What do we know about the impact of microfinance? The problems of power and precision[1]

Mahesh Dahal and Nathan Fiala[2]

May 2018

Abstract

Six randomized control trials were published simultaneously in one issue of an economics journal in 2015. The studies show no or minimal impact from providing microloans to clients and have led many researchers and policy makers to conclude that microfinance has been proven to have little or no positive impacts on people's lives. We review in detail these six studies and find three main results. First, unsurprisingly, the insignificant results are replicable using the researcher's original data. Many coefficients are large, but very noisy. Second, every one of the studies is significantly underpowered. This is generally due to low take-up of the financial product offered. Pooling the data from the six studies together improves power for most outcomes, but minimum detectible effect sizes are still very large. Third, when we run analysis on the pooled sample, we find a treatment effect of 29% increase in profits, significant at the 5% level. We also obtain large impacts on business growth and household assets, but not for overall consumption. These results suggest that existing research on the impact of microfinance is generally underpowered to identify impacts, whether modest or zero, reliably. We end by discussing ways to improve future research on this topic.

Keywords: Microfinance; RCTs; replication; power calculations

---

[1] We thank Chris Blattman, David McKenzie, Rachael Meager and Jörg Peters for their helpful comments. Funding for research assistant time for this study was provided by Opportunity International. The results and opinions are those of the authors only.

[2] Dahal: University of Connecticut, Agricultural and Resource Economics, mahesh.dahal@uconn.edu; Fiala (*corresponding author*): University of Connecticut, Makerere University and RWI – Leibniz Institute for Economic Research, nathan.fiala@uconn.edu.

1. Introduction

Microfinance has a long and complicated history with policy makers and researchers. Many initial proponents argued that lack of access to formal finance was a critical part of why people remained poor in developing countries. However, initial reviews of impacts did not show transformative changes for people beyond a few anecdotal stories. The large promises of microfinance solving world poverty were not panning out.

Decades after the beginning of the microfinance movement, there is still little conclusive evidence on the impact of microfinance on the lives of the poor. Any credible attempt at identifying the impact of microcredit on the wellbeing of people needs to overcome the concerns of double selection in credit markets—lenders selecting potential borrowers as well as borrowers self-selecting to borrow. Impact evaluations based on randomized control trials (RCT) have been increasingly used in the past decade to overcome identification issues. However, as we discuss in detail below, findings from RCTs remain inconclusive, mostly owing to low take-up of microfinance products that presents a statistical power challenge for RCT studies of microfinance. Many practitioners still believe in the promise of income growth, transitioning people out of poverty and solving other social issues, such as women's empowerment. However, the expected impacts have become more restrained.

Karlan and Zinman (2011) was one of the first studies to look at the impact of microfinance on poor households using data from an RCT in the Philippines. They found that the microcredit intervention they studied did not lead to bigger businesses, higher income, or higher subjective well-being, but instead resulted in better risk management, fewer businesses, and lower subjective well-being among those who received the microcredit treatment.

Several more recent randomized studies also look at the impact of the traditional microfinance model. Most notably, six randomized studies published in a 2015 special issue of the *American Economic Journal: Applied Economics* have come to a similar conclusion, showing lack of the "transformative" role of microfinance on the lives of poor households. We review the evidence presented in these six studies. We replicate the results of the studies using data provided by the authors, conduct ex-post power calculations to determine what can be reasonably expected from each study, pool the data to run a better powered test, and discuss the different contexts and models they evaluate. Table 1, reproduced from Banerjee, Karlan and Zinman (2015), provides information on key features and summary of each of the six studies.

As expected, we were able to replicate all of the findings in the six individual studies. Contrary to some perceived wisdom, the coefficients for many outcomes were actually very large when compared to control group means, but not statistically significant at traditional levels. This is likely due to serious power issues in each of the studies. The results of ex-post power calculations for the individual studies shows that most are significantly under-powered. The minimum detectible effect (MDE) sizes for main outcomes is up to 230% under perfect compliance, and up to 1,000% under actual compliance. Median (mean) MDE under perfect compliance is 22% (32%) while it is 132% (201%) under actual compliance.

We then combine the data into one sample and run power calculations on the pooled data. Even though the studies were designed independently, endline measures of outcomes were developed to be as comparable as possible across the studies, and so most outcomes can be easily combined. We find that this improves MDEs to between 8% and 44% under perfect compliance for most main outcomes. Using actual compliance rates, we find MDEs of between 31% to 176%. We conclude that the individual studies are significantly underpowered to make inferences on the outcomes they focus on. When combining data, the situation is significantly improved, but is still not ideal.

Because pooling data significantly improves power, we conduct analysis on the full sample, representing over 35,000 participants, running a single OLS regression. Our analysis follows Banerjee et al. (2015b) by assuming a common slope. We include country and region fixed effects to further improve power, weight the samples to account for cross-study imbalance in sample sizes[3], and adjust for timing of the surveys and purchasing power parity differences. We find impacts on business profits of about 29% above the control group, significant at the 5% level. We do not find statistically significant impacts on total consumption, but an increase of household assets of 13%, significant at the 5% level.

Our results from fully pooling the data are consistent with a recent study by Meager (2018), who looks closely at the data from seven randomized experiments, including Karlan and Zinman (2011) and the six *AEJ: Applied Economics* studies. She performs both a Bayesian

[3] We conduct both unweighted and weighted (our preferred approach) analysis, which allows us to control for the different sample sizes across studies. By weighting we are able to treat each study as equal to all others. We believe this is an important adjustment for the samples as some country studies, such as Mexico, represents almost half of the total sample, while India and Morocco are about 1/6 of the sample and Bosnia, where compliance with treatment was best, is less than 3% of the total sample.

hierarchical analysis with partial pooling and an OLS analysis on a fully pooled sample. She does not weight the samples in the fully pooled data and finds a coefficient estimate of 7.3[4]. This is very close to our unweighted coefficient and represents a 20% treatment effect size. However, Meager (2018) argues that fully pooling the samples is not appropriate as each study has different sampling procedures and represents different contexts. She finds a 7% effect size in the hierarchical analysis model and argues that the average effect of microcredit is positive but small relative to control group average levels, with a reasonably high chance of effectively zero impact.

We agree that fully pooling the sample has limitations. However, we see the issue of power as being of more immediate concern. Partial pooling of the data does not alleviate the power issues we highlight in the individual studies. By weighting the studies equally and controlling for region fixed effects, we believe our analysis addresses many concerns raised by Meager (2018), but not all. We thus trade one set of issues for another.

Our work in this paper builds off of others who have looked closely at the studies we discuss here. To the best of our knowledge, McKenzie (2012) was the first to note low statistical power in one of the six studies we discuss here. He finds that a sample size of 15 million would have been needed to obtain the power to identify effect sizes of 10% in the India study (Banerjee et al. 2015a).

While the original authors of the experiments we discuss here and Meager (2018) acknowledge there are issues with statistical power in these studies, we are the first to present clear evidence of the magnitude of the problem, and the first to discuss what this means for our ability to interpret the results. Banerjee, Karlan and Zinman (2015), in the introductory paper to the special issue, note that low take-up rates of credit presents a statistical power challenge to most randomized impact evaluations of microfinance published in the special issue. However, they never quantify just how big of an issue this is.

Our contribution to the literature on the impact of microfinance is to highlight the issue of statistical power more explicitly. Based on our review of the evidence presented in the six experiments we discuss here, we conclude that many of the studies that are presented by some as definitive evidence of the (lack of) effect of microfinance in fact fail to provide any answers, at least individually. But this does not mean that our estimates of impacts from the pooled sample

[4] We replicate the results of Meager (2018) in Appendix Table A5 for comparison and find the same results.

should be taken as the true impact of microfinance. While the most likely reason for our finding impacts when the individual studies did not is improved power, there are still some power issues in the pooled sample. As Gelman and Carlin (2014) argue, in under-powered studies, statistically significant results can still be misleading. In some cases, it is possible to overstate the magnitude of effects, and even get the sign wrong.

The problem of statistical power is pervasive in empirical studies. McKenzie and Woodruff (2013) show significant power issues in all of the 12 experimental studies of business skills training programs they review. Ioannidis, Stanley and Doucouliagos (2017) review 6,700 empirical economics studies and find more than half of them are under-powered.

Unlike the perception among many critics of microfinance, the studies reviewed here do not discredit the role of microcredit in poverty alleviation and improving livelihoods of poor households. Nor does combining the samples together definitively show impacts. What these results do suggest is that the impact of the microfinance programs studied in these experiments is likely not as transformative as it was once believed. But this is often used as a straw man by critics. The initial, unusually high expectations placed on microfinance to tackle mass poverty and fuel sustained economic growth should not be the basis for dismissing the potentially more modest findings of the impact of microfinance on improving livelihoods of poor households. Many of the null results found in the original six studies include economically meaningful effect sizes, but could not be taken as conclusive due to power issues.

This paper proceeds as follows. In section 2 we present the results from our replication of the six individual studies and from pooling the data. In section 3 we present our discussion of statistical power. In section 4 we offer our interpretation of what can reasonably be concluded from existing work on microfinance, including other studies that have been produced more recently. We end in section 5 with a discussion of what we feel is missing from the research on microfinance and what we think an ideal research design would look like.

2. Replication of previous experiments

We present in this section our replication of the main results for the six experimental studies. In addition to conducting a pure replication, we also pool data for common variables in all six studies and run regressions for common outcome variables on the combined dataset. Since there are large differences in sample sizes for the different studies, we report estimates from weighted regressions where each country study has an equal weight. We also provide estimates from unweighted regressions in the appendix. We begin by discussing the studies themselves.

2.1 The six experiments

The six studies we discuss here include experiments in Bosnia and Herzegovina (Augsburg et al. 2015), Ethiopia (Tarozzi, Desai and Johnson, 2015), India (Banerjee et. al, 2015), Mexico (Angelucci, Karlan and Zinman, 2015), Mongolia (Attanasio et al., 2015), and Morocco (Crépon et al., 2015). We obtained the data and analysis code for five of these studies online, downloaded on May 2017. Data and code for Tarozzi, Desai and Johnson (2015) was kindly made available by the authors upon request.

Five of the six microfinance studies reviewed here measure the impact of group liability loans. Attanasio et al. (2015) tested the impact of both group liability and individual liability loans in Mongolia, but they focus their analysis on group liability lending. Augsburg et al. (2015) only looked at individual lending in Bosnia and Herzegovina. Three of the studies were conducted in rural settings, two in urban settings and one in both urban and rural setting. Three studies provided microcredit to women only and three provided to both men and women. Interest rates (APR) varies widely between 12% in Ethiopia to 110% in Mexico.

Except for the Augsburg et. al (2015) study in Bosnia, which randomized loan access at the individual level, all other studies randomized microcredit programs at village, community or neighborhood level. In India, 52 out of 104 urban neighborhoods were randomly selected to be served by a microfinance institute. Tarozzi, Desai and Johnson (2015) uses data from a microcredit and family planning program in Ethiopia that was conducted using a 2 × 2 factorial design where 133 local administrative units (PAs) were randomly assigned to one of four groups: microlending only, family planning program only, both, or none. Their paper focuses on the microfinance component by treating PAs not receiving any microfinance program as control group and the PAs receiving microfinance program as treatment group. In Mexico, 238 geographic clusters that were not served by the partner Microfinance Institution (MFI) were randomized to receive credit access and loan promotion program. In Mongolia, 40 villages were randomized to either receive group liability (15), individual liability (15), and no (control group) microcredit program (10). In Morocco, the evaluation was implemented in 162 villages, divided into 81 pairs of similar villages. The pairs were chosen in areas the partner MFI was planning to start its operation. One village in each pair was randomly selected for the partner MFI to start its lending program.

All six studies measure the impact of increased access to microfinance. Most of these studies conduct an intent-to-treat (ITT) analysis, with a few focusing on a sample of "likely borrowers." In most cases, the average effect of easier access to microfinance on those who are primary targets of microfinance institutions is studied.

It could be the case that the impact of microfinance is greatest for those who are truly credit constrained. If this is the case, then it is not surprising that the studies reviewed here do not find much impact. Four of the six studies were conducted in contexts where majority of the households already had access to microfinance, and the interest rates charged by partner microfinance institutes of these studies were not much better than the market interest rates. We again refer to Table 1 for the summary of the six studies. 58 percent of respondents in Bosnia, 68 percent in India, 54 percent in Mexico, and 57 percent in Mongolia had at least one outstanding loan at baseline. Two out of six studies (Ethiopia and Morocco) were conducted in settings where households seem to have very limited access to microcredit. 13 percent respondents in Ethiopia and less than 25 percent (including 16 percent of households with loans from utility companies) in Morocco reported to have at least one outstanding loan from any source in study areas of these two papers. Net compliance rates for loans from an MFI or a bank are generally very low and vary significantly across countries. Difference in percentage of those with loans in treatment and control groups are 49.3 percentage points (ppts) for Bosnia, 25.2 ppts for Ethiopia, 8.4 ppts for India, 11.5 ppts for Mexico, 25.7 ppts for Mongolia, and 16.7 ppts for Morocco. This net compliance issue will be the main driving force behind the issues with power that we discuss below.

2.2 Results from the individual studies

We now present the results of replicating the individual studies. Because we find broadly the same results as the original authors, we do not present them in a table, but instead describe some of the more important findings.

The first impact we review is whether treatment led to an increase in access to finance. All of the treatments led to increased access to loans from microfinance institutes (MFIs) as compared to control groups. However, access to any type of loan did not significantly increase in treatment areas for some studies. Given that the interest rates provided by the MFIs are very similar to the market interest rates, it is important to observe increase in credit from all sources in order to expect any impact of the microcredit intervention. Take-up of MFI loans varied significantly across studies, with treatment groups in Mexico (6.9), India (8.4), Morocco (9.0) only slightly more likely to have an MFI loan than control group. Ethiopia (25.2), Mongolia (37) and Bosnia (43.9) saw greater access to any MFI loans. Access to any type of loan among the treatment group was higher by only 5.1 ppts in Mexico,7.6 ppts in Morocco, 19.3 ppts in Bosnia, 25.2 ppts in Ethiopia and 25.7 ppts Mongolia. In India, households in treated areas were not any more likely to have a loan from any source as compared to households in control areas.

There is some impact of increased access to microfinance on starting a business. Self-employment activity increased in Mongolia. Investment in business increased in India, Mexico, and Morocco. In India, 15 to 18 months after gaining access to loans, households are no more likely to have at least one business, but they invest more in the businesses they do have or the ones they start. There is an increase in the average profits of the businesses that were already in existence before microcredit intervention. In Mexico, households in treatment areas grow their businesses (both revenues and expenses increase), but there were no corresponding effects on business profits and entry. In Morocco, among those who take up microcredit loans, there are proportionally large average impacts on self-employment investments, sales and profits, although there is great heterogeneity in these effects. Although the impacts on profit are statistically indistinguishable from zero for most countries, the effect sizes are generally large and economically meaningful. For profit outcome, effect sizes as percentage of control mean are 23% in Bosnia, 68% in Ethiopia, 48% in India, 18% in Mongolia, 22% in Morocco, and close to 0% in Mexico.

None of the studies find impact on income. In Mongolia, there is positive impact of microcredit on total consumption as well as in food consumption. However, there is no impact on total consumption as well as food consumption in any other countries. Many of the null intent-to-treat results have confidence intervals that include economically meaningful effect sizes, particularly if one were to scale up the intent-to-treat estimates to infer treatment-on-the-treated effects.

2.3 Results for combining the studies

We next discuss the results of pooling the individual studies into a single dataset. We include country and region[5] fixed effects in the analysis. We control for regional effects even though the original studies do not as this is a more common approach to analysis of RCTs today as it increases statistical power without loss of internal validity. Our results are generally robust to including this control or not.

We obtain significant improvement in the precision of the estimates in the combined dataset. We find large and highly significant treatment effects on profits and revenues. We also find significant increase in the consumption of durable goods. However, we do not find any impact on food and total consumption. We also do not find any impact on assets and wage income.

As we discuss in the next section, some of the outcomes measured are severely underpowered, even when combining data. We are not able to control for baseline covariates, as they differ for each study, which is also a problem for proper inference of results. However, we are able to include region fixed effects for 4 out of 6 countries, which does help improve the precision of some of the estimates.

Results of the estimation using the combined sample is presented in Tables 2-4. Sample size for different outcome variables differ as some studies are missing variables common to other studies. For most of the common variables, we have a sample size of about 35,000 observations. Money figures such as loan amount, income, consumption values are expressed in purchasing power parity (PPP) 2009 USD terms for all countries. Following the schedule in most of the papers, all income, profit, and consumption variables are expressed in fortnightly terms. Estimates presented in Tables 2-4 are weighted to ensure equal weights are placed for each study. Unweighted estimates are presented in the Appendix, Tables A1 to A3.

Table 2 presents combined results on access to credit. Combined results suggest that microfinance treatment results in 13.1 ppts (7.5 ppts unweighted) increase in access to any type of loan, 21.5 ppts (13.1 ppts unweighted) increase in access to program MFI loans among persons/households in treatment areas. Combined results also suggest significant increase in loan

[5] There is no regional variable available for Bosnia and India. We control for woreda (16) fixed effects in Ethiopia, super cluster (45) fixed effects in Mexico, provinces (5 aimags) in Mongolia, and village pair (81) fixed effects in Morocco.

amount among treatment group, mostly resulting from increase in loan amount from program MFIs.

Table 3 presents combined results on the impact of microfinance on wage income and profits from self-employment activities. We find a significant effect (29% of control mean) of microfinance intervention on total business profits and production (revenues). We do not find any significant impacts on assets accumulation, expenses and wage income. Although statistically insignificant, the coefficient on wage income is negative, suggesting that there may be crowding out of income sources as a result of increases in business profits.

Combined results on the impact of microfinance on household consumption are presented in Table 4. We do not find significant impacts on total consumption, food consumption or temptation goods. However, we do find significant positive impact on consumption of durable goods of 13% relative to the control group.

2.4 Interpretation by the original authors

Do the modest take-up rates of microcredit in these studies suggest anything about the effectiveness of microfinance in helping microentrepreneurs grow their businesses and improve consumption? Banerjee, Karlan and Zinman (2015) suggest that the low-take up rates in these studies is the prima-facie case against that notion that microfinance being a panacea for poverty alleviation. Microfinance may not be a panacea for poverty, but these studies also do not have much to say about the relative effectiveness of microfinance compared to other development tools for improving livelihoods of the poor. The six studies considered here in this analysis are not able to show evidence of transformative effects of microfinance on the average borrower. However, the authors also caution the readers that the lack of transformative effects should not obscure other more modest but potentially important effects.

Summarizing the six papers, Banerjee, Karlan and Zinman (2015) come up with the following conclusion: First, the studies do not find clear evidence of transformative effects of reductions in poverty or improvements in living standards. Second, the lack of transformative effect does not mean absence of modest but important effects on investment in business growth. There is convincing evidence that businesses expand, and some evidence that profits increase. Failure of business expansion to translate into improvements in living standards may be because of tradeoff between business and wage income and heterogeneity on profitability (larger firms more profitable than smaller). Third, the lack of transformative effects should not obscure other more modest, but potentially important, effects leading to increase in freedom of choice: improvements in occupational and consumption choice, female decision power, and improved risk management. Fourth, the studies find little evidence of negative effects even in the context of high interest rates. Fifth, the presence of heterogenous effects suggest that the impact of microfinance can be transformative for some. And finally, many of the null results are estimated imprecisely even when these effect sizes are economically meaningful.

Most importantly, Banerjee, Karlan and Zinman (2015) believe that pooling across the studies would yield significant increases in business outcomes. Our results confirm this belief. While 29% effect sizes may not be transformative, they are still quite large relative to studies of other development programs. Our results suggest that microfinance can have important effects on business development and general economic growth. However, as we discuss in the next section, we must be cautious about the pooled sample impacts as even the pooled sample is still under-powered.

3. Power calculations for the six studies

Due to the low take-up of the interventions tested, it is highly likely that there are serious statistical power issues with these studies. We next compute ex-post power calculations and minimum detectable effect sizes for the main outcome variables (credit access, self-employment, business profits, income, and consumption) for each of the six studies. We present minimum detectable effect size (MDE) as percentage of mean of the outcome for control group. MDE are computed for situation with 100 percent compliance as well as for actual compliance rates reported in each of the papers.

Note that there is no consensus on what would constitute a high MDE as this is contingent on expectations of what could reasonably be accomplished by a program. The initial hype about microfinance led some to expect incomes of participants to double in a short period of time. More sober expectations of programs that have a good internal rate of return generally achieve increases in income of 15-20% per year.

We calculate the MDEs using a clustered randomized design with J groups of size n for a given power (κ), significance level (α), and proportion of subjects allocated to treatment group (P), inter-cluster correlation of ρ, and root mean square error (σ) is given by

$$MDE = \frac{M_{j-2}}{\sqrt{P(1-P)*J}}\sqrt{\rho + \frac{1-\rho}{n}} * \sigma$$

where $M_{j-2} = t_{\alpha/2} + t_{1-\kappa}$ for a two-sided test. P is the probability of being assigned in the treatment group.

We also conduct power calculations using the actual compliance rates. If c is the share of subjects initially assigned to the treatment group who receive the treatment and s is the share of subjects initially assigned to the comparison group who receive the treatment, MDE is given by[6]

$$MDE = \frac{M_{j-2}}{\sqrt{P(1-P)J}}\sqrt{\rho + \frac{1-\rho}{n}} * \frac{\sigma}{c-s}$$

All six studies studied in this exercise suffer from imperfect compliance. Net compliance rates for any MFI loans are low for all six studies and range between 6.9 ppts for Mexico and 43.9 ppts for Bosnia. The randomized designs used in each of these studies only influences the probability that someone receives a treatment. Even though these studies seem to have taken into consideration the possibility of imperfect compliance while determining sample sizes, actual compliance was much lower than expected. For example, the partner MFI for Banerjee et. al (2015) experiment in India expected that that 80% of eligible households would borrow. However, only 26.7 percent of the eligible households borrowed from the partner MFI and the net compliance rate ended up being only about 13% for partner MFI loans.

Controlling for baseline covariates helps improve precision of the estimates and thus increases the chances of detecting any effects. The minimum detectable effect sizes (MDE) computed in this exercise do not consider the fact that all six studies controls for baseline covariates and three of the studies control for baseline outcome values. However, in majority of the cases, MDEs are so large that controlling for baseline covariates would not improve the precision of the estimates by enough for the effects to be detected.

[6] See Bloom (2005) and Duflo, Glennerster and Kremer (2007) for further discussion.

Results of power calculations for the individual studies are presented in Table 5. We focus on discussing MDEs for variables in outcome groups for self-employment activities, income and consumption. MDEs under perfect compliance are generally reasonable, ranging between 12 to 53 percent of control mean in Bosnia, 21 to 227 percent in Ethiopia, 7 to 116 percent in India, 8 to 114 percent in Mexico, 2 to 105 percent in Mongolia, and 8 to 61 percent for Morocco. Very high MDE value for profits in Ethiopia is due to high root mean squared error. However, extremely low microfinance take-up rates (or low compliance rates) resulted in extremely large MDEs. MDEs for some variables are as high as 994% of the control group mean. As expected MDEs are largest for profits from self-employment activities for most of the studies.

Under actual compliance rates, MDEs for profits range from 367% in Morocco, 408% in Mongolia, 994% in Mexico, 895% in India, 903% in Ethiopia, and 273% in Bosnia. For Consumption, MDEs are 108% in Bosnia, 51% in India, 50% in Mexico, 7% in Mongolia, and 47% in Morocco. Similarly, MDEs for wage income are 119% in Bosnia, 156% in Ethiopia, 265% (end-line 1) in India, 66% in Mexico, 382% in Mongolia, and 138% in Morocco. These large MDEs, mostly a result of extremely low compliance rates, could explain the null results for many outcomes in these microfinance studies.

We also computed minimum detectable effect sizes for the combined sample in Table 6 (weighted) and appendix Table A4 (unweighted). For the combined sample, we assume one cluster for the Bosnia study, which does not follow a clustered design but instead randomizes at the individual level. This means that the Bosnia study adds very little to our power calculations. The pooled calculation is thus an upper bound for the MDEs.

To compute cluster or group size we took the total sample size in the combined data and divided it by the total number of clusters in all six studies. MDEs for combined sample are still very high as a percentage of control means. This probably explains why we get insignificant treatment effects for some outcome measures even with the combined sample. For the combined sample, MDEs are 11% for engagement in self-employment, 14% for starting a business in the near past, 44 % for wage income, 37% for assets, 32% for profits, and 9% for total consumption when we assume full compliance. With the 25% net compliance, more-or-less the average net compliance rates for the six studies, the MDEs increase by a factor of four.

Banerjee, Karlan and Zinman (2015) make it clear that generating sufficient statistical power is a challenge for randomized evaluations of microcredit, and one of the main caveats of the six studies considered in this meta-analysis is poor statistical power resulting from low take-up of microcredit. Many of the null results in these studies are within the confidence intervals that contain economically meaningful effect sizes of increased access to credit. This motivates them to suggest both additional studies and formal meta-analyses to better understand the impact of microfinance.

Five out of six individual papers considered in this analysis are forthright about one common major caveat in their papers: imprecisely estimated null effects even when the effect sizes are economically meaningful. The Ethiopia study speculates that the reason for the failure to identify statistically significant impacts on key outcomes such as net revenues or livestock ownership is likely insufficient statistical power or measurement error. The reason for the low power is mostly because the data used in the paper come from randomized experiment primarily designed to evaluate effectiveness of family planning programs and microloans on contraceptive choices.

The India paper notes that only a small difference in microfinance take-up between treatment and control areas means that the power and precision of the estimates are significantly lowered. Mexico paper notes that many of the statistically insignificant intent-to-treat estimates are economically meaningful effect sizes. They attribute the lack of precise nulls, even with a relatively large sample size, to a combination of the modest take up differential between treatment and control areas, heterogeneous treatment effects, and high variance and measurement error in outcomes.

To increase take-up rates and thus increase statistical power, the Mongolia paper offered credit to women who had expressed an initial interest in borrowing. Despite this attempt to increase statistical power, the paper still documents some quite substantial but imprecisely estimated impacts. The Morocco paper also designed and implemented a sampling strategy that would improve power to estimate the impact on borrowers, as well as to capture impacts representative at the village level. Still, they encounter issue of statistical power due to extremely low microfinance take-up rates. They do find that the impact on those who are more likely to borrow and those who actually borrow are much larger.

4. What do these studies and other studies of microfinance really tell us?

The six papers we look closely at here, in addition to a lot of other work that has been done on the topic, use random assignment of microfinance to answer the question: What is the impact of increased access to microfinance? Modest take-up differential between the treatment and control areas is one of the main reasons five out of the six study authors mention for the possible null effects of the impact of the access to microfinance. These studies are not able to reject the null of no impact even for some substantially large point estimates that are economically meaningful. The sample size for the Mexico study is relatively large, but this study also had a very low take-up differential between treatment and control neighborhoods and as a result many outcomes are imprecisely estimated null results. The authors attribute highly imprecisely estimated null effects to a combination of the modest take-up differential between treatment and control areas, heterogeneous treatment effects, and high variance and measurement error in outcomes.

It is also important to realize that most of the studies (with exception of Morocco and Ethiopia) were conducted in settings where access to credit were already high, which means that these studies are likely capturing marginal borrowers.

Even in rural Morocco, where the baseline access to credit is relatively low and there are no other sources of formal credit in study areas, take-up of credit is very low. The low take-up of credit also made it difficult for the authors to reject economically significant but statistically insignificant null results. The problem of low take-up differential between treatment and control areas was most severe in urban India, where households had access to other sources of credit. In India, access to MFI loan increases by a modest 8.4 percentage points in treatment areas relative to control areas, but there was no statistically significant increase in access to any loan type in treatment areas.

Outcomes such as profits and revenues are hard to measure and are also characterized by large variances. Thus, measurement errors in these outcome variables also makes it difficult to get precise estimates of the impact of microfinance on these outcomes.

Some evidence from the six studies also suggests that the impact of microfinance is not likely to be homogenous across the population or the prospective borrowers, which also makes it harder to detect impact on the average borrower. In the Ethiopia study, Tarozzi et al (2015) find that areas that were assigned microfinance treatment saw overall increases in earnings from self-employment activities, mostly affecting the right tail of the distributions. Banerjee et al. (2015a) find in India that the impacts on income-generating activities are concentrated in the upper tail of the distribution. They also find that those who choose to borrow are more likely to expand their existing businesses or to start female-owned businesses. Small business investment and profits of preexisting businesses increased, but new businesses were not that profitable.

In the Mexico study, Angelucci, Karlan, and Zinman (2015), examines the extent of heterogeneous effects of microfinance by estimating quantile treatment effects and show stronger effects at the upper end of the distribution for revenues and profits. However, they do not find any noticeable pattern across the distribution for most outcomes.

In Morocco, Crépon et. al (2015) find significant increase in total self-employment profit, although there also appears to be great heterogeneity in these effects. As in other studies they also find that the effect on profits is significantly positive at the higher quantiles but significantly negative at the lower quantiles of profitability. They also find that there are proportionally large average impacts on self-employment investments, sales, and profits among those who take-up microfinance.

Other, more recent studies, have produced evidence for a positive impact of microfinance. Studies that deviate from measuring the impact of traditional or standalone microfinance interventions suggest that a more creative approach to finance can lead to positive impacts for people. The success of microfinance is likely very context dependent and may depend on alleviating additional constraints that operate alongside credit constraints. For instance, Field et al. (2013) find that relaxing constraints to credit structure can have positive impacts. The study uses a field experiment to compare the classic microfinance contract, which requires that repayment begin immediately after loan disbursement, to a microfinance contract that includes a two-month grace period as well as less frequent loan-repayment schemes. They find that the provision of a grace period and less frequent repayment schemes increases short-run business investment and long-run profits.

Burke, Bergquist and Miguel (2017) also look at a modified model of finance and find that the timing of when the loan is given can matter. Finance given when farmers have harvested, rather than before planting as is the common approach, offers farmers inter-temporal arbitrage opportunities. This arbitrage leads to significant increases in farm profits and returns on investment.

It is possible that who receives the microfinance may matter for impacts. Most studies on microfinance, including the six studies we discuss here, look predominately at finance to women. In an experimental study of microloans, cash grants and business skills training where 40% of the sample are men, Fiala (2018) finds that male-owned microenterprises that are provided both access to loans and training report significantly higher profits in the short-run. These loans were also subsidized so that interest rates were 20% rather than the normal 27%. Thinking critically about the model and participants of microfinance programming could prove impactful for people.

Ogden (2016) presents a non-technical discussion of eight microfinance experiments, focusing on the fact that existing studies are measuring microcredit expansion to additional or marginal clients, and so modest results should be expected. He also discusses the need for creative approaches to microcredit designs, for which there is little quantitative evidence. We next present a few ideas on how to improve the evidence for microfinance.

5. Next steps for microfinance research

The results from our analysis of the six studies discussed here suggests that the current evidence for the impacts of microfinance has either been misinterpreted by critics of microfinance, or else is very weak. Clearly, researchers have not been able to find massively transformative impacts from microfinance. While none of the studies discussed here have the statistical power necessary to identify modest impacts, we are able to detect impacts on business profits and revenues when the data from the studies are combined together. But this result must be taken cautiously due to low power even in the pooled sample. In the end, in response to the question we pose in the title of this paper, we believe that researchers and policy makers actually know very little about the impact of microfinance.

What can be done to improve the state of evidence for microfinance? Field, Hollander and Pande (2014) conduct a review of microfinance literature and suggest five ways to better our understanding of the potential impact of microfinance. These suggestions include building more flexibility into the microfinance contract (as in Field et al., 2013), directly encouraging greater business investment (perhaps through training like in Fiala, 2018), using microfinance to build social capital, anticipating and measuring a broader range of development outcomes, and focusing more on the rural population. We believe these are all excellent suggestions that have paid off in the recent literature.

Improving the evidence for microfinance is ultimately based on the need for better designed studies. We end by discussing what an ideal experimental design might look like.

One of the major limitations with many of the studies we discuss is that most utilize an encouragement design. Power challenges are not necessarily inherent to encouragement designs. However, encouragement is unpredictable. Future studies that randomize at the cluster level will need to do a better job of identifying who is likely to take-up loans *ex ante*. This will both improve power, and ensure the ITT estimates are closer to actual impact sizes. The Morocco study estimates a model to predict the likeliness to borrow for each household and categorizes households based on their propensity to borrow. Those who have higher propensity to borrow are slightly more likely to have higher microfinance take-up rates and the impacts are higher for those with higher propensity to borrow. However, this is created *ex post* and suffers from concerns about how well researchers can identify likely borrowers. It also requires significantly more assumptions about the sample than a pure RCT.

Researchers could also consider collecting outcome data at multiple periods. As McKenzie (2012) suggests, taking multiple measurements of noisy outcomes such as business profits, income and expenditures at relatively short intervals allows for averaging out noise, increasing power significantly.

It is also necessary to identify contexts where households are truly credit constrained and borrowing is not frowned upon by local customs. One of the reasons for low take-up of loans even in rural Morocco, where the availability of other credit sources is extremely limited, could be because traditional borrowing models are frowned upon by most followers of Islam. The product must be truly appreciated, or else there will be nothing meaningful to measure.

Of course, the best power situation is obtained when randomization is at the individual level. A few of the studies we discuss above were able to do this. This greatly improves power, but at the cost of increasing the likelihood of spillovers. It is also not feasible in most circumstances, either due to resource constraints, concerns about control individuals simply finding other finance options, or microfinance institutions not willing to turn down eligible applicants.

Researchers also need to be clear about what exactly is being evaluated. Many of the studies here can be described as measuring the expansion of microfinance to either new areas or marginal clients. When individuals have other options for microfinance, it is possible that researchers are simply comparing one microfinance option against other credit options.

It is also important to realize that even if liquidity constraints are a binding constraint for income growth, there may be other constraints that need to be loosened for microfinance to deliver on its promise. For example, providing business training may be important for encouraging poor households to take-up self-employment activities and make them profitable. Fiala (2018) finds male-owned microenterprises that are provided both access to loans and training report significantly higher profits. Thus, microfinance interventions combined with other interventions, like business or skills training, may improve the chances of finding impacts.

Finally, the impacts of microfinance are likely heterogeneous. More studies are needed that allow for estimating meaningful heterogenous effects. This could mean conducting a rich baseline with a relatively large sample size, or doing high quality qualitative and exploratory work before conducting an experiment on a product.

References

Angelucci, Manuela, Dean Karlan, and Jonathan Zinman. "Microcredit impacts: Evidence from a randomized microcredit program placement experiment by Compartamos Banco." *American Economic Journal: Applied Economics* 7, no. 1 (2015): 151-182.

Attanasio, Orazio, Britta Augsburg, Ralph De Haas, Emla Fitzsimons, and Heike Harmgart. "The impacts of microfinance: Evidence from joint-liability lending in Mongolia." *American Economic Journal: Applied Economics* 7, no. 1 (2015): 90-122.

Augsburg, Britta, Ralph De Haas, Heike Harmgart, and Costas Meghir. "The impacts of microcredit: Evidence from Bosnia and Herzegovina." *American Economic Journal: Applied Economics* 7, no. 1 (2015): 183-203.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster, and Cynthia Kinnan. "The miracle of microfinance? Evidence from a randomized evaluation." *American Economic Journal: Applied Economics* 7, no. 1 (2015a): 22-53.

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Pariente, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. “A multifaceted program causes lasting progress for the very poor: Evidence from six countries.” *Science* 348, no. 6236 (2015b).

Banerjee, Abhijit, Dean Karlan, and Jonathan Zinman. "Six randomized evaluations of microcredit: Introduction and further steps." *American Economic Journal: Applied Economics* 7.1 (2015): 1-21.

Bloom, Howard S. "Randomizing groups to evaluate place-based programs." *Learning more from social experiments: Evolving analytic approaches* (2005): 115-172.

Burke, Marshall, Lauren Falcao Betgquist and Edward Miguel. “Selling low and buying high: An arbitrage puzzle in Kenyan villages.” *Working paper* (2017).

Crépon, Bruno, Florencia Devoto, Esther Duflo, and William Parienté. "Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco." *American Economic Journal: Applied Economics* 7, no. 1 (2015): 123-150.

Duflo, Esther, Rachel Glennerster, and Michael Kremer. "Using randomization in development economics research: A toolkit." *Handbook of development economics* 4 (2007): 3895-3962.

Fiala, Nathan. "Stimulating microenterprise growth: Results from a loans, grants and training experiment in Uganda." *Working paper* (2018).

Field, Erica, et al. "Does the classic microfinance model discourage entrepreneurship among the poor? Experimental evidence from India." *American Economic Review* 103.6 (2013): 2196-2226.

Field, Erica, Abraham J. Hollander, and Rohini Pande. "*Micro finance: Points of Promise*." (2014).

Gelman, Andrew, and John Carlin. "Beyond power calculations: assessing type S (sign) and type M (magnitude) errors." *Perspectives on Psychological Science* 9.6 (2014): 641-651.

Ioannidis, John, Tom D. Stanley, and Hristos Doucouliagos. "The power of bias in economics research." *The Economic Journal* 127.605 (2017).

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.

Karlan, Dean, and Jonathan Zinman. 2011. “Microcredit in Theory and Practice: Using Randomized Credit Scoring for Impact Evaluation.” *Science* 332 (6035): 1278–84.

Karlan, Dean, et al. "Agricultural decisions after relaxing credit and risk constraints." *The Quarterly Journal of Economics* 129.2 (2014): 597-652.

McKenzie, David. "Beyond baseline and follow-up: The case for more T in experiments." *Journal of development Economics* 99.2 (2012): 210-221.

McKenzie, David, and Christopher Woodruff. "What are we learning from business training and entrepreneurship evaluations around the developing world?" *The World Bank Research Observer* 29.1 (2013): 48-82.

Meager, Rachel. "Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments." *American Economic Journal: Applied Economics (Forthcoming)*

Ogden, Timothy. "The Case for Social Investment in Microcredit." *Financial Access Initiative*. 2016

Tarozzi, Alessandro, Jaikishan Desai, and Kristin Johnson. "The impacts of microcredit: Evidence from Ethiopia." *American Economic Journal: Applied Economics* 7.1 (2015): 54-89.

Table 1: Loan information and sampling for the six studies

	Bosnia and Herzegovina	Ethiopia	India	Mexico	Mongolia	Morocco
Unit of randomization	1,196 individual applicants	133 peasant associations	104 neighborhoods	238 clusters (neighborhoods or villages)	40 villages	162 villages
Gender of borrowers	41%	13% female household head	100%	100%	100%	7% female household head)
Targeted to Microentrepreneurs?	Yes (91 percent of respondents planned to invest in new or existing business)	Yes (Plans for starting business considered "salient" criteria	No	Yes (Has business or interested in starting one)	Yes	Yes
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	22 percent APR	12 percent APR	24 percent APR (12 percent nondeclining)	110 percent APR	26.8 percent APR	14.5 percent APR
Market interest rate	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
Baseline credit access rate	58.3%	13.1%	68%	53.7%	57.3%	24% (including 16% from utility companies and 6% informal)
Sample size	994	6,263 (endline)	6,811	~16,150	611	4,934
Net compliance rate (Any MFI loan)	43.9 ppts	25.2 ppts	8.4 ppts	6.9 ppts	37 ppts (approx.)	9.0 ppts (approx.)
(Any loan)	19.3 ppts	25.2 ppts	0 (approx.)	5.1 ppts	25.7 ppts	7.6 ppts

Source: Based on Banerjee, Karlan and Zinman (2015) and original studies.

Table 2: Loan take-up for the combined sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Took loan offered		**Has any loan**		**Amount of loan taken**		**Amount of any loan**	
Assigned to treatment	0.205***	0.215***	0.123***	0.131***	262.629***	275.557***	409.084***	432.935***
	(0.019)	(0.019)	(0.019)	(0.017)	(30.537)	(30.384)	(103.644)	(106.195)
N	34,464	34,464	35,842	35,842	34,774	34,774	35,804	35,804
Adjusted R2	0.157	0.182	0.218	0.267	0.078	0.095	0.069	0.067
F-test	120.374	129.871	42.157	62.035	73.966	82.252	15.579	16.620
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	0.049	0.049	0.505	0.505	46.4	46.4	1,472.9	1,472.9
Treatment effect as % of control mean	419.86%	439.27%	24.34%	25.92%	566.04%	593.90%	27.77%	29.39%

Notes: Significance levels: (***1%, **5%, *10%). Standard errors in brackets clustered as in original studies. Regions include: 1 each in Bosnia and India, 16 Woredas in Ethiopia, 45 super clusters in Mexico, 5 provinces in Mongolia, and 81 village pairs in Mongolia. All monetary variables are expressed in 2009 PPP USD terms using data obtained from WDI. Observations are weighted so that the sum of weights is equal for all countries.

Table 3: Income outcomes for the combined sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Assets		**Wage income (2 weeks)**		**Profits (2 weeks)**		**Revenue (2 weeks)**		**Expenses (2 weeks)**	
Assigned to treatment	-784.882	-824.766	-95.924	-121.415	11.003**	12.200***	30.418**	32.356**	15.767	16.313
	(710.025)	(747.386)	(108.687)	(112.088)	(4.998)	(4.297)	(13.496)	(12.666)	(10.484)	(10.127)
N	29,893	29,893	35,785	35,785	35,047	35,047	34,895	34,895	35,072	35,072
Adjusted R2	0.342	0.339	0.025	0.029	0.021	0.029	0.020	0.027	0.014	0.018
F-test	1.222	1.218	0.779	1.173	4.847	8.060	5.079	6.525	2.262	2.595
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	24,097.4	24,097.4	387.5	387.5	41.8	41.8	143.1	143.1	95.0	95.0
Treatment effect as % control mean	-3.26%	-3.42%	-24.75%	-31.33%	26.31%	29.17%	21.26%	22.61%	16.60%	17.17%

Notes: Significance levels: (***1%, **5%, *10%). Standard errors in brackets clustered as in original studies. Regions include: 1 each in Bosnia and India, 16 Woredas in Ethiopia, 45 super clusters in Mexico, 5 provinces in Mongolia, and 81 village pairs in Mongolia. All monetary variables are expressed in 2009 PPP USD terms using data obtained from WDI. Observations are weighted so that the sum of weights is equal for all countries.

Table 4: Consumption outcomes for the combined sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total consumption (2 weeks)		**Food consumption (2 weeks)**		**Durable consumption (2 weeks)**		**Temptation goods (2 weeks)**	
Assigned to treatment	22.979	26.801	3.454	3.595	5.260*	5.771**	-1.029	-1.253
	(26.075)	(25.552)	(4.971)	(5.050)	(2.833)	(2.524)	(1.133)	(1.137)
N	13,357	13,357	29,854	29,854	13,311	13,311	29,792	29,792
Adjusted R2	0.035	0.062	0.245	0.262	0.056	0.054	0.071	0.078
F-Stat	0.777	1.100	0.483	0.507	3.448	5.230	0.824	1.214
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	341.8	341.8	155.2	155.2	45.5	45.5	22.1	22.1
Treatment effect as % of control mean	6.72%	7.84%	2.22%	2.32%	11.55%	12.68%	-4.66%	-5.68%

Note: Significance levels: (***1%, **5%, *10%). Standard errors in brackets clustered as in original studies. Regions include: 1 each in Bosnia and India, 16 Woredas in Ethiopia, 45 super clusters in Mexico, 5 provinces in Mongolia, and 81 village pairs in Mongolia. All monetary variables are expressed in 2009 PPP USD terms using data obtained from WDI. Observations are weighted so that the sum of weights is equal for all counties.

Table 5: Ex-post minimum detectable effect sizes (MDES)

					Full Compliance		Actual Compliance	
Country	Variable description	Inter-cluster correlation	Treatment effect	Control mean	MDE	% Control mean	MDE	% Control mean
Bosnia	Asset value	0.033	-4388.22	111229.30	22940.13	21%	118860.79	107%
	Profit (respondent business)	0.009	671.87	2902.89	1529.89	53%	7926.88	273%
	Any self-employment income	0.033	0.06	0.67	0.08	12%	0.42	63%
	Self-employment income (profit)	0.004	74.50	6122.05	1648.41	27%	8540.98	140%
	Wage income	0.028	322.89	6896.86	1577.90	23%	8175.64	119%
	Total consumption per capita	0.022	-647.88	4166.78	867.40	21%	4494.32	108%
	Food expenditure	0.031	-4.10	117.47	15.53	13%	80.47	69%
Ethiopia	Started business in last 3 years	0.015	-0.02	0.07	0.02	21%	0.06	84%
	Net revenues (profits)	0.000	525.80	146.24	332.61	227%	1319.87	903%
	Self-employment (profits)	0.000	513.31	755.06	346.36	46%	1374.46	182%
	Wage income	0.031	48.66	293.86	115.69	39%	459.10	156%
India	Assets (stock)	0.009	597.51	2497.55	1132.11	45%	8708.55	349%
	Profit	0.001	354.34	744.90	866.41	116%	6664.71	895%
	Started a business (12 months)	0.003	0.01	0.05	0.02	35%	0.13	270%
	Self-employment (profit)	0.001	354.34	744.90	866.41	116%	6664.71	895%
	Wage income	0.221	-526.35	2988.03	1027.94	34%	7907.19	265%
	Total consumption per capita	0.038	-12.11	524.67	34.81	7%	267.78	51%
	Food expenditure per capita	0.025	-8.79	84.29	14.05	17%	108.06	128%
Mexico	Profits in the last two weeks	0.000	0.11	145.47	166.20	114%	1445.22	994%
	Started a business (12 months)	0.006	-0.01	0.10	0.02	15%	0.13	132%
	household business income	0.014	58.27	839.82	186.43	22%	1621.10	193%
	HH wage income	0.020	-29.79	4540.71	346.46	8%	3012.73	66%
	Value of assets	0.016	-1533.71	8318.57	1437.92	17%	12503.66	150%
	Amount spent on food (weekly)	0.040	3.65	874.26	50.45	6%	438.69	50%
Mongolia	Assets (stock)	0.065	-29.29	2236.46	692.37	31%	2694.05	120%
	HH business profit	0	-4.79	-26.85	15.94	-59%	62.02	-231%
	Respondent business profit	0	-7.85	-12.11	12.69	-105%	49.40	-408%
	Respondent started business	0.001	0.01	0.07	0.06	90%	0.23	350%
	Self-employment (profit)	0	-4.79	-26.85	15.94	-59%	62.02	-231%
	Wage income	0.012	-252.82	413.86	406.78	98%	1582.81	382%
	Total consumption per capita	0.115	0.11	10.95	0.21	2%	0.82	7%
	Monthly HH food expenditure	0.142	0.14	10.34	0.24	2%	0.92	9%
Morocco	Assets	0.129	1448.42	15984.40	4420.52	28%	26470.17	166%
	Profit	0.041	2004.55	9055.73	5546.59	61%	33213.13	367%
	Has a self-employment activity	0.186	-0.02	0.83	0.06	7%	0.37	44%
	Wage income	0.049	446.69	27669.28	6378.01	23%	38191.65	138%
	Self-employment profit	0.041	2004.55	9055.73	5546.59	61%	33213.13	367%
	Total monthly HH consumption	0.096	-45.65	3057.00	302.64	10%	1812.21	59%
	Monthly HH food consumption	0.138	2.76	1784.02	139.66	8%	836.30	47%

Table 6: Minimum detectable effect sizes for the pooled sample

Variable description	Intra-cluster correlation	Treatment effect	Control mean	Full Compliance MDE	Full Compliance % Control mean	25% Compliance MDE	25% Compliance % Control mean
Wage income (2 weekly)	0.070	-95.924	387.535	170.60	44%	682.41	176%
Assets	0.375	-784.882	24097.390	8871.45	37%	35485.81	147%
profits	0.013	11.003	41.826	13.51	32%	54.04	129%
Revenues	0.051	30.418	143.107	45.07	31%	180.27	126%
Expenses	0.035	15.767	95.003	33.61	35%	134.43	142%
Self-employment activity	0.225	0.024	0.455	0.05	11%	0.19	42%
Started a business	0.014	0.005	0.079	0.01	14%	0.04	56%
Total consumption	0.108	22.979	341.771	31.74	9%	126.95	37%
Food Consumption	0.167	3.454	155.250	12.00	8%	48.00	31%
Durable Consumption	0.040	5.260	45.528	9.18	20%	36.73	81%
Temptation goods	0.068	-1.029	22.065	2.07	9%	8.30	38%

Notes: Observations are weighted so that the sum of weights is equal for all countries.

Online Appendix

Table A1: Loan take-up for the combined sample without weighting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Took loan offered		**Has any loan**		**Amount of loan taken**		**Amount of any loan**	
Assigned to treatment	0.127***	0.131***	0.066***	0.075***	132.913***	140.440***	225.127**	255.752**
	(0.009)	(0.009)	(0.012)	(0.010)	(12.494)	(12.378)	(96.334)	(109.780)
N	34,464	34,464	35,842	35,842	34,774	34,774	35,804	35,804
Adjusted R2	0.065	0.095	0.165	0.209	0.026	0.043	0.071	0.068
F-test	183.019	209.466	31.500	59.620	113.169	128.733	5.461	5.427
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	0.045	0.045	0.505	0.505	38.0	38.0	1,469.6	1,469.6
Treatment effect as % of control mean	280.62%	289.74%	13.06%	14.84%	349.83%	369.64%	15.32%	17.40%

Notes: Significance levels: (***1%, **5%, *10%). Standard errors in brackets clustered as in original studies. Regions include: 1 each in Bosnia and India, 16 Woredas in Ethiopia, 45 super clusters in Mexico, 5 provinces in Mongolia, and 81 village pairs in Mongolia. All monetary variables are expressed in 2009 PPP USD terms using data obtained from WDI. Observations are not weighted, so each country study is not weighted equally.

Table A2: Income outcomes for the combined sample without weighting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Assets		**Wage income (2 weeks)**		**Profits (2 weeks)**		**Revenue (2 weeks)**		**Expenses (2 weeks)**	
Assigned to treatment	-131.889	-131.686	-8.811	-15.129	5.531	8.705**	19.722*	22.251*	13.072	13.044
	(157.321)	(172.110)	(12.126)	(13.296)	(4.754)	(3.836)	(11.133)	(11.966)	(9.733)	(10.373)
N	29,893	29,893	35,785	35,785	35,047	35,047	34,895	34,895	35,072	35,072
Adjusted R2	0.384	0.382	0.034	0.040	0.007	0.013	0.021	0.027	0.014	0.016
F-test	0.703	0.585	0.528	1.295	1.354	5.150	3.138	3.458	1.804	1.581
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	4826.0	4826.0	188.0	188.0	27.7	27.7	105.6	105.6	78.3	78.3
Treatment effect as % control mean	-2.73%	-2.73%	-4.69%	-8.05%	19.95%	31.40%	18.68%	21.07%	16.70%	16.66%

Notes: Significance levels: (***1%, **5%, *10%). Standard errors in brackets clustered as in original studies. Regions include: 1 each in Bosnia and India, 16 Woredas in Ethiopia, 45 super clusters in Mexico, 5 provinces in Mongolia, and 81 village pairs in Mongolia. All monetary variables are expressed in 2009 PPP USD terms using data obtained from WDI. Observations are not weighted, so each country study is not weighted equally.

Table A3: Consumption outcomes for the combined sample without weighting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Total consumption (2 weeks)		**Food consumption (2 weeks)**		**Durable consumption (2 weeks)**		**Temptation goods (2 weeks)**	
Assigned to treatment	5.059	5.852	1.805	0.208	3.534**	3.616**	-0.772	-1.263***
	(8.696)	(7.097)	(2.703)	(1.910)	(1.724)	(1.485)	(0.495)	(0.446)
N	13,357	13,357	29,854	29,854	13,311	13,311	29,792	29,792
Adjusted R2	0.035	0.065	0.124	0.151	0.045	0.044	0.046	0.059
F-Stat	0.338	0.680	0.446	0.012	4.201	5.929	2.429	8.003
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	312.3	312.3	165.4	165.4	23.8	23.8	21.5	21.5
Treatment effect as % of Control Mean	1.62%	1.87%	1.09%	0.13%	14.82%	15.17%	-3.58%	-5.86%

Notes: Significance levels: (***1%, **5%, *10%). Regions include: 1 each in Bosnia and India, 16 Woredas in Ethiopia, 45 super clusters in Mexico, 5 provinces in Mongolia, and 81 village pairs in Mongolia. All monetary variables are expressed in 2009 PPP USD terms using data obtained from WDI. Observations are not weighted, so each country study is not weighted equally.

Table A4: Minimum detectable effect sizes for pooled sample without weighting

Variable description	Intra-cluster correlation	Treatment effect	Control mean	Full Compliance MDE	Full Compliance % Control mean	25% Compliance MDE	25% Compliance % Control mean
Program MFI loan	0.17	0.21	0.05	0.03	65%	0.13	261%
Any loan	0.24	0.12	0.51	0.05	9%	0.19	38%
Program MFI loan amount	0.10	262.63	46.40	52.12	112%	208.48	449%
Any loan amount	0.08	409.08	1472.95	480.99	33%	1923.94	131%
Wage income (2 weekly)	0.07	-95.92	387.53	170.60	44%	682.41	176%
Assets	0.37	-784.88	24097.39	8871.45	37%	35485.81	147%
profits	0.01	11.00	41.83	13.51	32%	54.04	129%
Revenues	0.05	30.42	143.11	45.07	31%	180.27	126%
Expenses	0.03	15.77	95.00	33.61	35%	134.43	142%
Self-employment activity	0.22	0.02	0.45	0.05	11%	0.19	42%
Started a business	0.01	0.00	0.08	0.01	14%	0.04	56%
Total consumption	0.11	22.98	341.77	31.74	9%	126.95	37%
Food Consumption	0.17	3.45	155.25	12.00	8%	48.00	31%
Durable Consumption	0.04	5.26	45.53	9.18	20%	36.73	81%
Temptation goods	0.07	-1.03	22.06	2.07	9%	8.30	38%

Notes: All monetary variables are expressed in 2009 PPP USD terms using data obtained from WDI. Observations are not weighted, so each country study is not weighted equally.

Table A5: Replication of Meager (2018) pooled OLS results

	Profits (2 weeks)	**Expenses (2 weeks)**	**Revenue (2 weeks)**	**Total consumption (2 weeks)**	**Durable consumption (2 weeks)**	**Temptation goods (2 weeks)**
Assigned to treatment	7.208*	12.938*	22.374**	4.574	2.288	-0.637***
	(4.334)	(7.699)	(8.745)	(2.873)	(13.492)	(0.216)
N	35,303	35,303	35,303	30,830	14,224	30,706
Adjusted R2	0.031	0.058	0.086	0.009	0.118	0.236
F-Stat	2.765	2.824	6.546	2.535	0.029	8.702
Coefficients reported in Meager (2018)	7.3	13	22.5	4.6	2.3	-0.6
Control mean	35.4	96.9	139.9	298.5	91	13.1
Treatment effect as % of Control Mean	20%	13%	16%	2%	3%	-5%

Notes: Significance levels: (***1%, **5%, *10%). This table recreates the analysis conducted in Meager (2018) and so we use her PPP adjustments and clustering for standard errors. Observations are not weighted, so each country study is not weighted equally.