Six Randomized Evaluations of Microcredit: Introduction and Further Steps

Abhijit Banerjee, Dean Karlan, and Jonathan Zinman*

September 11, 2014

Abstract

Causal evidence on microcredit impacts informs theory, practice, and debates about its effectiveness as a development tool. The six randomized evaluations in this volume use a variety of sampling, data collection, experimental design, and econometric strategies to identify causal effects of expanded access to microcredit on borrowers and/or communities. These methods are deployed across an impressive range of locations—six countries on four continents, urban and rural areas—borrower characteristics, loan characteristics, and lender characteristics. Summarizing and interpreting results across studies, we note a consistent pattern of modestly positive, but not transformative, effects. We also discuss directions for future research.

* banerjee@mit.edu (MIT, J-PAL): MIT Department of Economics, 40 Ames Street Building E17, Room 201A, Cambridge, MA 02142; dean.karlan@yale.edu (Yale University, IPA, J-PAL, and NBER): Yale University, P.O. Box 208269, New Haven, CT 06520-8269; jzinman@dartmouth.edu (Dartmouth College, IPA, J-PAL, and NBER): Department of Economics, 314 Rockefeller Hall, Dartmouth College, Hanover, NH 03755-3514. The authors would like to thank the coauthors and research teams of all papers that appear in this issue, and Zachary Groff, William Nowak, Glynis Startz, and Martin Sweeney for outstanding research assistance.

I Motivation

Microcredit, typically defined as the provision of small loans to underserved entrepreneurs, has been both celebrated and vilified as a development tool. In its heyday, microcredit was the basis for the 2006 Nobel Peace Prize and embraced by policymakers, donors, and funders worldwide as an effective policy tool.[1] But around this time, some observers, including many of the researchers represented in this volume, pointed out that the evidentiary base for anointing microcredit was quite thin. Various theories—of poverty traps, behavioral decision making, general equilibrium effects, and/or credit market competition—suggest that the impacts of expanding access to credit on poor people need not be positive, and could even be negative.[2]

The empirical evidence invoked by microcredit's proponents was largely based on anecdotes, descriptive statistics, and impact studies that failed to disentangle causation from correlation. More recently, these same questionable methods are often invoked to criticize microlenders for high interest rates, serial borrowing, default crises, and other symptoms of debt traps.

The six studies in this volume grew out of the debates that started in the 2000s and continue today. They generate causal evidence on the impacts of microcredit on its intended users with research designs that rely on some randomness in the allocation of credit offers by individual microlenders.

Randomization permits the identification of causal effects by minimizing the selection biases that can confound observational studies. In the case of microcredit, these biases can come from both the demand-side and the supply-side, as we discuss in Section III. The magnitude of these selection

[1] An example quote, from the Prime Minister of Bangladesh at the 1997 Microcredit Summit, "In our careful assessment, meeting the credit needs of the poor is one of the most effective ways to fight exploitation and poverty. I believe that this campaign will become one of the great humanitarian movements of history. This campaign will allow the world's poorest people to free themselves from the bondage of poverty and deprivation to bloom to their fullest potentials to the benefit of all – rich and poor."

[2] See Banerjee (2013) and Zinman (forthcoming) for related reviews of theory and evidence on such issues.

biases is an empirical question, and few studies focus on estimating how important they are in practice (Beaman et al. 2014 being an exception). Nevertheless, the studies in this volume do generate some circumstantial evidence of various types. Beyond the relatively obvious—comparing naïve OLS estimates to those relying on experimental variation within-study, and comparing randomized to non-randomized estimates across studies—models estimating correlations between borrowing likelihood and baseline characteristics also contain some useful information.

"Who borrows?" is of course the combination and demand- and supply-side factors, and a key question for non-randomized identification strategies is whether an econometrician can plausibly capture all the important factors. The modest R-squareds in the borrowing-likelihood models estimated in this volume suggest not. It is difficult to predict microcredit use even with a rich set of baseline characteristics of potential borrowers.

Aside from the use of rigorous methods for identifying causal effects, another noteworthy feature of the studies here is their collective spanning of contextual, social, and market environments, thus lending external validity to their results, especially given the consistent pattern of findings summarized below. Taken together, the studies cover microcredit expansions by seven different lenders in six different countries—Bosnia, Ethiopia, India, Mexico, Morocco and Mongolia—from 2003-2012. These lenders, products, and settings strike us a fairly representative of the microcredit industry/movement worldwide. Combined with two previously published field experiments from other settings (Karlan and Zinman 2010, Karlan and Zinman 2011), the evidence here adds to the evidentiary base for evaluating and refining the theories and practices of microcredit.

Drawing lessons across the six studies has been greatly facilitated by the efforts of the six research teams and the Editor, Esther Duflo, to make the papers readily comparable. Although each study was designed and implemented independently, everyone has exerted substantial effort in the analysis and editorial phases to use comparable outcomes, estimation strategies, and expositional organization. The data used in each study is posted on the AEJ website and we encourage researchers to use it; e.g., it is well-suited for more formal meta-analysis than we do here.

What are the key takeaways, in our estimation? One is the existence of modest take-up rates of credit among (prospective) microentrepreneurs, which is a prima facie case against microcredit being a panacea (a cure-*all*) in the literal sense, and presents a statistical power challenge for randomized identification strategies. Second is the difficulty of meeting the power challenge by predicting microcredit take-up (as noted above, this speaks to the likely importance of unobserved heterogeneity in borrowing and lending decisions, and hence to the value of large-sample randomization). Third is the lack of evidence of transformative effects on the average borrower. Fourth is that the lack of transformative effects does not seem to be for lack of trying in the sense of investment in business growth. Fifth is that the lack of transformative effects should not obscure other more modest but potentially important effects. If microcredit's promise were increasing freedom of choice it would be closer to delivering on it. Sixth, just as there is little support for microcredit's strongest claims, there is little support for microcredit's harshest critics, at least with respect to the average borrower. Seventh, the limited analysis of heterogeneous treatment effects in these studies does suggest hints of segmented transformative effects—good for some, bad for others.

Another key lesson also leads to a critical caveat: statistical power still poses a major challenge to microcredit impact studies. This motivates both additional studies and formal meta-analyses going forward. But at this moment we cannot emphasize enough that, at the level of individual studies, many of the null results described here are part of confidence intervals that contain economically meaningful effect sizes of credit access in one or both directions. Given the modest take-up rates ("low compliance" in the program evaluation sense), the confidence intervals on estimates of the effects of *borrowing* are almost certainly much wider still. (How much wider is difficult to identify, for reasons we discuss below.) In short, *most* of the null results in the studies here would lack precision if they were converted into treatment-on-the-treated units. The individual studies may lack strong evidence for transformative effects on the average borrower, but they also lack strong evidence against transformative effects.

Another key caveat is that these studies have nothing to say about impacts on *infra*marginal borrowers. It may well be the case that impacts are substantially different on the borrowers and/or communities already being served before the lenders in these studies began experimenting on the margin. Although marginal borrowers are the most relevant ones for current policy, funding, and practice decisions, understanding effects on inframarginal borrowers definitely has value for testing theory and informing the design of future interventions. We could imagine experiments on inframarginal borrowers—e.g., on loan amounts or maturities, or by allocating repeat loans by lottery in settings where loanable funds are scarce—but are not aware of any.

A final caveat is that these studies do not have much to say about the benefit-cost proposition of expanding access to microcredit. E.g., only one of the papers considers the (un)profitability of marginal loans for the lender.

The rest of this Introduction proceeds as follows. Section II summarizes the study settings. Section III discusses potential biases in observational studies of microcredit impacts and the methods these studies use to address them, focusing on important nuances in the unit of randomization, data collection, and econometric strategies. Section IV summarizes some key takeways from the results. Section V uses some simple theory to help interpret the results. Section VI highlights some unanswered questions and directions for future research.

II The microcredit products

Table 1 summarizes the six study “settings”: characteristics of the loans, lender, borrower, and locations where each study took place. We present the data on time periods covered in Table 2, which shows start dates ranging from 2003 to 2010, and end dates ranging from 2006 to 2012. Of the six programs evaluated, four are traditional group lending, one is individual lending (Bosnia), and one includes both group and individual loans (Mongolia). Group sizes vary considerably across the studies, ranging from three or four borrowers in Morocco to as many as 50 in Mexico (with a minimum of 10). Each of the group lenders relies heavily on groups for screening and monitoring. Joint liability gives groups incentives to perform these functions, although anecdotal evidence indicates that lender enforcement of joint liability is imperfect and varies across organizations. The individual loan underwriting relies heavily on loan officers. At least three of the lenders provide dynamic incentives (better terms on subsequent loans) for loan repayment. According to data from MixMarket.org, loan loss rates of the lenders range from 0 to 3.2 percent at baseline (excluding the anonymous Bosnian lender).

Each lender targets borrowers with certain characteristics. Three lend only to women. Most have minimum and/or maximum ages. All but one (India) targets microentrepreneurs. Some require explicit business plans to get loans, while others restrict lending (at least nominally) to people who already had a business or were planning to start one. However, none of the lenders actually restrict how disbursements are spent or attempt to monitor whether business investment increases.

Nominal interest rates (APRs) vary from 12 percent to 27 percent with the one exception of Mexico, which has a rate of approximately 110 percent. Each of the APRs but India’s is substantially below the median market microloan interest rate for the country.[3]

[3] We calculated market interest rates using MIX Market data. The following formulas were used to determine APR:

Regarding non-price loan terms, average loan sizes vary substantially in PPP terms across studies, from about \$450 in Mexico to about \$1,800 in Bosnia. Average loan amount as a proportion of family income ranges from 6 percent in Mexico to 118 percent in Ethiopia. Average loan maturities range from four months in Mexico to 16 months in Morocco, with four of the studies falling in the 12-16 month range. Each lender except Ethiopia has a fixed repayment schedule, with two requiring weekly, two monthly, and Morocco offering a variety of schedules. Half of the lenders required or requested collateral, broadly defined. Borrowing groups in Mongolia have to put 20 percent of the loan amount in a joint savings account before loan proceeds are disbursed. 77 percent of Bosnian borrowers post "collateral", although in most cases this is just a guarantee provided by the borrower and/or a co-signer. In Ethiopia the lender's guideline is to not ask for collateral, but in practice most borrowers report, in the endline data, being asked for some.

In each case the lending function—the provision of liquidity—is performed by the lender (i.e., these are not ROSCAs). The seven lenders across these studies include a mix of for-profits (India, Mexico, and Mongolia) and non-profits. Most of the microlenders in these studies can be described as large: five of the six for which we have data have at least \$190 million in microloans outstanding as of 2012 (\$190 million is the 93rd percentile of microloans outstanding in MixMarket.org data), and the anonymous Bosnian microlender is described as a large MFI by the researchers. As of 2012, the Mexican lender was one of three publicly-traded microlenders in the world.

All told, the six settings represented in this volume strike us as fairly representative of the distribution of lenders, loan terms, borrowers, and mar-

1. $weekly\ rate = (\%\ nominal\ yield\ on\ gross\ portfolio + loan\ loss\ rate)/(1 - loan\ loss\ rate) \times 7/365.25$

2. $APR = (1 + weekly\ rate)^{365.25/7} - 1$

For each study, we selected data from the year that was closest to the onset of the treatment and restricted the analysis to the ten largest lenders (by gross portfolio) in each country for which data was available.

kets that comprise the microcredit world. MixMarket.org and MFTransparency.org are useful sources of comparison data.

III Research methodology

The motivation for relying on randomized variation to identify microcredit impacts is pretty standard: concerns about selection biases. In the case of microcredit one should worry about both supply- and demand-side biases.

Concerns about demand-side selection biases stem from the likelihood that people who choose to borrow may be different, or trending differently, than those who choose not to borrow. If these differences are unobserved—not fully controlled for—by the econometrician, and correlated with downstream outcomes of interest, then estimates of the effects of microcredit will be biased: they will not capture true causal effects.

Biases can go in either direction. People may borrow when they have experienced, or expect to experience, a negative shock, a la the "Ashfenfelter dip" in job training. This will produce downward bias: impact estimates will be lower (less positive) than their true causal effects. But we could also get an upward bias: people may borrow when they expect an improvement, or because they are more capable in ways that are difficult to measure, a la the longstanding concern with the confounding effects of "spirit" and "spunk" in studies of labor markets and entrepreneurship.[4]

Similarly, concerns about supply-side selection biases stem from the likelihood that lenders, like borrowers, make strategic decisions based on factors that are difficult for researchers to fully observe. Lenders choose which neighborhoods and markets to enter, and depending on their motivation may thus select relatively vibrant and growing markets (because of profitability) or stagnant and particularly poor markets (because of social concerns). Lenders (and/or borrowers, in the case of group lending) also select or ra-

[4]See Beaman et al. (2014) for empirical evidence that there is selection on returns.

tion from a pool of prospective borrowers; again, selection biases can go in either direction, depending on the objective function of whomever is doing the underwriting.

While all the studies discussed here are randomized control trials, there are actually two types of experimental designs here. Five of the studies use randomized program placement. The Bosnia study uses individual-level randomization.

Randomized program placement requires lenders (or lenders alongside researchers) to identify a large set of communities (or neighborhoods) and randomly assign each community to either treatment or control. Lenders then employ one or more levers for ensuring that those in areas assigned to treatment end up being more likely to be offered a loan. One lever is an approach often used in encouragement designs - marketing only (or more intensively) in treatment areas. Another lever is using address verification to try to prevent people in control areas from borrowing. Of course, even applying both of these levers does not guarantee perfect supply-side compliance with the treatment assignment (we discuss demand-side issues affecting power below). For example, the lender's field staff may have incentives to deviate, and address verification may be gameable (home vs. work addresses may fall in different communities, etc.). This is one reason why it is important to verify whether and how much random assignment to treatment predicts increased borrowing from the lender; i.e., to test how strong the threshold component of the first stage is. A study should have sufficiently large effects here to have hope of identifying downstream effects of borrowing or credit access.

The Bosnia study uses credit scoring to engineer individual-level randomization, similar to Karlan and Zinman (2010) and Karlan and Zinman (2011). This approach typically randomizes applicants who are deemed by the lender to be on the margin of creditworthiness. Individual-level randomization can present compliance challenges as well, particularly in cases where loan officers retain some operational freedom to deviate from the random assignment and/or individuals have outside options that depress take-up rates conditional on applying.[5]

The two different levels of randomization each have their advantages and disadvantages. Randomized program placement offers the potential benefit of capturing treatment effects at the community level (more precisely, at the level of the unit of randomization). In many cases this level of analysis is more interesting because it internalizes any spillovers or general equilibrium effects that occur within the community (see the Morocco paper for a direct analysis of such spillovers). As long as such effects are contained within communities (no cross-community spillovers),[6] randomized program placement thus addresses an important challenge in identifying the impact of microcredit: non-borrowers may be affected as well, either positively or negatively. Individual-level randomization, on its own, is not well-suited to measuring spillovers or general equilibrium effects. In principle, however, one could combine the two levels of randomization—for instance, by randomly approving 10 percent of marginal rejects in some communities, and 90 percent in other communities—to capture the community-level effects in situations where individual-level randomization is attractive for other reasons (see below).

The main downside of randomized program placement is typically a loss of statistical power, in large part because take-up rates are relatively low. Take-up rate differentials between treatment and control—the primary determinant of statistical power—are as low as 9 percent among the studies here. In contrast, with individual-level randomization, it is easier to create a sample of people who have already expressed interest in credit (e.g., by

[5] As is always the case, one must consider the sources and magnitudes of imperfect compliance with random assignment when interpreting treatment effect estimates.

[6] A promising way to test the assumption of no cross-community spillovers is to use an additional randomization, varying the intensity of credit offering across (groups of) treated communities. This randomization was done in the Mexico study and is the subject of work-in-progress.

applying for it), and thus obtain relatively high differential take-up rates. In some cases this boils down to a budget issue where one can offset lower power by increasing the sample size. In other cases increasing sample size may not be operationally feasible.[7]

IV What do the studies find?

In summarizing and distilling results across the six papers in this issue, we follow the papers in focusing on average intent-to-treat effects. We also provide a bit of discussion on treatment-on-the treated and heterogeneous effects. Where applicable we focus on the outcomes measured at the second of two endline surveys, unless otherwise noted.

Each of the six studies starts by estimating treatment effects on credit access and usage (Tables 2a and 2b in each paper), as these are a first stage of sorts for establishing the plausibility of effects on downstream outcomes. There are three challenges to engineering a powerful first stage through random assignment, each typical of experimental designs that rely on encouragement—rather than compulsion or strong added inducements—to use a treatment. One is the potential modest demand for microcredit. In practice,

[7]Aside from the potential power benefit, the main advantage of individual-level randomization is that it is becoming relatively attractive to lenders over time. In contrast, opportunities for randomizing program placement in tandem with planned geographic expansions may dry up as microcredit markets approach maturity/saturation. Many lenders are adopting credit scoring models and are open to the proposition that experimentation delivers bottom-line benefits in the form of model refinements. This approach could even work for group lending, either as an input to group screening or as automation of final-stage underwriting in cases where the lender retains discretion to deny loans approved by a group.

It is also important to note that the two approaches will often estimate treatment effects on different margins of borrowers. Randomized program placement typically identifies effects of credit expansion to new communities. Individual-level randomization typically identifies effects of credit expansion (or contraction) within markets that are already being served by the lender. Both margins are interesting, but there may be cases where, for example, using individual-level randomization to focus on a relatively narrow margin of borrowers is less interesting for policy purposes or more interesting for testing theories.

these studies estimate take-up rates of study-specific loan products ranging from approximately 17 percent to 31 percent among their target populations (the Bosnia study, which targeted marginal applicants, had a take-up rate of 100 percent, and the Mongolia study, which targeted women who had expressed interest in receiving a loan, had take-up rates of 50 percent and 57 percent for individual and group products, respectively).

The second challenge combines the first with the challenges of experimental implementation: potentially low or modest take-up rate differentials between treatment and control areas. In practice we see this in India, Mexico, and Morocco, where treatment effects on the likelihood of borrowing from the implementing MFI range from 9 to 12 percentage points. Fortunately each of these studies anticipated the modest take-up differential at the research design phase and compensated with a large sample size (Mexico) or methods for predicting take-up (India and Morocco). The other studies have differentials in roughly the 25 to 50 percentage point range.

The third challenge combines the first two with the potential presence of close substitutes for the offered treatment: additional microlending is unlikely to have much of an effect in a competitive market. This was less a concern in places where researchers could be confident ex-ante that the study area was largely untouched by other MFIs. But in Bosnia, India, and Mexico it is particularly critical to examine whether borrowing from any other MFI goes up to compensate, and each of the studies find that it does—modestly in India and Mexico ($<$ 10 pp), and quite substantially in Bosnia.

Some of the effects on credit are also interesting in their own right. In particular the question of whether expanded access from one MFI substitutes for or complements other credit sources relates to questions about the nature of liquidity constraints, risk-sharing, borrower production functions, and mechanisms producing downstream outcomes. For example, findings of complementarity would suggest a combination of liquidity constraints still binding and lumpiness in borrower opportunity sets, and/or the need to re-finance marginal debt instead of paying it off (raising concerns about debt traps). Instead, we see some evidence of substitution for at least some segment of the borrowers. Informal borrowing falls in the one site where it was prevalent at baseline (India). This suggests some substitution of informal for formal credit—reduced demand—but leaves unanswered the related question of whether increased formal access disrupts informal risk-sharing and hence the supply of informal credit. The Mongolia study finds some evidence of partial (perhaps 20-25 percent) crowd-out of other formal credit, while none of the other studies finds strong effects. In all, there is some evidence of substitution among different credit types, and no evidence that expanded access from one MFI leads borrowers to take on additional debt from other sources, though in this case, as for other outcomes, this may conceal important heterogeneity–it could be substitutes for some and complements for others.

The studies also estimate treatment effects on a large variety and number of "downstream" outcomes (study mean number of downstream treatment effects estimated = 50, min = 36, max = 82). Fortunately there is substantial outcome comparability across studies, especially with respect to what we and the studies consider particularly key outcomes.

Conducting inference on the effects of microcredit access on downstream outcomes is complicated by measurement challenges, fungibility, and heterogeneity. Measurement error is always a challenge when income or consumption is among the main outcomes of interest. Measurement error and reporting biases complicate even the seemingly straightforward exercise of identifying how loan proceeds are spent (Karlan, Osman, and Zinman 2013). Heterogeneity in loan uses can make some effects hard to detect, and standard corrections for multiple hypothesis testing may be too conservative. For example, if some households use loans to grow businesses, while others use them for consumption smoothing, and still others for solidarity or empowerment, then the effects of microcredit will diffuse across multiple outcome families, and may be too small to detect on many of the individual outcome families. In light of these issues, we find it unsurprising that none of the studies finds statistically significant effects at the 10 percent level on even half of the downstream outcomes tested, with a range from 6 percent to 39 percent. Given these challenges, the studies in this issue make judicious use of theory and practice to identify especially key (families of) outcomes.

One key outcome family is microentrepreneurial activity (Table 3 in each of the papers). Even though this might be considered more of an intermediate outcome—a means to the end of greater utility rather than a great proxy for utility in and of itself—the fact that most microlenders target potential microentrepreneurs means that treatment effects on entrepreneurship (or at least self-employment) constitute a litmus test of sorts. If we do not find increases in business likelihood, size, and/or profitability, it is unlikely that microcredit, at least as traditionally defined, is delivering on its promise of reducing poverty by relaxing credit constraints that inhibit business growth. The full picture of evidence suggests at least a partial passing of the litmus test. The effects on extensive margins (ownership, starts, closures) are modest, with three of the studies finding no effects, Bosnia finding an effect on ownership, Mongolia finding an effect on ownership from group borrowing only, and India finding an effect on starts at the first endline only that is quite small in level terms (although large in percentage terms).

The effects on measures of investment, business size, and profits, which combine the intensive and extensive margins, are more promising. Five of the studies have measures of business assets and/or investment, and eight of the 10 point estimates on these measures are positive, with two of the positive ones (and none of the negatives) reaching statistical significance.[8] This suggests that the average effect, pooling across studies, is likely statistically as well as economically significant. Five of the studies have measures

[8]Outcomes from Mongolia are counted twice (joint and individual liability arms). There are two measures that were only considered in Ethiopia: value of livestock owned and value of large animals owned.

of revenue and/or expenses, and the point estimates on all 13 of these measures are positive, with 6 of them reaching statistical significance.[9] Each of the studies measures profits, and here we have seven positive point estimates and one zero, with one statistically significant result.[10] Again, our eyeballing suggests that pooling across would yield significant increases in business size and profits.

All told, each study finds at least some evidence, on some margin, that expanded access to credit increases business activity.

Another key outcome family is income (Table 4 in each paper); after all, increased income is essential to poverty reduction. None of the six studies finds a statistically significant increase in total household income, although key point estimates are positive in four of the six studies.[11]

The results on income composition are somewhat more encouraging. Of the four studies with measures of wage income and business income, two find evidence of increases in business income offsetting reductions in wage income. The two remaining studies find increases in both wage and business income. Out of eight point estimates of effects on income from remittances/transfers or government aid/benefits, five are negative. These results suggest that although microcredit may not be transformative in the sense of lifting people or communities out of poverty, it does afford people more freedom in their choices (e.g., of occupation) and the possibility of being self-reliant.

Yet another important outcome family is consumption (or, more precisely in some cases, consumption expenditures), which is a widely-used proxy for living standards (Table 6 in each of the papers except Ethiopia, which did

[9] Again, outcomes from Mongolia are counted twice (joint and individual liability arms). There are three measures that were only considered in Ethiopia: cash revenues from crops, expenses for crop cultivation, and livestock sales.

[10] There is one measure that was only considered in Ethiopia: net revenues from crops. The profit measures in the Mongolia paper for respondent businesses only, both of which are negative, are not counted here.

[11] Only one study (Morocco) calculates a point estimate on total household income. For all other studies, total household income is derived by summing all reported income components.

not measure consumption). The results from the four studies with a measure of total household consumption find no evidence of an increase. Three find fairly precise null effects, at least in intention-to-treat terms (India, Mongolia, and Morocco). Bosnia finds a significant reduction, although this may be due to the fact that most borrowers in the sample were still paying back their initial loans (see the theory section, Table 1 re: averages for loan term, and Table 2 re: time between treatment onset and endline). Each study, including Ethiopia, has some measure of food consumption, and the results are mixed at best. Four studies find null effects (though only Morocco can rule out effects greater than +/- 5 percent of the control group mean), Mongolia finds evidence of a modest increase in the group lending treatment, and Ethiopia finds evidence of substantial decrease (more precisely, there is an increase in food insecurity reported in their Table 7).

Other results on the composition of consumption, as with income, suggest some potentially important if not transformative effects. In measuring durable consumption we think it is important to start with measures of durable stocks rather than expenditures, since increases in the latter may indicate increased churn of assets from the strain of the debt service (see the Mexico paper for a discussion of this issue, and Morocco's Table 4 for related evidence of a decline in asset sales). The three studies that measure durable stock(s) find mixed results. In Mongolia, microcredit access increases the stock of household durables (the effect is statistically significant in the individual lending treatment). In Bosnia and India, however, microcredit decreases the stock of durables, and the effect is statistically significant in Bosnia. One robust finding on consumption is a decrease in discretionary spending (temptation goods, recreation/entertainment/celebrations). Five of the studies estimate treatment effects on 10 such measures, finding seven negative point estimates, three of which are statistically significant. Whether this belt-tightening is indicative of improved self-discipline, changes in bargaining power within the family, the lumpiness of investment, and/or some other mechanism is an open question. Another robust finding on expenditures is the lack of significant effects on other types of spending (health and education), although many of these nulls are imprecisely estimated.

The final outcome family we consider focuses on social indicators (Table 7 in each of the papers). Two particularly important outcomes here are child schooling (as an indicator of child welfare and leading indicator of family income growth) and female empowerment. Each of the six studies estimates treatment effects on schooling, and the effects are a mix of more and less precisely estimated nulls. The one exception is Bosnia, which finds a significant decline in school attendance among 16-19 year-olds. Four of the studies estimate effects on female decision power and/or independence within the household, and three find no effect. India's null is precisely estimated (in intent-to-treat terms at least), ruling out effects larger than +/- 0.05 standard deviations. Mexico finds a small but significant increase in female decision power.

Other results from Mexico (Tables 7 and 8) raise the possibility of meaningful effects on other aspects of subjective well-being: happiness and trust in others each increase by an estimated 0.05 standard deviations, although many other indicators are unaffected. All told, the studies find no evidence of transformative effects on social indicators, but do find some hints of positive effects on female empowerment and well-being, at least in Mexico.

The full picture of the evidence suggests several tentative conclusions. Reassuringly, these echo the conclusions in a previous survey which covers a number of studies not in this issue (Banerjee 2014).

First, there is little evidence of transformative effects. The studies do not find clear evidence, or even much in the way of suggestive evidence, of reductions in poverty or substantial improvements in living standards. Nor is there robust evidence of improvements in social indicators.

Second, the lack of transformative effects is not for lack of trying in the sense of investment in business growth. There is pretty strong evidence that businesses expand, though the extent of expansion may be limited, and hints (eyeballing the pattern of positive coefficients across studies) that profits increase. The evidence on why expansion does not produce strong evidence of increases in household living standards is mixed: some studies find evidence suggesting that households trade off business income for wage income, while others suggest that larger businesses are no more profitable, even in level terms, than smaller ones (at least on average; we discuss heterogeneity below).

Third, the lack of transformative effects should not obscure other more modest, but potentially important, effects. The studies find some, if not entirely robust, evidence of effects on occupational choice, business scale, consumption choice, female decision power, and improved risk management. As we stated previously, microcredit's promise were increasing freedom of choice it would be closer to delivering on it.

Fourth, just as there is little support for microcredit's strongest claims, nor is there much support for microcredit's harshest critics. The studies find little evidence of harmful effects, even with individual lending (Bosnia, Mongolia), and even at a high real interest rate (Mexico).

Fifth, the limited analysis of heterogeneous treatment effects in these studies does suggest the possibility of transformative effects—good for some, bad for others—on segments of microlenders' target populations. Morocco and India find evidence of large positive effects on business profits in the right tail of the distribution, and Morocco finds negative effects on the left tail. Mexico finds an increase in financial decision power among the left tail. Bosnia finds a decrease in teen schooling among lower-educated households. Microcredit's strongest supporters and harshest critics may each be correct for segments of borrowers, if not on average.

Finally, we emphasize that many of these inferences lack precision, at least at the level of individual studies. Many of the null results are part of confidence intervals that contain economically meaningful effect sizes in one or both directions. Most of the null results would lack precision if they were converted into treatment-on-the-treated units. Statistical power still poses a major challenge to microcredit impact studies, although there are many hopeful indicators for future work, including the progress in the India and Morocco studies in predicting take-up, the Mexico study in obtaining a large sample size, and the prospect of pooling data across studies.

V Some simple theory

The results reported above highlight the fact that many things change when a family gets access to microcredit. Even if microcredit is beneficial, we do not expect all things to change in the same direction at the same time. The model in this section provides a simple framework for thinking about potentially diffuse and dynamic impacts of microcredit. The basic idea of the model is that potential borrowers have a lumpy expenditure opportunity that would generate benefits both in the present and the future. Additional credit facilitates that investment because otherwise, the required cut in present consumption to finance the investment may be unacceptably large.

As long as the marginal amount borrowed is not too large, expanded access to microloans can create interesting dynamics for borrowers. They may cut consumption (including leisure) in the short run, and may permanently reduce certain types of consumption (while increasing others). Labor supply may also go up in the short run to mitigate the negative effect of consumption, but may go down thereafter. In other words, neither an immediate fall in consumption nor an eventual fall in labor supply is necessarily evidence that microcredit has failed to deliver the goods.

The results on the (lack of) impact on consumption spending could be construed as broadly consistent with this prediction of the model. Perhaps some borrowers have already gained in terms of consumption while others are currently consuming less but will consume more later so that there is no net effect. Similarly, while we do not emphasize changes in labor supply, the India study does find an increase in labor supply on the first round of data collection, which goes away by the second round. Of course, all these interpretations are subject to the concern that we do not know how to define "short run" and "long run". E.g., five of the six studies have endline data collected after enough time has elapsed for several loan cycles. Is this short enough to capture short-run effects or should we think of all the measured effects reported here as long-run effects? We return to this question in the Conclusion.

A Basic Model

A consumer "lives" for 2 periods. She can spend money on two goods which we will call non-durable and durable. The non-durable is fully divisible and is consumed during the period in which it is bought. Denote non-durable consumption by c_n. The durable lasts for two periods, and yields durable services in both periods. If it is a business durable then the services are just outputs of the consumption good; if it is a consumer durable they are flows of instantaneous consumption emanating from the durable.

The durable is indivisible, costs an amount c_d, and yields durable services of ac_d in each period. Moreover there are no additional benefits from owning a second durable. Assume that durable services and non-durables are perfect substitutes, in the sense that the consumer's per-period utility function is $u(c)$, where $c = c_n$ if she has not purchased the durable in the current or previous period and $c = c_n + ac_d$ otherwise, and that $0 < a < 1$. Therefore in the current period purchasing the durable leads to a net loss in flow utility, but it might still be optimal because a could be greater than $1/2$. The consumer discounts the future at rate δ and maximizes total of present and discounted future utility.

The consumer earns a labor income of y in units of the non-durable every period and there is no savings, so the total amount y is spent every period.

However, the household has the option of borrowing up to an amount b^{max} for one period at a gross interest rate r. We assume that the durable costs more than the maximum possible amount of debt: $c_d > b^{max}$.

B Analysis of the model

It turns out that we can study the consumer's decision diagrammatically. In Figure 1, the horizontal axis represents consumption in period 1 and the vertical axis is consumption in period 2. UU and $U'U'$ are two potential indifference curves. They both have slope $1/\delta$ when they intersect the 45 degree line, OO', at points E and E'. The point E represents the endowment, the vector (y, y). The line EF, which has the slope r, represents the set of options open to the consumer if she borrows in period 1 but does not purchase the durable. The distance along the horizontal direction from E to F represents b^{max}, the maximum possible loan size. As drawn, we are assuming that $r < 1/\delta$, which gives the consumer a reason to borrow–the highest indifference curve reachable on EF is typically higher than the one through E.

Another option is to buy the durable without borrowing. The point A represents this case, i.e. it is the point $(y - (1 - a)c_d, y + ac_d)$.

The third option is to borrow and buy the durable. The line segment AB represents the set of choices for someone who does so. The horizontal distance from A to B is b^{max} and the slope of the line is r. As drawn, it is clear that the point B lies on the highest indifference curve that is available and the consumer will choose both to borrow and to buy the durable. However, her first-period consumption is still lower than at point E. Total consumption goes down in the first period as a result of purchasing the durable.

However, this is not the only possibility. The point B' represents what happens when b^{max} is higher (F' is the corresponding point where the consumer borrows without purchasing the durable). In this case, borrowing and buying the durable is still the best option, but total consumption goes up in both periods. Finally, the point B'' represents the case where b^{max} is small. F'' is the corresponding value in the case where there is no durable purchase. In this case, borrowing without buying the durable is the best option, and first-period consumption goes up.

Figure 2 captures the case where $r\delta > 1$. In this case there is no reason to just borrow–the line EF lies everywhere under the indifference curve through E. However, borrowing to buy the durable still makes sense and improves welfare.

In general, more credit (weakly) increases the incentive to buy the durable relative to either not buying but borrowing or not buying and not borrowing. To see this denote the utility of buying the durable as $v_d(b^{max})$, and that of not buying the durable by $v_n(b^{max})$.

$$\frac{d}{db^{max}}v_d(b^{max})=\max\{\frac{d}{db}[u(y-(1-a)c_d+b)+\delta u(y+c_d-rb)],0\}=\max\{u'(y-(1-a)c_d+b)-\delta ru'(y+ac_d-rb),0\}$$

which, by the concavity of u is always at least as large as $\frac{dv_n(b^{max})}{db^{max}} = \max\{\frac{d}{db}[u(y+b) + \delta u(y-rb)], 0\} = \max\{u'(y+b) - \delta ru(y-rb), 0\}$.Therefore this is also true at the point where $v_d(b^{max}) = v_n(b^{max})$, which tells us that if at any level of b^{max} $v_d(b^{max}) > v_n(b^{max})$, then this is also true at all higher values of b^{max}. In this sense, increased access to credit favors buying the durable.

Moreover, it is evident that when the consumer switches to buying the durable as a result of increased credit access, her borrowing must go up. Hence, compared to someone who has less credit access, her second-period non-durable consumption, $y - rb$, must be lower.

Result 1: Compare two people, one of whom has better access to credit. She is more likely to buy the durable, but her first-period total non-durable consumption and even total consumption may be higher or lower. If she buys a consumer durable, her second-period non-durable consumption will be lower. If she buys a business durable, her second-period non-durable consumption will be higher.

If the durable is for business, then it makes sense to think of a higher a as a more productive project. For a high enough a, the investment will be made even when access to credit is very limited or absent. Conversely, increased access to credit will encourage consumers with relatively low values of a to invest.

Result 2: Increased access to credit increases the likelihood that the consumer makes a fixed investment but reduce the average product of the projects that get implemented.

Next, consider a variant of the model where the consumer also has a labor supply decision. Assume that the consumer can earn w units of non-durable consumption per unit of labor and supplies l_1 and l_2 units of labor in periods 1 and 2. The disutility of labor is given by the function $v(l)$ which is assumed to be increasing, convex, differentiable everywhere and satisfying the Inada condition at $l = 0$. The consumer now maximizes

$$u(y - (1-a)c_d + b + wl_1) - v(l_1) + \delta[u(y + c_d - rb + wl_2) - v(l_2)$$

if she buys the durable and

$$u(y + b + wl_1) - v(l_1) + \delta[u(y - rb + wl_2) - v(l_2)]$$

if not.

By our assumptions about v, an interior optimum for l always exists and is given by

$$u'(c) = v'(l).$$

It is evident that l is decreasing in c. Furthermore, if $u_l(x) = \max_{\mathrm{l}}\{u(x + wl) - v(l)\}$, it is easy to show that $u_l(x)$ inherits the concavity of $u(c)$ and therefore Result 1 extends to this case. In other words, improved loan access may lead to a reduction in non-durable and even total consumption in the first period. If total consumption goes down, labor supply will go up in that period.

Result 3: Increased access to credit can lead to an increase in labor supply in the first period.

Finally, the assumption that durables and non-durables are perfect substitutes is convenient for diagrammatic analysis but not essential for our results. Suppose, on the contrary, durable consumption of c_d leads to an utility equal to the service flow from the durable ac_d, which is separable from the utility from non-durables. Then it is easily shown by following the same argument that Result 1 will still hold. The only change is that now labor supply only depends on non-durable consumption, and since non-durable consumption can be lower in both periods, labor supply may be permanently raised by improved credit access.[12]

Result 4: If durables and non-durables are not perfect substitutes, increased access to credit may raise labor supply in both periods.

VI Further research

The six studies in this paper greatly add to the evidentiary foundation on the impacts of microcredit. But they stop short of fully answering many questions, highlighting several lines of inquiry for research going forward.

One is continuing to develop methods that address the power challenges seemingly inherent to encouragement designs. The India and Morocco studies in this volume make encouraging strides.

A second is to continue exploring impacts—on borrowing and spending decisions, and downstream outcomes—at different horizons, from the very short run (Karlan, Osman, and Zinman 2013) to even longer horizons than those studied here. It should be noted that many of the studies in this

[12]The same result also holds when instead of durables and non-durables, the consumer chooses between a divisible consumption good and a non-divisible one (say, a wedding).

volume do consider the long run in a sense: they measure impacts after several (potential) loan cycles have elapsed. Moreover, looking across studies, there is little evidence of treatment effects on investments with relatively long gestation periods, such as education or health, and there is little evidence of transformative effects on social indicators—e.g., decision power, locus of control—that might later feed back into greater success in other domains/outcome families. There is also a practical challenge in studying long-run impacts: absent a strong motivation to do so, and with evidence that microcredit delivers some benefits, the case for withholding microcredit from a control group for several (more) years weakens. Having said that, two longer-run follow-ups to the experiments in this volume are in the works: the India team conducted surveys 7 years post-random assignment, and the Morocco team is in the field for another follow-up survey, 8 year post-random assignment.

A third line of inquiry is to continue exploring whether and why results replicate across different "settings", which we define loosely as credit delivery models (for-profit vs. nonprofit; joint vs. individual liability, etc.), credit terms, market conditions, and borrower characteristics. The studies here offer impressive variety in settings, and do not produce strong evidence that effects vary substantially across settings, but there is more work to do. Studies of external validity would benefit from greater interplay between theory that generates predictions on where and why microcredit should work best and empirical work that tests those predictions. More broadly, there are probably bridges to build between work on microcredit and other small-dollar credit markets for consumers and their closely-held businesses, and between work on these markets and the many literatures that touch on some aspect of whether and why markets supply credit (in)efficiently (Zinman forthcoming).

Fourth, and closely related, is the identification and interpretation of heterogeneous treatment effects on (potential) borrowers. In particular there is growing concern among policymakers, advocates, and funders that one or more behavioral tendencies leads some, perhaps many, people to do themselves more harm than good by borrowing. It is worth emphasizing that there is scant evidence on how behavioral tendencies actually mediate credit impacts (Zinman forthcoming), and in any case, the presence of behavioral deviations from rationality may in some cases strengthen the case for microcredit rather than weaken it (Banerjee and Mullainathan 2011, Bernheim, Ray, and Yeltekin 2013, Mullainathan and Shafir 2013, Carrell and Zinman 2014). More broadly, we believe that understanding distributional effects is important in a world with growing concerns about debt traps, and here the increasing potential to develop screening and targeting technologies that maximize benefits while minimizing harm offers exciting possibilities.

Fifth, as some of the studies note, we have only scratched the surface of identifying spillover and general equilibrium effects. Much as modest intent-to-treat effects could obscure heterogeneity in effects on different types of borrowers (as discussed directly above), they could also or instead obscure heterogeneity in effects on borrowers and non-borrowers. Non-borrowing businesses could be harmed by business stealing, or benefit from agglomeration. Non-borrowing wage earners could benefit from increased employment opportunities or lower prices, or be harmed if successful borrowers acquire market power. The Morocco study shows one method for pursuing some of these questions, and the Mexico study includes randomized variation in treatment intensity that should allow the identification of spillover effects in future work, but there is much more to do.

Sixth, the studies here identify impacts on marginal but not inframarginal borrowers (e.g., on those who borrowed before the studies here started). This is a strength in the sense that marginal borrowers are the focus of much theory, practice, and policy. But it is a weakness in the sense that impacts on inframarginal borrowers are key to understanding the totality of microcredit's success or failure as a development tool. Different methods will be required to identify impacts on inframarginal borrowers.

Seventh, despite its success in numbers, microcredit institutions could innovate more; in particular, discovering lending models that match more closely to cash flow needs of borrowers may prove more transformative. Field et al. (2013), for example, demonstrates that in India, delaying the initial payment gives borrowers the opportunity to make larger durable investments, and thus improves short- and long-run income for enterprises. In Mali, Beaman et al. (2014) shows that agricultural lending with payments matched to the cash flows of farms (i.e., repayment at harvest, not immediately after loan disbursal in weekly or monthly payments) can lead to increased investment and farm revenues. The nonprofit sector led innovation in microcredit to get the industry to where it is today, and could further lead the industry in exploring innovations to improve the impact on the poor (Karlan 2014).

Finally, we emphasize that the microcredit studies in this volume are silent on the impacts of many other promising noncredit microfinance activities. Many microfinance institutions (MFIs) now focus on savings, not just credit, and the evidence from randomized evaluations of microsavings is quite promising (Karlan, Ratan, and Zinman 2013). MFIs have also been expanding into other payments media (Jack and Suri 2011) and insurance (Mobarak and Rosenzweig 2012, Karlan et al. 2013), making them increasingly bank-like (Burgess and Pande 2005). The impacts of modern microfinance on the lives of the poor and vulnerable are still unfolding, and we hope that researchers will continue working to help identify and shape them.

Bibliography

Banerjee, Abhijit. 2013. "Microcredit Under the Microscope: What Have We Learnt in the Last Two Decades, What Do We Need to Know?" *Annual Review of Economics* 5: 487–519.

Beaman, Lori, Dean Karlan, Bram Thuysbaert, and Christopher Udry. 2014. "Self-Selection into Credit Markets: Evidence from Agriculture in Mali." *Working Paper*.

Bernheim, B. Douglas, Debraj Ray, and Sevin Yeltekin. 2013. "Poverty and Self-Control." *National Bureau of Economic Research*.

Burgess, Robin, and Rohini Pande. 2005. "Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment." *American Economic Review* 95 (3): 780–95.

Carrell, Scott, and Jonathan Zinman. 2014. "In Harm's Way? Payday Loan Access and Military Personnel Performance". Working paper. University of California-Davis, Davis, CA.

Field, Erica, Rohini Pande, John Papp, and Natalia Rigol. 2013. "Does the Classic Microfinance Model Discourage Entrepreneurship Among the Poor? Experimental Evidence from India." *American Economic Review* 103 (6): 2196–2226.

Jack, William, and Tavneet Suri. 2011. "Mobile Money: The Economics of M-PESA." *NBER Working Paper No. 16721*, January.

Karlan, Dean. 2014. "The Role of Nonprofit Organizations in Financial Inclusion." *Stanford Social Investment Review*, 29–35.

Karlan, Dean, Isaac Osei-Akoto, Robert Darko Osei, and Christopher R. Udry. 2013. "Agricultural Decisions after Relaxing Credit and Risk Constraints." *Quarterly Journal of Economics, Forthcoming*.

Karlan, Dean, Adam Osman, and Jonathan Zinman. 2013. "Follow the Money: Methods for Identifying Consumption and Investment Responses to a Liquidity Shock." *National Bureau of Economic Research Working Paper*.

Karlan, Dean, Aishwarya Ratan, and Jonathan Zinman. 2013. "Savings by and for the Poor: A Research Review and Agenda." *Review of Income and Wealth*, October.

Karlan, Dean, and Jonathan Zinman. 2010. "Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts." *Review of Financial Studies* 23 (1): 433–64.

———. 2011. "Microcredit in Theory and Practice: Using Randomized Credit Scoring for Impact Evaluation." *Science* 332 (6035): 1278–84.

Mobarak, Ahmed Mushfiq, and Mark Rosenzweig. 2012. "Selling Formal Insurance to the Informally Insured", February. http://papers.ssrn.com/abstract=2009528.

Mullainathan, Sendhil, and Eldar Shafir. 2013. *Scarcity: Why Having Too Little Means so Much*. New York: Times Books, Henry Holt and Company.

Zinman, Jonathan. forthcoming. "Consumer Credit: Too Much or Too Little (or Just Right)?" *Journal of Legal Studies*, no. Special Issue on Benefit-Cost Analysis of Financial Regulation.

Table 1: Country, Lender, and Loan Information

Study:	Bosnia and Herzegovina (1)	Ethiopia (2)	India (3)	Mexico (4)	Mongolia (5)	Morocco (6)
GDP per capita (PPP USD)[a]	$8,431 (year = 2009)	$507 (year = 2003)	$3,662 (year = 2007)	$14,667 (year = 2010)	$6,109 (year = 2008)	$5,455 (year = 2007)
Annual income (PPP USD)	$19,780	$424	$2,700	$7,828	$1,620	$5,059
Implementing lender	Not revealed	Oromiya Credit and Savings Share Company, Amhara Credit and Savings Institute	Spandana	Compartamos Banco	XacBank	Al Amana
Lender organization type	n/a	Non-profit	For-profit	For-profit	For-profit	Non-profit
Lender amount of microloan debt outstanding in 2012 (USD)[b]	n/a	$91.9 million (Oromiya), $194.6 million (Amhara)	$409.6 million	$1.1 billion	$462.1 million	$229.8 million
Region	14 regions throughout country	2 Western regions	1 major metropolitan area (Hyderabad)	4 regions in North-Central Sonora	5 Northern regions	11 regions throughout country
Rural or Urban?	Both	Rural	Urban	Both	Rural	Rural
Gender of borrowers	Male, Female	Male, Female	Female	Female	Female	Male, Female
Targeted to microentrepreneurs?	Yes (91 percent of respondents planned to invest in new or existing business)	Yes (Plans for starting business considered "salient" criteria for loan grant)	No	Yes (Has business or interested in starting one)	Yes	Yes
Loan eligibility	Sufficient collateral, repayment capacity, crediworthiness, business capacity, credit history, other (inlcuding characteristics)	Poverty status, viable business plan, and other criteria	Women ages 18-59 who have resided in the same area for at least one year and have valid identification and residential proof (at least 80 percent of women in a group must own their home)	Women ages 18-60 with proof of address and valid identification	Women who own less than MNT 1 million ($869 exchange rate) in assets and earn less than MNT 200,000 ($174 exchange rate) in monthly profits from a business	Men and women ages 18-70 who hold a national ID card, have a residency certificate, and have had an economic activity other than non-livestock agriculture for at least 12 months
Sampling frame	Marginal loan applicants considered too risky and “unreliable” to be offered credit as regular borrowers under the terms above	Random selection of households	Households with at least one woman age 18-55 that have resided in the same area for at least three years	Mexican women ages 18-60 who either have a business/economic activity, would start one if they had enough money, or would consider taking credit from an institution	Women who met eligibility criteria and signed up to declare interest in receiving loan from lender	(1) Households deemed likely borrowers, (2) random selection of households
Loan term length	Average 14 months	12 months	12 months	4 months	3-12 months group (average 6 months); 2-24 months individual (average 8 months)	3-18 months (average 16 months)
Repayment frequency	Monthly	Borrowers were expected to make regular deposits and repayments	Weekly	Weekly	Monthly	Weekly, twice monthly, or monthly
Interest rate[c]	22 percent APR	12 percent APR	24 percent APR (12 percent non-declining)	110 percent APR	26.8 percent APR	14.5 percent APR
Market interest rate[b]	27.3 percent APR	24.7 percent APR	15.9 percent APR	145.0 percent APR	42.5 percent APR	46.3 percent APR
Liability	Individual lending	Group (joint liability)	Group (joint liability)	Group (joint liability)	Two treatment arms: group (joint liability) and individual	Group (joint liability)
Group size	No data	No data	6-10 people	10-50 people	7-15 people	3-4 people
Collateralized	Yes (77 percent)	Yes (majority asked to provide)	No	No	Yes (100 percent) for group loans, often for individual loans	No (yes for few individual loans)
Loan loss rate at baseline[b]	No data	0.3 percent (Oromiya), 0.0 percent (Amhara)	2.0 percent	3.2 percent	0.1 percent	0.5 percent
Initial treatment loan size (local currency)	Average 1,653, median 1,500 (2009 BAM)	Median 1,200 (2006 birr)	10,000 (2007 Rs)	Average 3,946 (2010 peso)	Average group: 320,850 (per borrower), average individual: 472,650 (2008 MNT)	Average 5,920 (2007 MAD)
Initial treatment loan size (PPP USD)	Average $1,816, median $1,648	Median ~$500 (median)	$603	Average $451	Average $696 (group), average $472 (individual)	Average $1,082
Loan size as a proportion of income	Average 9 percent, median 8 percent	118 percent	22 percent	6 percent	43 percent (group), 29 percent (individual)	21 percent
Better terms (greater amount and/or lower interest rate) on subsequent loans	No data	No data	Yes	Yes	Yes	No data

[a] Source: World Bank

[b] Source: MIX Market

[c] APR calculated using the upper bound of the interest rate ranges reported for each study (when applicable)

Table 2: Study Information and Results

Study:	Bosnia and Herzegovina (1)	Ethiopia (2)	India (3)	Mexico (4)	Mongolia (5)	Morocco (6)
Baseline survey date	December 2008-May 2009	2003	2005	April 2010 (for panel sample)	February 2008	April 2006-December 2007
Treatment start date	December 2008-May 2009	2003	2006-2007	April 2009 (June 2010 for panel sample)	March 2008	2006
Panel data	Yes	No	Yes	Yes	Yes	Yes
Treatment end date	February-July 2010	March-July 2006	August 2007-April 2008	November 2011-March 2012	September 2009	2009
Unit of randomization	1,196 individual applicants	133 peasant associations	104 neighborhoods	238 clusters (neighborhoods or villages)	40 villages	162 villages
Endline 1 survey sample size	995	6,263	6,862	16,560 (additional cross sectional endline surveys)	964	5,551 (additional endline surveys)
Panel (endline 1) response rate	83 percent (995/1,196)	n/a	74 percent	63 percent (1,823/2,912)	84 percent (964/1,148)	92 percent (4,118/4,465)
Endline 2 survey sample size	n/a	n/a	6,142	n/a	n/a	n/a
Panel (endline 2) response rate	n/a	n/a	89.5 percent	n/a	n/a	n/a
Time between treatment, endline survey 1	14 months	36 months	15-18 months	Average exposure 16 months	19 months	24 months
Time between baseline, endline survey 2	n/a	n/a	39-42 months	n/a	43 months (follow-up with smaller sample - 3 villages)	n/a
Loan take-up in treatment (from study lender only)	100 percent	31 percent	18 percent endline 1, 17 percent endline 2	19 percent	57 percent group, 50 percent individual	17 percent
Repayment rates	46 percent late repayment, 26 percent written off	No data	At endline 1, 43 percent of control group and 49 percent of treatment group were ever late for a payment	90-day delinquency rate of 9.8 percent, default rate about 1 percent	7 percent group, 5 percent individual 90-day individual delinquency rate	No data
Randomization process	Across individuals, after baseline survey	Across clusters, after baseline survey	Across clusters, after baseline survey	Across clusters	Across clusters, after baseline survey	Across clusters, after baseline survey

Figure 1: Interest rate < time preference

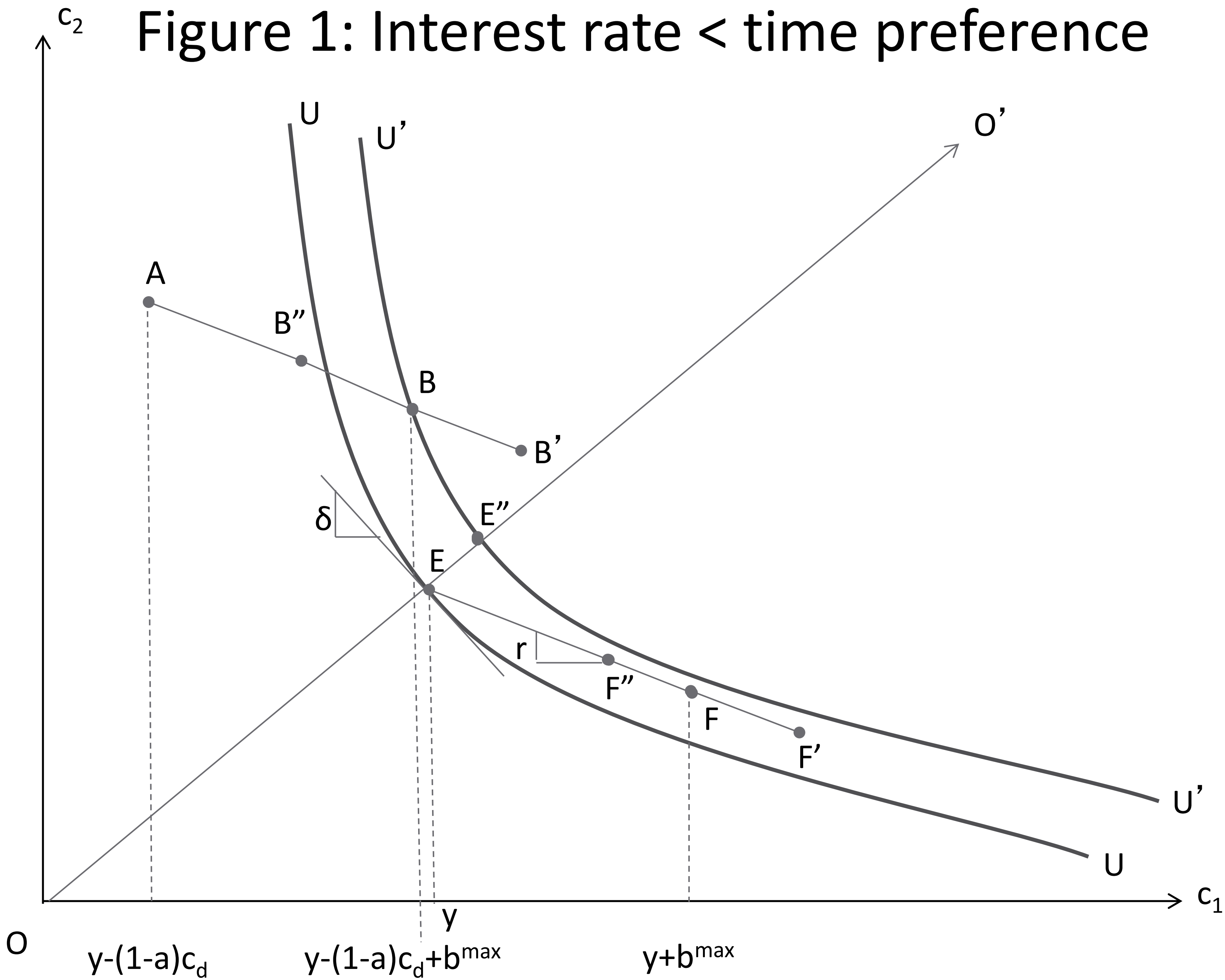

Figure 2: Interest rate > time preference