pre-treatment parallel trends between villages that received credit first and those who received it later. Using simple *t*-tests and difference-in-differences regression, we then compare (a) differences between the two villages during the years before the first village in each region gained access to microfinance, with (b) differences between the two villages during the years when the first village enjoyed microfinance access while the second village did not.

The fundamental events we used in this study recorded major capital investments in a household enterprise (purchase of livestock, new agricultural land, major machinery, a new store or kiosk), discrete changes in home improvements (building a new roof on a house, new walls, a major floor upgrade, a new indoor toilet), and the first purchase of major household consumer durables (television, bicycle, cell phone, stove).

A history of these 1/0 events for each household are recorded in a matrix of a backcast panel data over the previous ten years, from year 2001 through 2010. The psychology literature has shown that in collecting this type of survey data it is often helpful to place events in a time context with other events (Conway & Bekerian, 1987). Indeed in recent work in among Indian fisherman, Giné & de Francesca (2011) show that the use of these benchmarks is important for accuracy in collecting recall data. In our study when it was difficult for respondents to determine the year of an event, enumerators used ages of children, births and deaths of family members, years of important cricket matches, and other key benchmarks in the household history to pinpoint the year in which the event took place. Changes in the probabilities of these events are estimated around the timing of the treatment (program access or credit take-up).

In our sample population, 58% household heads are illiterate, 33% have primary education and 8% have high school education. Average age of the household heads is 39.9 years, number of members in a family is 4.9, and average land holding per family is 0.61 acres. Agricultural production is main occupation of 26% families, (agricultural) retail trade[11] of 54% families and both agriculture production and retail trade of 20% families. Two-thirds (67%) of the households in our sample were borrowers of JBS. The dropout rate was very low (2%) from the program in our survey sample. Descriptive statistics and pre-credit baselines of the dependent and control variables are shown in Table 1. In the final two columns we give means of dependent variables in the two years before the rollout occurred in first village of the three village pairs. We see that consumption of televisions and bicycles were somewhat higher in the first villages to receive the microfinance program and land purchases were higher in the second villages to receive the treatment within the pairs. But all other variables regarding business investment, housing improvements, and consumer durables were not insignificantly different between the village pairs before program rollout.

(c) Empirical Estimations

In our event-study estimations we use a linear probability model (LPM) instead of probit or logit models, since as a linear estimator it produces more robust estimates when flat panel datasets are used with fixed-effect estimations (Chamberlin, 1979). Moreover, with low-frequency events such as those in our survey (where the average probability of one of these fundamental events is only 0.037), the LPM fares well because there is little risk of projecting probabilities outside of the [0, 1] probability space when expected effects are positive. Our RETRAFECT model estimates

$$y_{ijt} = \alpha_{jt} + X'_{it}\beta^C + \sum_{t=\bar{t}-k}^{\bar{t}+k} \tau_{it}^C C_{it} + u_{ijt}^C \qquad (9)$$

$$y_{ijt} = \alpha_{jt} + X'_{it}\beta^T + \sum_{t=\bar{t}-k}^{\bar{t}+k} \tau_{it}^T T_{it} + u_{ijt}^T \qquad (9')$$

where y_{ijt} is binary economic impact variable that is equal to 1 if a given event occurs for household i living in region j at time t (e.g. an upgrade to a dwelling). Right-hand-side variables include a region-year fixed effect, α_{jt}, a vector of control variables, X', and the event study treatment window with coefficients τ_{ijt} and lead/lag indicator dummies C_{ijt} and T_{ijt}, which represent credit take-up and credit availability, respectively. Thus, the coefficient τ_{it}^C, for example, gives the difference in outcome y_{ijt} between the treatment villages when they had access to microfinance for one year relative to the partner villages in that same year. We set $k = 4$, and thus estimate a symmetric 9-year credit-treatment-window containing four leads and lags variables around treatment.

An advantage of the RETRAFECT event-study estimations is that they are able to illustrate the dynamics of impact over a number of years. However, our data also allow for difference-in-difference estimations which are more appropriate for estimating standard impact statistics. But first, to obtain an estimate of what practitioners observe, we look at the differences in outcomes between borrowing and non-borrowing examining changes on fundamental events that occur when a household takes a microfinance loan, corresponding to equation (3) in our theoretical framework. For the “apparent effect” observed from credit take-up we estimate

$$y_{ijt} = \alpha_j + \boldsymbol{X_i'\beta} + \pi Y_t + \delta C_{it} + e_{it} \tag{10}$$

where α_j is a regional fixed effect, $\boldsymbol{X_i'}$ is a vector of controls that includes gender, age, age-squared, education, and type of enterprise, Y_t is a year time trend, C_{it} is a dummy variable turning to 1 when household *i* receives credit through a microfinance loan at time *t*, and e_{it} is an error term. The coefficient δ represents the observed difference in the dependent variable between a household taking and a household not taking a microfinance loan, adjusted for these controls. Note that by utilizing a regional fixed effect, this difference incorporates differences between households with and without credit as well as differences over time in one household between periods when it has a loan and when it does not.

To compare what practitioners observe to what is more likely to reflect actual microfinance impact, we estimate Intention to Treat (ITT) effects using difference-in-differences that yield microfinance effects across the entire population of our sample. These estimations show the impact of the program, not just on borrowers, but on our entire sample, which involves those eligible for the JBS program.

We divide our data into three periods: period 1, period 2, and period 3. The first village in the regional pair to receive credit is referred to as Village A, the second Village B. In a given region, Period 1 is defined as the period beginning in 2001 in which neither of the two villages had access to microfinance and ending the year before Village A gained access to the JBS microfinance program. Period 2 begins in the first year, by region, when the A-Villages gained credit access and ends in the year before the B-Villages also gained credit access. Period 3 begins in the first year in which the B-Villages received the JBS program and ends in 2010, so that in the third period both villages had access to microfinance.

Thus we estimate the equation

$$y_{ijt} = \alpha_j + \boldsymbol{X_i'\beta} + \boldsymbol{P_t'\theta_t} + \delta T_{it} + e_{i,} \tag{11}$$

(where i now represents a period as opposed to a year) and y_{ijt} is the average annual probability of a fundamental event in a given period. In (11), α_j is a region fixed-effect, $\boldsymbol{X_i'}$ is our vector of controls, $\boldsymbol{P_t'}$ is the vector of (two) period dummies (where period 1 is omitted as the baseline) and T_i is equal to one if household i is a member of a village with microfinance access in the second or third period.

We also estimate a LATE (instrumented ATT) using two-stage least squares with first-stage

$$C_{it} = \alpha_j^1 + \boldsymbol{X_i'\beta^1} + \boldsymbol{P_t^{1\prime}\theta_t} + \varphi T_i + \varepsilon_i \tag{12}$$

and second stage

$$y_{ijt} = \alpha_j^2 + \boldsymbol{X_i'\beta^2} + \boldsymbol{P_t^{2\prime}\theta_t} + \tau \hat{C}_i + \epsilon_i \tag{12'}$$

by instrumenting for credit take-up C_{it} with credit availability T_{it}.[12] Estimation of the local average treatment effect yields an estimate of microfinance impact on "compliers," those who were induced by the existence of the program to undertake microfinance who would not have done so otherwise, and τ gives an estimate of the instrumented ATT. Since at this time microfinance was not available through any microfinance provider other than JBS, we would expect the instrumented ATT to provide reasonable estimates for the impact of the program more generally on microfinance borrowers.

4. DATA ANALYSIS AND RESULTS

(a) Event Study Results

The RETRAFECT methodology is useful exploring the dynamics of changes in impact variables over a number of years surrounding treatment. Under strong assumptions regarding the orthogonality of program placement with respect to impact variables, it is able to yield estimates of program impact. We formally lay out these assumptions in the Appendix. Equations (9) and (9') illustrate the dynamics of apparent impact and causal impact on fundamental household events. Results of these estimations are given in Tables 2A, 2B, and 2C. The five columns to the left show breakdowns for individual events associated with development for credit availability with the fifth column showing cumulative impacts on all four of the events for each of the three table categories (enterprise investment, housing improvements, and consumer goods). The sixth column on the right shows the cumulative impacts for credit take-up. Figures 1A – 3B show the dynamics of apparent impact (credit take-up) and causal impact (credit availability) in graphical form. We cluster standard errors of our estimated coefficients at the household level.

Table 2A shows the RETRAFECT estimations of the impact of credit availability on enterprise investment. Credit availability appears to display no significant impact on investments in machinery or on land purchases, but does significantly impact investment in livestock and new stores (new physical retail space). The livestock impact makes sense given that many of the loans were granted to rural households for herd expansion and to retailers also seeking to expand. These impacts are large in magnitude, beginning mainly two years after initial borrowing, and ranging from a 2.4 to 7.8 percentage point increase in the probability of a major investment in either of these two categories over a baseline of less than half a percentage point. The summary category for any enterprise investment is also large, increasing the probability of any major enterprise investment in the second, third, and fourth years after initial credit by 8.0, 7.1, and 8.5 percentage points, respectively.

These magnitudes are smaller than the increases in the probability of enterprise investment after credit *take-up*, which begin in the year the first microfinance loan is received and continue to the fourth year after credit, ranging from 9.2 to 13.8 percentage points. Given that JBS take-up was 51% in our sample, the apparent and causal impacts for enterprise investment are fairly similar, with the apparent impact simply occurring earlier in time than the causal impact. For both types of estimations, Table 2A gives *F*-statistics testing differences between estimated probabilities of enterprise investment before and after credit access/take-up that are highly significant. All estimations use region-year fixed effects and controls for age, age-squared and type of enterprise. The estimations on credit access are weighted so that borrowers in the data accurately represent their proportion in the random sample. Differences between the two RETRAFECT estimations can be seen visually in Figures 1A and 1B, where the dotted line represents a 95% confidence interval around each dynamic point estimate.

RETRAFECT estimations for housing improvements are given in Table 2B. In the credit access estimations, only one housing improvement variable is significant in one year: new walls in the first year after credit. Given that there are 16 such impact coefficients, the outcome probably reflects expected random variation in the estimated parameters. None of the other measures of home improvements show any level of significance in any year subsequent to the first year of credit access.

With the home improvement variables, we see the first real divergence between the apparent impact of microfinance and causal impact. The estimations around credit take-up show large increases in the probability of a home improvement subsequent to taking a microfinance loan, ranging from a 4.8 percentage point increase in the year the first microfinance loan was taken to 11.7 in the third year after taking a loan. The result is nearly identical with McIntosh et al.'s (2011) finding of large increases in home improvements subsequent to the taking of microfinance loans in Ghana, Guatemala, and India.

Apparent impact of microfinance on consumer durables overstates causal impact only slightly less than on home improvements. Whereas estimates of increases in the probability of a major consumer good range from 9.2 to 10.3 percentage points after the year of credit take-up, the RETRAFECT window around credit access shows only modest evidence of increases in the probability of bicycles, cell phones, and stoves around the second year after credit access. While the F-statistic testing for "before-and-after" differences in major consumer good purchases around credit take-up is significant at the 2% level, it is insignificant for credit access.

(b) Difference-in-difference Results

We show simple difference-in-difference outcomes for business investment, home improvements and consumer durables for periods 1 and 2 in Figures 4A, 4B, and 4C, respectively. In period 1 neither A-Villages nor B-Villages in each of the three regions had access to credit. In period 2, A-Villages had credit access, but not B-Villages. Results of t-tests of these raw difference-in-differences are given in the figures.

Figure 4A shows that while difference-in-differences seem to result in substantially larger increases in business investment (2.75 percentage points, baseline in Table 1= 2.70%) for the A-Villages with microfinance access, the difference is only marginally statistically significant (p-value 0.105). Similarly, the A-Villages with microfinance access also display higher rates of increase in home improvement over this period (2.36 percentage points, baseline in Table 1 = 3.00%), but the increase is statistically insignificant (p-value 0.128). In contrast, consumer durable purchases are significantly *lower* in A-Villages, increasing much faster among the three villages with delayed microfinance access (p-value 0.002) as shown in Figure 4C.

Table 3 presents three types of estimations:

(1) Take-up estimations, which show the differences in outcome variables that occur with microfinance borrowing and estimate equation (10). These isolate differences in impact variables associated with a household having a microfinance loan.

(2) Intention-to-treat (ITT) effects, that estimate equation (11) and show the impact of microfinance availability on all households in a village, who were all eligible for the JBS program, weighted to account for our over-sampling of borrowers.

(3) Estimates of the instrumented ATT in (12'), which show the impact of the program on "compliers" who took the program at any point during the respective period it was offered. These offer interesting comparisons to the "apparent impacts" given in the take-up estimations.

The instrumented ATT estimates use microfinance access as an instrument for take-up. Because of the nature of JBS program rollout, where programs were introduced by village based on a transparent geographical expansion strategy rather than by village economic factors, the instrument satisfies the exclusion restriction provided that this geographical expansion strategy is independent of differences in economic outcomes and trends between A-Villages and B-Villages. Note that in the presence of externalities to microfinance investment, the SUTVA assumption would not hold and any positive spillovers from microfinance borrowing are imputed from non-borrowers to the treated such that while these impacts exist, the treatment effect actually accruing to the borrower would be lower. However, we feel reasonably comfortable that microfinance externalities are small if they exist at all. The largest controlled study carried out in a virgin microfinance area, Crépon et al. (2013), for example, tests for and finds no empirical evidence for microfinance externalities.

We test for similarity in pre-treatment outcomes between villages as well as pre-treatment parallel trends in our three major categories of impact variables: business investment, home improvements, and consumer durable purchases. We find no significant differences between A-Villages and B-Villages in pre-treatment (2001-2004) outcomes among the three aggregated impact variables (*t*-statistics = 0.84, 0.23, -1.02, respectively). We also test for pre-treatment parallel trends, differences in trends between A- and B-Villages between 2001-02 and 2003-04. Two of these three tests (business investment and consumer durables) yield no significant results (*t*-statistics = 0.75, -1.61, respectively), but the test on home improvements shows a significantly higher trend in the B-Villages during this time ($t = 2.64$), a result is driven by an abnormal spike in home improvement in 2003 in Bahundangi of 0.171, which reverted back to 0.014 in 2004. If this, however, constitutes part of a true underlying trend, it would tend to bias our results for home improvement *against* rejecting a null hypothesis in favor of positive impact from microcredit, since the pretreatment "trend" would be greater in the B-Villages.

The estimations in Table 3 show that practitioner observations considerably overstate the true impacts of microfinance, but that microfinance still appears to have significant impacts. These are limited mainly to business investment, but the impact on the probability of a new business investment in our Nepal study is substantial. Although the apparent impact on business investment--differences in the probability of a major capital investment between when a borrower does and does not have a microfinance loan--is nearly 40% larger than what is measured by the instrumented ATT for business investment, the causal effects estimated by the ITT and the instrumented ATT are large, with percentage point increases of 2.5 and 10.8, respectively over a baseline probability of a 2.7 percent chance of any type of major investment in a given year. These are larger effects than have been reported in most recent experimental and quasi-experimental studies, but not necessarily larger than what we might expect in a region in which microfinance has been newly introduced and where take-up is widespread.

Similarly, the apparent affect is over 50% larger for home improvements than the instrumented ATT estimates. For consumer durables, the apparent effect is enormous (12.7 percentage points) while the instrumented ATT estimate is strongly negative (-20.4 percentage points). Over our 12 non-aggregated impact variables, apparent impact is significant for 11 of these variables, while the local average treatment effect is significant for only two: new store capital and new livestock, and only at the 5% and 10% levels respectively, although the coefficients indicate a large 9.4 percentage point increase in the probability of new livestock investments and a 3.3 percentage point increase in the probability of new store capital.

We cluster standard errors at the household level in these estimations. This seems to be the most reasonable approach to us given the nature of our data, where the strongest correlation the residuals is most likely to be within a given household over time. As a robustness check, however, we also cluster standard errors at the village level (Bertrand, Duflo, and Mullainathan, 2008), but clustering is problematic due to the small number of villages (six). The solution to this problem is to implement the Cameron, Gelbach, and Miller (2008) wild bootstrap which allows for clustering with small for a smaller number of clustering units (<40). The wild bootstrap resamples using a multiplicative correction to residuals in the context of cluster-level heteroskedasticity to allow for unbiased estimates of standard errors. As seen in Table 3, after carrying out the wild bootstrap and clustering at the village level, significance for business investment on microfinance take-up remains under the wild bootstrap estimation and significance is retained for capital investment in store investment, however it falls to insignificance for investment in new livestock. Thus, using this alternative method of clustering, the evidence falls slightly more toward the lower levels of impact noted by researchers and away from that claimed by practitioners.

Under either type of clustering procedure, we find positive microfinance impact results that appear to exceed those estimated by the slew of recent randomized evaluations carried out in places with high levels of existing microfinance lending (Karlan and Zinman, 2011; Angelucci et al., 2012; Attanasio et al., 2011; Augsburg et al., 2012; Banerjee et al., 2013.) But there are strong parallels between our findings and those of Crépon et al. (2013) in Morocco, where microfinance was absent prior to the studied credit intervention. Like many of these experimental papers, we find positive impacts of microfinance borrowing on business expansion but small or negative impacts on consumer spending. Our instrumented ATT impact estimates indicate a point estimate of 10.8 percentage point increase in the probability of some kind of business investment. This comes from significant increases in livestock (when standard errors are clustered at the household level) and physical investment in store buildings, but not significantly from investments in land and machines, although both of these have positive point estimates.

Like many of the recent studies on microfinance based on randomized evaluations, we also find a reduction in consumer-oriented goods, which similarly appear to be sacrificed in favor of business expansion, but this cannot be verified with certainty. Nearly all of these recent studies cited above find reductions in "temptation goods" after microfinance borrowing (goods which households indicate they would like to spend less on such as alcohol and tobacco), and similarly we find a negative impact on television purchases in our data—the probability of purchasing a television falls dramatically by 14.5 percentage points—which parallels a movement toward enterprise investment. Thus our findings are similarly consistent with the idea that microfinance fosters a movement away from "temptation" goods and toward "efficiency" goods. Our difference-in-difference findings diverge slightly from the Crépon et al. (2013) study in that we also find some very modest evidence for a positive impact on home improvements in that while none of our impact variables (improvements in roof, wall, floor, and toilet) are statistically significant, all except improved floor carry positive coefficients.

(c) Other Robustness Checks

Our methodology uses recall data to construct a backcast panel data set of fundamental events to the household. Clearly it is of interest to know how robust these results are to the accuracy of the timing of these events. While there is little doubt that many of the events in question took place (home improvements, purchase of certain consumer durables, business investment since their existence in the present can be quickly verified; the main question of interest is how robust the results are to inaccuracy in the timing of the events.

To carry out robustness checks for recall accuracy, we generate random (0,1) normal distribution z-scores for every observation in the data set to represent departures from perfect recall of the timing of the fundamental event. We assume that recall over the particular year a fundamental event took place is distributed normally with variance σ^2. We then create a recall parameter, ρ, which is equal to standard deviations per year away from the true year of the event in the recall distribution. A smaller ρ would thus represent less accurate memory in pinpointing the year a given event took place; larger ρ would represent more accurate recall. For example, with $\rho = 1$, a simulated random z-score for an observation where $z \in [-1, 1]$ means that the year of the event was recalled with accuracy, $z \in [-2, -1)$ means that the event was recalled one year earlier than the event actually occurred, $z \in [-3, -2)$ means that the event was recalled two years earlier than it actually occurred. With $z \in (1, 2]$ the event was recalled one year after it actually occurred, and so forth. A lower value of ρ indicates poorer quality recall so that, for example, with $\rho = 0.50$, a simulated random z-score for an observation where $z \in (1.5, 2]$ would mean that the year of the event was recalled as occurring three years after it actually occurred.

We focus on our most significant positive result, the impact of microfinance on business investment in our sample, considering first our event-study estimations. Consider the nine-year window results in Table 2A, where the F-statistic on the significance of post- to pre-microfinance years is 7.96, with p-value =0.0048. At recall error $\rho = 1$, the F-statistic falls to 6.52, but impact on business investment is still strongly significant ($p = 0.0124$). At our previous example of $\rho = 0.50$, the F-statistic falls to 0.67 ($p = 0.4131$) and significance disappears entirely. The borderline cases when the quality of recall is just large enough to yield significance are at $\rho = 0.75$, which yields an F-statistic of 2.74 ($p = 0.097$), making the impact on business investment marginally significant at the 10% level, and $\rho = 0.82$, which yields an F-statistic of 3.86 ($p = 0.049$), where the impact is marginally significant at the 5% level.[13] Thus in our borderline 5-percent-significance case, 58.8% of our subjects would recall the event year with accuracy, 15.5 percent each would mistakenly recall the event one year too closely to the present and one year too far into the past, and 4.5 percent each would recall the event two years too closely to the present and two years too far into the past, and about 0.7 percent each would recall the event three years too closely to the present and three years too far into the past.[14] Because *p*-values are relatively low for our difference-in-difference estimations, the robustness of our significant results for our ITT and instrumented ATT for business investment are more sensitive, reaching a 10% significance level at ρ = 1.96 and 1.94, meaning that the statistical significance of the results are sensitive to more than about 5% of the subjects failing to recall the correct year in which business investment took place.

Because we have multiple hypotheses, a further robustness check controls the familywise error rate over the family of hypotheses in our paper. Using the Holm-Bonferroni step-down procedure (see Holm, 1979) at α =0.10 across out three summary indices, where we establish a ranking of *m* *p*-values indexed by $p(k)$ in decreasing order of magnitude: $p(1), p(2), \ldots p(m)$,. The null hypothesis rejection criterion is to reject all null hypothesis for which $p(k) < \frac{\alpha}{m-k^*+1}$, where k^* corresponds to the marginal *p*-value (hypothesis) that satisfies this criterion. For our RETRAFECT estimations, seven out of our ten significant post-credit take-up coefficients retain significance for our summary indices on investment, dwelling improvements, and consumer goods, while we are no longer able to reject three of our null hypotheses, and we are unable to reject the null in any of our ITT estimations, although business investment is close to marginally significant. For difference-in-difference estimations, we find that each of our take-up estimations is robust to controlling the familywise error rate, but that because our statistical significance for new investments is only marginally significant at the 10% level, we cannot reject the null hypothesis of no impact after we control the familywise error rate. Thus, in general we find that our robustness checks push our RETRAFECT results closer to the weaker microfinance impacts found in recent well-known experimental studies.

5. CONCLUSION

Our paper explores the impact of institutional microcredit on the improvement of household welfare of the poor in rural Nepal, while attempting to explain the paradox between the high levels of microfinance impact reported anecdotally by practitioners and the low levels of impact reported by recent experimental studies.

An empirical shortcoming of our study is the relatively small number of villages (six) where credit was newly introduced to sequential pairs of highly similar villages never having previous access to microfinance. Although implementation of the wild bootstrap yields an unbiased estimation of standard errors in the presence of a small number of clusters, it still means that statistical power in our study--to reject false null hypotheses of no microfinance impact--is lower than ideal. Nevertheless the data from these six villages are unique for three reasons: First, we have reason to believe that credit was allocated to new villages in a manner orthogonal to changes in welfare variables over time. Second, the credit introduced by JBS was the first microcredit available to the population in our study area. Third, credit take-up was extremely high among eligible households after it was introduced, far higher than the differential rates induced by recent experimental designs. So while the data from our study comes from a small number of villages, the unique data from these villages helps us to avoid some of the pitfalls that plague many recent studies that have sought to analyze the impact of microfinance, but in areas in which microfinance had already been introduced.

We demonstrate in a simple model why the before-and-after observations of practitioners *overstate* the average treatment effect of microfinance, and why the results from recent studies of microfinance—those carried out where microfinance pre-existed experimental interventions--*underestimate* average treatment effects.

While our empirical results appear to reveal greater microfinance impacts than most recent randomized evaluations, they display far smaller and more limited impacts than are often claimed in the microfinance industry. Many of these impacts claimed in the microfinance industry are derived from the difference in the lives of borrowers after taking microfinance loans. There are many ways one could measure this overstatement of impact. Our results indicate that of the 13 post-credit coefficients that are significant in our RETRAFECT estimations for apparent impact (credit take-up), only 4 are significant in our causal impact estimations related to credit availability. In our difference-in-difference estimations, out of 14 coefficients significant for apparent impact, only 3 are significant in instrumented ATT estimates, and only one (impact on investment in rural enterprise capital) under the wild bootstrap robustness check. Thus about 75% these coefficients indicating apparent impact do not show actual impact.

Moreover, if we look at all 12 of our individual dependent (impact) variables in terms of their point-estimated magnitudes of impact in Table 3, the *median overstatement* of apparent impact is also close to 75% (where impact on new walls is overstated by 75%, and on new cell phones by 73%.) Thus based on a consensus of all of these criteria in our Nepal data, we conclude that approximately 3/4 of *apparent* microfinance impact to be an illusion driven by correlated unobservables.

Nevertheless, of the true causal impact that remains, we do find evidence of positive and significant effects from microfinance provision in Nepal: The substantial increase in enterprise expansion we observe in our data is an encouraging impact result from a microcredit program which provides loans that are quite small (only US$192 on average) in these rural areas of Nepal, where agriculture and small rural enterprises are the major income-generating activities.

REFERENCES

Amendariz, B. & Morduch , J. (2005). Measuring impacts. *Economics of microfinance* (pp. 207-229). The MIT Press: Cambridge, Massachusetts.

Angelucci, M., D. Karlan, & J. Zinman (2012). Win some lose some? Evidence from a randomized microcredit program placement experiment by Compartamos Banco. J-PAL working paper.

Attanasio, O., B. Augsburg, R. De Haas, E. Fitzsimons, & H. Harmgart (2011). Group lending or individual lending? Evidence from a randomised field experiment in Mongolia. Pub ref: MPRA Paper No. 35439.

Augsburg, B., R. D. Haas, H. Harmgart, & C. Meghir (2012). Microfinance, poverty and education. IFS working paper.

Banerjee, A., Duflo, E., Glennerster, R. & Kinnan, C. (2013). The miracle of microfinance? Evidence from randomized evaluation. *Centre for Micro Finance, IFMR Research*, Working Paper Series No. 31.

Becchetti, L. & S. Castriota (2011). Does microfinance work as a recovery tool after disasters? Evidence from the 2004 tsunami. *World Development*, 39(6), 898–912.

Berhane, G. & Gardebroek, C. (2011). Does microfinance reduce rural poverty? evidence based on household panel data from northern Ethiopia. *American Journal of Agricultural Economics, forthcoming.*

Bertrand, M., E. Duflo, & S. Mullainathan. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119 (1), 249.275.

Bhalotra, S., Copestake, J., & S. Johnson. (2001). Assessing the Impact of Microcredit: A Zambian Study. *The Journal of Development Studies*, 37(4), 81-100.

CGAP (2002). Microfinance and the millennium development goals, CGAP donor brief No. 9, December 2002, CGAP, http://www.cgap.org/gm/document-1.9.2422/DonorBrief_09.pdf.

CGAP (2003). Helping to improve donor effectiveness: the Impact of Microfinance. CGAP donor brief No. 13.

Cameron, A. C., J. B. Gelbach, & D. L. Miller. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90 (3), 414-427.

Chamberlain, G. (1979). Analysis of covariance with qualitative data. *National Bureau of Economic Research, Working Paper No. 325.*

Coleman, B. (2002). Microfinance in northeast Thailand: Who benefits and how much? ADB, ERD working paper series no. 9.

Conway, M.A., & D.A. Bekerian. (1987). Organization in autobiographical memory. *Memory and Cognition*, 15(2): 119.

Crépon, B., F. Devoto, E. Duflo & W. Parienté (2011). Impact of microcredit in rural areas of Morocco: Evidence from a randomized evaluation" MIT/JPAL working paper.

Dingcong, C., Infantado, C., Kondo, T. & A. Orbeta Jr. (2008). Impact of microfinance on rural households in the Philippines. *Philippines Institute for Development Studies.* Discussion Paper series no. 2008.05.

Dizon, F. (2009). Rural out-migration and farm asset accumulation: the case of Guatemala. (Master's thesis and presented at the Pacific Development Conference, March 2009)

Duflo, E., Glennerster, R. & M. Kremer. (2007). Using randomization in development economics research: a tool kit. *Centre for Economic Policy Research*, Discussion paper series no. 6059.

Field, E., R. Pande, J. Papp, & N. Rigol (2013). Does the classic microfinance model discourage entrepreneurship among the poor? Experimental evidence from india." *American Economic Review*, 103(6): 2196-2226.

Giné, Xavier & Francesca de Nicola (2012). How accurate is recall data: Evidence from coastal India. World Bank Policy Research paper #6009.

Giné, Xavier, & Ghazala Mansuri (2011). Money or ideas? A field experiment on constraints to entrepreneurship in rural Pakistan. World Bank Working Paper.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6: 65-70.

Imai, K., G. Raghav, T. Ganesh, and S.K. Annim (2012). Microfinance and poverty: A macro perspective. *World Development*, 40 (8), 1675–1689

Karlan D. (2001). Microfinance impact assessments: the perils of using new members as a control group. *Journal of Microfinance*, 3(2), 75-85.

Karlan, D. & Zinman, J. (2011). Microcredit in theory and practice: using randomized credit scoring for impact evaluation, *Science* 10(332), 1278-1284

Khandker, S. R. (2005). Microfinance and poverty: evidence using panel data from Bangladesh, *World Bank Economic Review*, 19(2): 263-286.

Khandker, S. R., & Pitt M. M. (1998). The impact of group-based credit programs on poor households in Bangladesh: does the gender of participants matter? *Journal of Political Economy*, 106(2), 958-996.

MacInlay, C. (1997). Event studies in economics and finance. *Journal of Economic Literature*, 35, 13-39.

McIntosh C., Villaran G., & Wydick, B. (2011). Microfinance and home improvement: using retrospective panel data to measure program effects on fundamental events. (Symposium on Microfinance) *World Development*, 39(6), *922-37*.

Microcredit Summit, State of the Microcredit Summit Campaign Report 2014. Stable URL: http://stateofthecampaign.org/data-reported/

Morduch, J. (1998). Does microfinance really help the poor? New evidence from flagship programs in Bangladesh. NYU Wagner Graduate School working paper.

Morduch, J. & Roodman, D. (2009). The impact of microcredit on the poor in Bangladesh: revisiting the evidence. Working paper no. 174, Centre for Global Development.

Rosenberg, R. (2010). Does microcredit really help poor people? CGAP Focus Note 59.

Rubin, D. (1974). Estimating the causal effects of treatments in randomized and non-randomized studies. *Journal of Education Psychology* 66, 688-701.

Rubin, D. (1977). Assignment to a treatment group on the basis of a covariate. *Journal of Educational Statistics* 2, 1-26.

Stiglitz, J.E. & Weiss, A. (1981). Credit rationing in markets with imperfect information. *American Economic Review*, vol. 71, no. 3, pp. 393-410.

White, H. (2006). Time Series Estimation of the Effects of Natural Experiments. *Journal of Econometrics*, 135, 527-566.

FOOTNOTES

[1] Microcredit Summit, 2014. Data reported are as of December 31, 2012.

[2] Field et al. (2013) find statistically significant impacts on profits in a treatment group in which microfinance borrowers were given a grace period before their initial interest payment at the expense of higher default rates.

[3] To avoid notational clutter, we will generally suppress the i subscript notation as all of our arguments apply equally to comparing a single household across time as to comparing across households at a single period in time.

[4] Clearly there loans can be triggered by events that are unrelated to a positive shock in productivity of an enterprise, however an important feature of our model demonstrates that it is rational for borrowers to take loans when these positive opportunity shocks do occur, and that this may lead to overstatement of program impact.

[5] The websites of leading microfinance institutions contain many stories intended to illustrate the impact of microfinance loans. For example, a story that appears on the website of Opportunity International relates that "Rayusa Muzalila of Kiganda, Uganda, has used Opportunity loans to expand her grocery and start a textile business. With her increased profits, she has opened her first-ever savings account…" A similar story on the ACCION website reads that "(After a series of microloans) Juan now has a full workshop of tools, including a modern table saw. Employing several neighborhood boys part-time, he can now turn out one couch a day…"

[6] Terai is southern plain area of Nepal where climate is sub-tropical and population density is high as compared to Northern hilly region of Nepal.

[7] The MFI's portfolio status as of mid-July 2010 included outstanding loans of NRs. 553 million (US$7.7 million approx.), 46,914 total borrowers , 60,623 individuals with outstanding savings of NRs. 227.5 million (US$3.16 million), a loan recovery rate of just under 100%, operational self-sufficiency ratio of 155%, and a financial self-sufficiency ratio of 125%.

[8] A main feature of Grameen Classical System is group liability in repayment of individual loans. Grameen Bank, Bangladesh has moved to Grameen Generalized System in 1992 from its Grameen Classical System.

[9] A village development committee (VDC) is a political area consists of many small clusters (villages).

[10] The year in parenthesis indicates the year when microfinance program was rolled-out in the respective VDC.

[11] Retail trade includes petty trading of vegetables, groceries, grains, livestock.

[12] IV-first stage results for credit availability as an instrument for credit take-up show an F-statistic equal to 95.3 in the first-stage, indicating credit-availability (not surprisingly) to be strong instrument for credit take-up.

[13] Five-year window results are robust at the 5-percent level to $\rho = 1.70$.

[14] In Giné and Francesca's (2012) data on 32 observations of boat purchases, they find that coefficient of variation for boat purchase recall data (s.d. =16.7 months / mean = 63.3 months) that is only 23% higher than that for administrative data, but that subjects appear to negatively telescope purchases about six months too far into the past. If a similar sized negative telescoping phenomenon were to be present in our data, it would leave impact results unaffected because the significant effects we estimate on business investment (see Table 2A) occur two years after the implementation of the microfinance program.

Table 1: Descriptive Statistics

	Sample Statistics		Baseline Statistics		Two Years Before Rollout in first village of village pair		
	Mean	Std. Dv.	Mean	Std. Dv.	Mean 1st Village	Mean 2nd Village	*t*-test
Dependent variables							
Roof	0.032	0.176	0.020	0.139	0.020	0.0165	0.467
Wall	0.005	0.068	0.002	0.050	0.002	0.000	0.926
Floor	0.007	0.084	0.004	0.062	0.003	0.007	1.071
Toilet	0.014	0.118	0.008	0.088	0.007	0.142	1.275
Home improvement	0.048	0.214	0.030	0.169	0.027	0.035	0.817
Television	0.041	0.199	0.027	0.162	0.031	0.009	2.433
Bicycle	0.062	0.241	0.036	0.187	0.049	0.018	2.692
Cell phone	0.052	0.222	0.009	0.092	0.007	0.004	0.511
Stove	0.007	0.081	0.003	0.056	0.006	0.002	0.910
Consumer durables	0.133	0.339	0.067	0.250	0.083	0.0332	3.423
Livestock	0.049	0.215	0.016	0.126	0.020	0.014	0.778
Machine	0.007	0.085	0.003	0.053	0.004	0.0023	0.489
Store	0.007	0.085	0.003	0.056	0.005	0.000	1.467
Net-land purchase	0.015	0.120	0.005	0.070	0.000	0.007	2.652
Business Investment	0.073	0.261	0.027	0.161	0.029	0.023	0.604
Control variables							
Age	39.920	12.239					
Education*	1.498	0.644					
Agriculture	0.259	0.438					
Retail trade	0.543	0.498					
Agriculture and retail	0.198	0.398					

*illiterate =1, primary=2, high school=3

Table 2A : OLS RETRAFECT Regression on Business Investment

	--- Credit Availability Window ---					Uptake Window
VARIABLES	New Machine	New land	New livestock	New Store	Any new investment	Any new investment
4 years before	-0.001	0.004	-0.020*	0.002	-0.017	-0.004
	(0.003)	(0.007)	(0.012)	(0.003)	(0.014)	(0.008)
3 years before	-0.001	0.002	-0.012	0.001	-0.010	0.010
	(0.003)	(0.006)	(0.010)	(0.002)	(0.012)	(0.009)
2 years before	0.003	0.000	-0.007	-0.000	-0.007	-0.003
	(0.005)	(0.008)	(0.015)	(0.004)	(0.017)	(0.011)
1 year before	0.001	-0.008	-0.007	0.002	-0.008	0.004
	(0.006)	(0.010)	(0.016)	(0.004)	(0.020)	(0.013)
Credit access yr	0.006	-0.005	0.014	0.012*	0.020	0.110***
	(0.007)	(0.012)	(0.020)	(0.006)	(0.025)	(0.020)
1 year after	-0.010	-0.003	-0.001	0.001	-0.010	0.092***
	(0.009)	(0.016)	(0.024)	(0.007)	(0.030)	(0.020)
2 years after	0.007	0.003	0.057**	0.024***	0.080**	0.138***
	(0.007)	(0.016)	(0.027)	(0.007)	(0.032)	(0.026)
3 years after	-0.000	0.014	0.049*	0.016**	0.071**	0.106***
	(0.008)	(0.018)	(0.027)	(0.008)	(0.033)	(0.032)
4 years after	0.003	-0.024	0.078**	0.031***	0.085*	0.129***
	(0.014)	(0.024)	(0.038)	(0.008)	(0.044)	(0.044)
education	0.004**	0.003	0.002	0.002	0.010**	0.005
	(0.002)	(0.002)	(0.004)	(0.001)	(0.005)	(0.005)
age	0.000	-0.000	0.002**	0.000	0.002**	0.002
	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
age-squared	-0.000	0.000	-0.000**	-0.000	-0.000*	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
retail trade	-0.002	0.001	-0.019***	0.005***	-0.012*	-0.010
	(0.003)	(0.003)	(0.006)	(0.002)	(0.007)	(0.007)
agric*retail	-0.001	0.006	0.018**	0.006**	0.029***	0.023**
	(0.004)	(0.005)	(0.008)	(0.003)	(0.010)	(0.010)
Constant	-0.013	0.046**	0.018	-0.008	0.051	0.050
	(0.010)	(0.020)	(0.027)	(0.009)	(0.035)	(0.034)
Observations	7,020	7,020	7,020	7,020	7,020	7,020
R-squared	0.014	0.027	0.051	0.012	0.070	0.024
F-stat: post- vs. pre-treatment	0.000	0.04	9.02***	10.93***	78.23***	39.55***
p-value	0.95	0.853	0.003	0.001	0.004	0.000

Standard errors clustered at the household level in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 2B: Fixed Effect OLS RETRAFECT Regression on Housing Improvement

	--- Credit Availability Window ---					Uptake Window
VARIABLES	New roof	New wall	New floor	New toilet	Any home improvement	Any home improvement
4 years before	-0.001	0.005*	0.002	-0.016***	-0.015*	-0.007
	(0.006)	(0.002)	(0.003)	(0.006)	(0.009)	(0.009)
3 years before	-0.000	0.007*	-0.000	-0.036**	-0.033*	-0.018**
	(0.011)	(0.004)	(0.006)	(0.014)	(0.019)	(0.009)
2 years before	-0.005	0.010**	0.003	-0.028***	-0.033*	-0.012
	(0.013)	(0.005)	(0.007)	(0.010)	(0.017)	(0.010)
1 year before	-0.015	0.008	-0.002	-0.045***	-0.062***	-0.012
	(0.015)	(0.006)	(0.009)	(0.017)	(0.023)	(0.011)
Credit access year	-0.020	0.007	-0.007	-0.035**	-0.056**	0.048***
	(0.017)	(0.008)	(0.010)	(0.017)	(0.023)	(0.015)
1 year after	-0.012	0.014	-0.007	-0.029*	-0.044*	0.073***
	(0.018)	(0.009)	(0.012)	(0.015)	(0.024)	(0.017)
2 years after	0.016	0.006	-0.008	-0.011	0.005	0.052***
	(0.019)	(0.010)	(0.012)	(0.018)	(0.026)	(0.018)
3 years after	0.035**	0.004	-0.007	0.002	0.030	0.117***
	(0.017)	(0.011)	(0.012)	(0.015)	(0.022)	(0.028)
4 years after	0.036	0.007	-0.006	0.008	0.043	0.036
	(0.027)	(0.010)	(0.012)	(0.019)	(0.033)	(0.027)
education	-0.005*	0.001	0.006***	0.012***	0.008	0.005
	(0.003)	(0.001)	(0.002)	(0.004)	(0.005)	(0.005)
age	0.001	-0.000	0.000	0.002**	0.002**	0.001
	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
age-squared	-0.000*	0.000	-0.000	-0.000**	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
retail trade	-0.003	-0.004*	-0.002	-0.005	-0.007	-0.005
	(0.005)	(0.002)	(0.002)	(0.005)	(0.006)	(0.006)
agric*retail	0.007	-0.001	0.005	0.001	0.007	0.007
	(0.006)	(0.003)	(0.004)	(0.007)	(0.009)	(0.009)
Constant	-0.014	0.004	-0.016*	-0.039**	-0.047*	-0.034
	(0.018)	(0.012)	(0.010)	(0.019)	(0.027)	(0.027)
Observations	7,020	7,020	7,020	7,020	7,020	7,020
R-squared	0.027	0.011	0.018	0.039	0.042	0.014
F-stat: post- vs. pre-treatment	2.35	0.01	0.55	6.31**	5.51**	15.4***
p-value	0.125	0.939	0.459	0.012	0.019	0.001

Standard errors clustered at the household level in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 2C: FE OLS RETRAFECT Regression on Consumer Durables Purchase

	--- Credit Availability Window ---					Uptake Window
VARIABLES	New television	New bicycle	New cellphone	New stove	Any Consumer Durable	Any Consumer Durable
4 years before	0.008	-0.002	-0.001	0.002	0.008	0.004
	(0.010)	(0.012)	(0.005)	(0.002)	(0.014)	(0.012)
3 years before	0.004	-0.010	-0.002	0.000	0.003	0.025*
	(0.013)	(0.013)	(0.007)	(0.002)	(0.018)	(0.015)
2 years before	0.044***	0.004	-0.000	0.006	0.040*	0.023
	(0.015)	(0.017)	(0.010)	(0.004)	(0.023)	(0.016)
1 year before	0.034*	0.040**	0.002	0.005	0.079***	0.031
	(0.017)	(0.019)	(0.014)	(0.005)	(0.027)	(0.019)
Credit access yr	0.015	0.002	-0.009	0.003	0.002	0.092***
	(0.022)	(0.023)	(0.018)	(0.007)	(0.032)	(0.023)
1 year after	0.005	0.014	0.024	0.011	0.040	0.088***
	(0.023)	(0.025)	(0.023)	(0.008)	(0.036)	(0.024)
2 years after	0.001	0.072**	0.059**	0.017**	0.115***	0.103***
	(0.027)	(0.030)	(0.026)	(0.007)	(0.041)	(0.029)
3 years after	-0.024	0.029	0.016	0.002	0.020	0.092**
	(0.024)	(0.027)	(0.025)	(0.007)	(0.037)	(0.036)
4 years after	-0.029	0.050	0.015	-0.012	0.006	0.103**
	(0.037)	(0.044)	(0.045)	(0.021)	(0.064)	(0.049)
education	0.019***	0.005	0.021***	0.004*	0.034***	0.029***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.007)	(0.006)
age	-0.001	-0.001	-0.001	-0.000	-0.003	-0.004**
	(0.001)	(0.001)	(0.001)	(0.000)	(0.002)	(0.002)
age-squared	0.000	0.000	0.000	0.000	0.000	0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
retail trade	-0.003	-0.006	-0.003	0.001	-0.009	-0.006
	(0.005)	(0.005)	(0.005)	(0.002)	(0.010)	(0.009)
agric*retail	0.010	0.016**	0.007	0.002	0.019	0.015
	(0.006)	(0.007)	(0.007)	(0.003)	(0.013)	(0.012)
Constant	0.007	0.092***	-0.005	-0.008	0.086*	0.108**
	(0.023)	(0.030)	(0.020)	(0.011)	(0.045)	(0.043)
Observations	7,020	7,020	7,020	7,020	7,020	7,020
R-squared	0.031	0.043	0.093	0.019	0.113	0.013
F-stat: post- vs. pre-treatment	2.98*	2.43	2.02	0.03	0.18	5.63**
p-value	0.085	0.119	0.156	0.853	0.671	0.025

Standard errors clustered at the household level in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Difference-in-Difference Estimations

Take-Up, ITT, and instrumented ATT Estimations for
Enterprise Investment, Housing Improvements, and Consumer Durables

VARIABLES	New machines	New lands	New livestock	New store	New investments
Microfinance Take-up	0.010**	0.037***	0.095***	0.020***	0.149***
	(0.005)	(0.006)	(0.011)	(0.005)	(0.013)
Microfinance ITT	0.001	0.002	0.022*†	0.008**	0.025*†
	(0.006)	(0.005)	(0.013)	(0.003)	(0.015)
Microfinance instrumented ATT	0.002	0.007	0.094*†	0.033**	0.108*†
	(0.025)	(0.022)	(0.056)	(0.013)	(0.065)

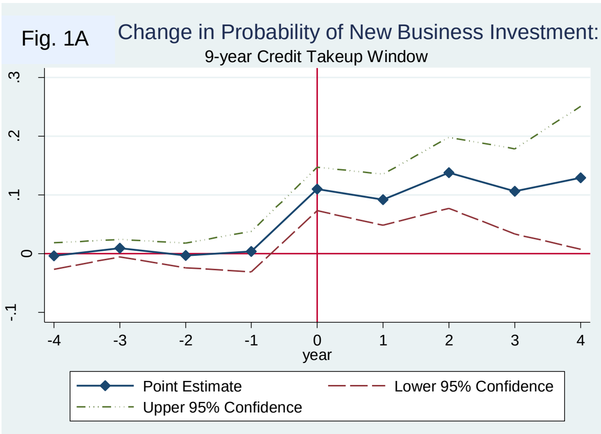
VARIABLES	New roof	New wall	New floor	New toilet	All Home Improvements
Microfinance Take-up	0.046***	0.004	0.010**	0.022***	0.071***
	(0.008)	(0.003)	(0.004)	(0.007)	(0.011)
Microfinance ITT	0.009	0.000	-0.008	0.005	0.011
	(0.011)	(0.004)	(0.007)	(0.007)	(0.014)
Microfinance instrumented ATT	0.039	0.001	-0.034	0.023	0.047
	(0.049)	(0.019)	(0.031)	(0.031)	(0.059)

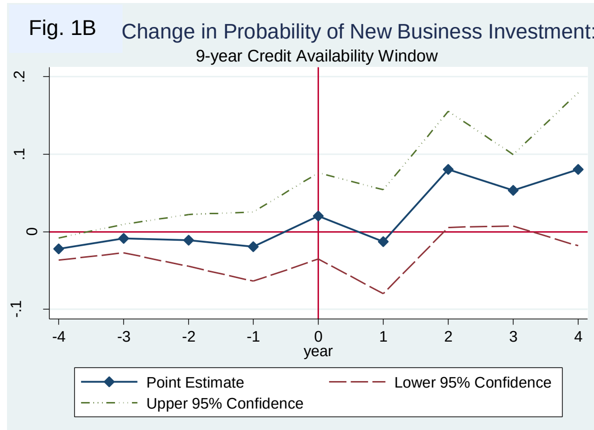
VARIABLES	New televisions	New bicycles	New cell phone	New stove	All Consumer Durables
Microfinance Take-up	0.027***†	0.061***	0.084***	0.009**†	0.127***
	(0.009)	(0.010)	(0.011)	(0.004)	(0.015)
Microfinance ITT	-0.033**	-0.019	0.005	0.003	-0.047**
	(0.014)	(0.015)	(0.011)	(0.004)	(0.023)
Microfinance instrumented ATT	-0.145**	-0.085	0.023	0.011	-0.204**
	(0.061)	(0.066)	(0.048)	(0.018)	(0.100)

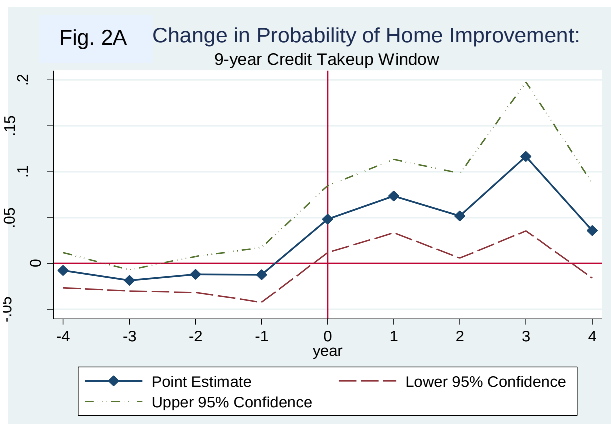
Each cell gives a single impact coefficient. Regressions include fixed effects at the region level and controls for education age, age^2, and type of enterprise. Clustered-robust standard errors at the household level in parentheses. Credit availability instruments for take-up in instrumented ATT estimations (First-stage F-statistic = 95.33, p-value < 0.001).
***$p < 0.01$, ** $p < 0.05$, * $p < 0.10$. †$p > 0.10$ using standard errors clustered with wild bootstrap.

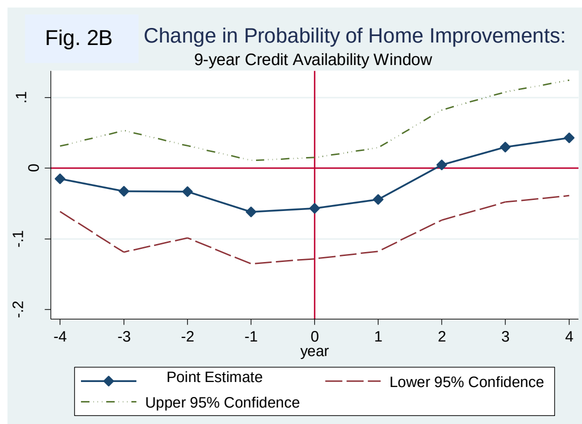
Figure 1: Map of Six Villages and Survey Site

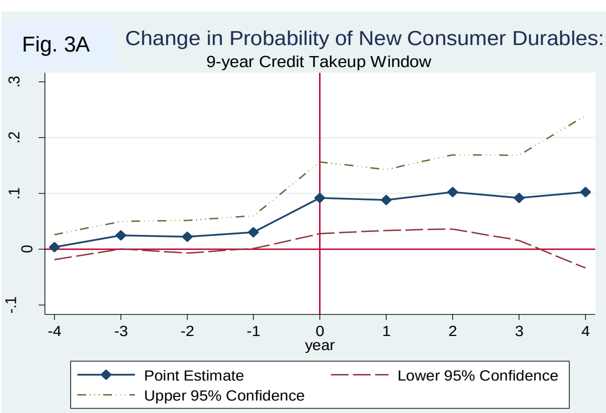
Figures 1A-3B: Changes in Impact Variables, Before and After Take-Up and Credit Availability











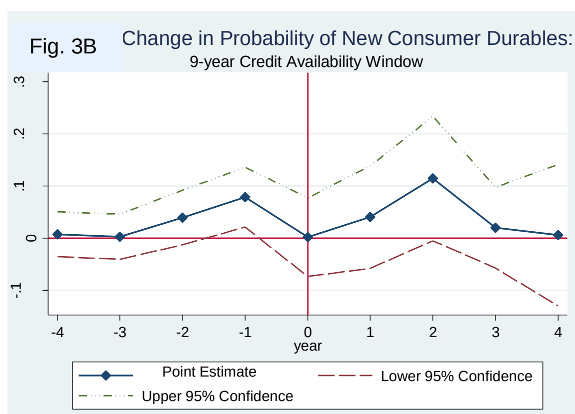


Figure 4A

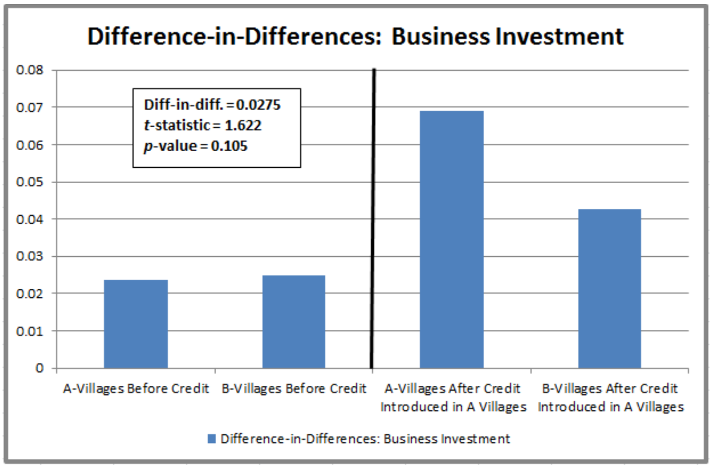


Figure 4B

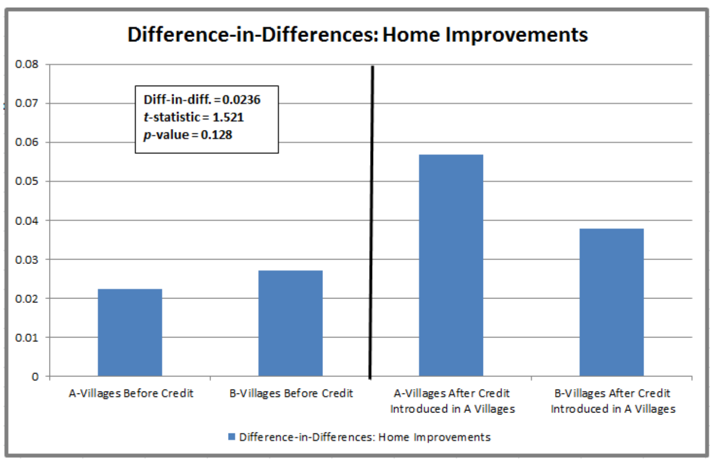
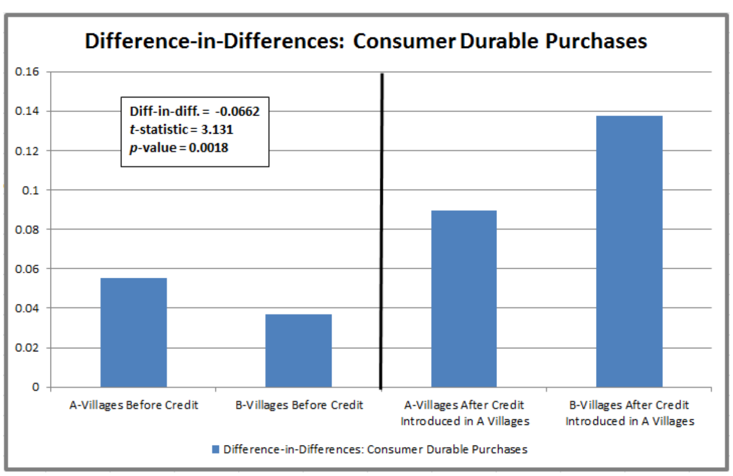
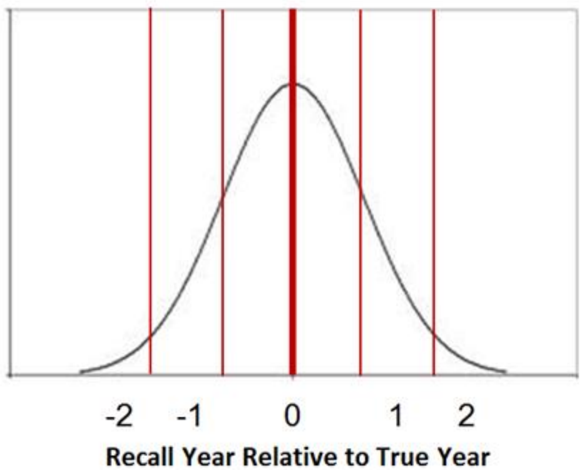


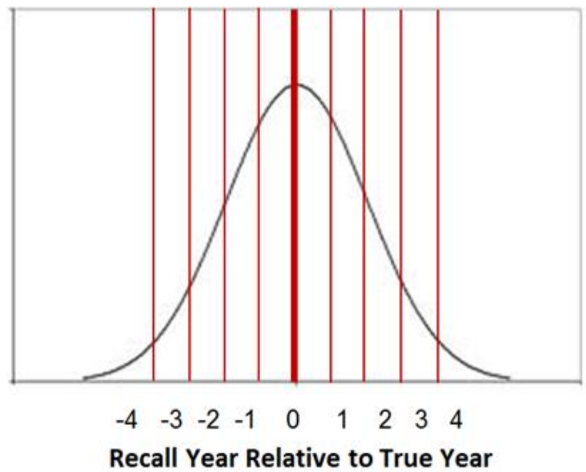
Figure 4C





Recall Timing Error, $\rho = 1$

Figure 5A



Recall Timing Error, $\rho = 0.50$

Figure 5B

APPENDIX: CAUSAL INTERPRETATION OF RETRAFECT ESTIMATION

The causal question of interest is how would the observed impact measure sequence $\{Y_t\}_{t=\tau}^{N}$ change if a development program were to enter and remain in an area beginning in year $\bar{t}$ relative to no program access for this local population during the same sequence of time?

We define the following notation so that we can analyze RETRAFECT estimation in a counterfactual framework.

- $t = 1,2,\cdots,N$ denotes the time period. Note that we use N here because T has been used to indicate treatment in section 2.
- D_i^t: a sequence of binary treatment history for individual i up to time t. Note that the subscript here denotes the *history* up to time t. That is, $D_i^t = (D_{i1}, D_{i2}, \cdots, D_{it})$. $D_i^N = (D_{i1}, D_{i2}, \cdots, D_{iN})$, or the sequence of binary treatment for the whole period.
- δ_0^N: reference treatment, $\delta_0^N = (0, 0, \cdots, 0)_{N\times 1}$. This is a particular realization of D_i^N.
- δ_1^N: comparison treatment, $\delta_1^N = (0, 0, \cdots, 1, \cdots, 1)_{N\times 1}$, where 1 starts from year $\bar{t}$. This is another particular realization of D_i^N.
- Y_{it}^0 and Y_{it}^1: *potential* outcomes for individual i at time period t given reference treatment and comparison treatment, respectively, where $t = 1, 2, \cdots, N$. Note that at most one of the potential outcomes can be observed for any individual i.
- $D_i^{t-k,t+k}$: a sequence of binary treatment variables during the event window for t, where superscript $(t-k, t+k)$ implies the event window from period $t-k$ to $t+k$. That is, $D_i^{t-k,t+k} = (D_{it-k}, D_{it-k+1}, \cdots, D_{it}, D_{it+1}, D_{it+k})$.
- $\tilde{Z}_i^t$: a sequence of the history of observed potential causes other than the treatment for individual i up to time t. That is, $\tilde{Z}_i^t = (\tilde{Z}_{i1}, \tilde{Z}_{i2}, \cdots, \tilde{Z}_{it})$.
- $\ddot{Z}_i^t$: a sequence of the history of unobserved potential causes other than the treatment for individual i up to time t. That is, $\ddot{Z}_i^t = (\ddot{Z}_{i1}, \ddot{Z}_{i2}, \cdots, \ddot{Z}_{it})$.
- W_i^t: a sequence of the history of observed proxies for individual i up to time t. That is, $W_i^t = (W_{i1}, W_{i2}, \cdots, W_{it})$.
- $\tilde{X}_i^t$: a set of observed covariates, including observed causes and proxies. That is, $\tilde{X}_i^t \equiv \left(\tilde{Z}_i^t, W_i^t\right)$.

- X_i^t: $X_i^t \equiv \left(D_i^{t-k,t+k}, \tilde{X}_i^t\right)$.

Note that RETRAFECT assumes that only treatments within this event window can have effects on Y_{it}. Following the framework established in White (2006), the outcome is determined by a response function:

$$y_{it} = c_i^t\left(d_i^{t-k,t+k}, \tilde{z}_i^t, \ddot{z}_i^t\right), \tag{A1}$$

where we assume that the outcome at time t depends on i and depends on the sequence of treatment realizations during the event window $d_i^{t-k,t+k}$, the observed causes history $\tilde{z}_i^t$ and the unobserved causes history $\ddot{z}_i^t$. If we further assume that the outcome depends on i only through the arguments of the response function, the functional form c does not depend on i. Thus,

$$\begin{aligned} y_{it} &= c_i^t\left(d_i^{t-k,t+k}, \tilde{z}_i^t, \ddot{z}_i^t\right) \\ &= c^t\left(d_i^{t-k,t+k}, \tilde{z}_i^t, \ddot{z}_i^t\right). \end{aligned} \tag{A2}$$

Given $(\tilde{z}^t, \ddot{z}^t)$, the effect of the event at time t is

$$\begin{aligned} \Delta_{it}(\tilde{z}^t, \ddot{z}^t) &= c^t\left(\delta_1^{t-k,t+k}, \tilde{z}_i^t, \ddot{z}_i^t\right) - c^t\left(\delta_0^{t-k,t+k}, \tilde{z}_i^t, \ddot{z}_i^t\right) \\ &\equiv c_1^t(\tilde{z}_i^t, \ddot{z}_i^t) - c_0^t(\tilde{z}_i^t, \ddot{z}_i^t). \end{aligned}$$

Write

$$\begin{aligned} Y_{it}^0 &\equiv c_0^t\left(\tilde{Z}_i^t, \ddot{Z}_i^t\right), \text{and} \\ Y_{it}^1 &\equiv c_1^t\left(\tilde{Z}_i^t, \ddot{Z}_i^t\right). \end{aligned} \tag{A3}$$

At period t, the expected effect for treated for given x_i^t is

$$\begin{aligned} \Delta_{1it}(x_i^t) &\equiv E\left(Y_{it}^1 | D_i^{t-k,t+k} = \delta_1^{t-k,t+k},\ \tilde{X}_i^t = \tilde{x}_i^t\right) \\ &\quad - E\left(Y_{it}^0 | D_i^{t-k,t+k} = \delta_1^{t-k,t+k},\ \tilde{X}_i^t = \tilde{x}_i^t\right) \\ &\equiv \tilde{\mu}_{1t}(x_i^t) - \tilde{\mu}_{01t}(x_i^t), \end{aligned} \tag{A4}$$

where

$$\tilde{X}_i^t \equiv \left(\tilde{Z}_i^t, W_i^t\right), X_i^t \equiv \left(D_i^{t-k,t+k}, \tilde{X}_i^t\right)$$

and $\{W_{it}\}$ are observable proxies. Note that $\tilde{X}_i^t$ includes observed causes and proxies.

For the estimation of $\tilde{\mu}_{01t}$, assuming the unconfoundedness condition

$$\ddot{Z}_i^t \perp D_i^{t-k,t+k} | \tilde{X}_i^t, \text{ then}$$

$$\tilde{\mu}_{1t}(x_i^t) = \tilde{\mu}_{0t}(x_i^t) \equiv E\left(Y_{it}^0 | D_i^{t-k,t+k} = \delta_0^{t-k,t+k},\ \tilde{X}_i^t = \tilde{x}_i^t\right). \tag{A5}$$

Assuming only the contemporary values $\tilde{Z}_{it}$ and $\ddot{Z}_{it}$ affect the outcome, and the true relationship is a linear parametric one, i.e.,

$$
\begin{aligned}
Y_{it} &= c^t\left(D_i^{t-k,t+k}, \tilde{Z}_i^t, \ddot{Z}_i^t\right) \\
&= \textstyle\sum_{t=\bar{t}-k}^{\bar{t}+k} \tau_t D_{it} + \tilde{Z}'_{it}\theta + \ddot{Z}_{it}\,,
\end{aligned} \tag{A6}
$$

Where $\tilde{Z}'_{it}$ includes a constant term. In addition assume

$$\ddot{Z}_{it} \perp D_i^{t-k,t+k} | \tilde{X}_{it},$$

then

$$
\begin{aligned}
&E\left(Y_{it} | D_i^{t-k,t+k}, \tilde{X}_i^t\right) \\
&= \tilde{Z}'_{it}\theta + \sum_{t=\bar{t}-k}^{\bar{t}+k} \tau_t D_{it} + E\left[\ddot{Z}_{it} | D_i^{t-k,t+k}, \tilde{X}_i^t\right] \\
&= \tilde{Z}'_{it}\theta + \sum_{t=\bar{t}-k}^{\bar{t}+k} \tau_t D_{it} + E\left[\ddot{Z}_{it} | \tilde{X}_i^t\right] \\
&= \tilde{Z}'_{it}\theta + \sum_{t=\bar{t}-k}^{\bar{t}+k} \tau_t D_{it} + \tilde{Z}'_{it}\gamma_1 + W'_{it}\gamma_2
\end{aligned}
$$

$$= \tilde{X}'_{it}\beta + \textstyle\sum_{t=\bar{t}-k}^{\bar{t}+k} \tau_t D_{it}. \tag{A7}$$

Thus, τ_t can be consistently estimated by RETRAFECT.

Note that

$$Y_{it}^1 = \sum_{t=\bar{t}-k}^{\bar{t}+k} \tau_t \delta_{1it} + \tilde{Z}'_{it}\theta + \ddot{Z}_{it}$$

and

$$Y_{it}^0 = \sum_{t=\bar{t}-k}^{\bar{t}+k} \tau_t \delta_{0it} + \tilde{Z}'_{it}\theta + \ddot{Z}_{it}.$$

So the causal effect at time t would be

$$Y_{it}^1 - Y_{it}^0 = \textstyle\sum_{t=\bar{t}-k}^{\bar{t}+k} \tau_t (\delta_{1it} - \delta_{0it}). \tag{A8}$$

Because the reference treatment and comparison treatment are known, and because we can estimate τ_t, we can estimate the causal effect.

In the setting of credit availability, T_{it} is the treatment D_{it}, and we assume the effect lasts for up to k periods after the initial availability. The effect from the treatment at $t+1$ to $t+k$ should be viewed as the effect from the expected treatment which is known at time t, and we assume the effect does not depend on the initial year that microfinance is available.