What do we know about the impact of the microfinance?

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 these identification issues. However, 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.

Eight recent randomized studies, six randomized studies published in a 2015 special issue of the *American Economic Journal: Applied Economic,* Karlan and Zinman (2011), and Fiala (2018) form the basis of our knowledge on the impact of the traditional microfinance model. Most of these studies have come to a similar conclusion, showing lack of the "transformative" role of microfinance on the lives of poor households. A recent working paper reviews the evidence presented in the six studies that are most comparable with one another. The six papers closely analyzed in the review paper, in addition to the few other papers that delve into this specific topic, use random assignment of microfinance to answer the question: What is the impact of increased access to microfinance? The review paper replicates the results of the studies and conducts ex-post power calculations to determine if the studies have sufficient statistical power to distinguish tightly identified null results from imprecisely estimated insignificant results for which the researchers cannot rule out treatment effects with confidence. The paper also pools the data from all six studies to run a better powered test and discusses the different contexts and models these studies evaluate.[1]

Authors of the original studies also acknowledge the issue of low power. Five out of the six papers suggest that the null (rather than negative) effects on the impact of access to microfinance could be due to low take-up differential between the treatment and control areas. Low statistical power resulting from low take-up rates mean that these studies are not able to reject the hypothesis that microfinance does not have an impact, even though some point

[1] The context and the models of microfinance evaluated in these six studies is presented in Annex Table 1.

estimates are substantially large and are economically meaningful. Additionally, many of the highly imprecisely estimated null effects can be attributed 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.

Careful reading of the papers reveals that the authors of the studies reviewed here do not discredit the role of microcredit in poverty alleviation and improving livelihoods of poor households. This is in contrast to many critics' perceptions of microfinance, who likely base their opinion on a cursory look at the results without critically engaging the issue of low power and cautious interpretations of the results from the authors.

Although not statistically significant at traditional levels, the coefficients for many outcomes in these studies are actually very large when compared to control group means. 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 coefficients 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. This means that the treatment group could demonstrate high levels of change and results would still be considered inconclusive and statistically insignificant.

Pooling data from similar studies can improve statistical power and help identify overall impact that individual studies may fail to detect. However, pooling data does not necessarily guarantee sufficient power to detect impact. In this case, when combining data, the situation is significantly improved, but is still not ideal. Running power calculations on the pooled data, MDEs improves to between 8% and 44% under perfect compliance for most main outcomes. Under actual compliance rates, MDEs increases to between 31% and 176%. This suggests that even the pooled sample may not have sufficient power to detect impact on some of the outcomes.

Because pooling data improves power to some extent, the paper conducts analysis on the full sample, representing over 35,000 participants, running a single ordinary least squares (OLS) regression.[2] The pooled analysis includes country and region fixed effects to further improve power, weights the samples to account for cross-study imbalance in sample sizes,[3] and adjusts for timing of the surveys and purchasing power parity differences. Results of pooled analysis show impacts on business profits of about 29% above the control group, significant at the 5% level. No statistically significant impacts were found on total consumption, but there was a 13% increase of durable consumption, significant at the 5% level.

One of the major limitations of pooling the data is the concern of heterogeneity in the sampling and context across studies, as raised in Meager (2018). By weighting the studies equally and controlling for region fixed effects, the pooled analysis addresses some, but not all, concerns raised by Meager (2018). This strategy allows for the review paper to both address the heterogeneity concerns and focus on the issue of power, which—they believe—is of more immediate concern. The problem of statistical power is pervasive in empirical studies. McKenzie and Woodruff (2013) show significant power issues in all 12 of the 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. As the review paper shows, the issue of underpowered RCT studies of microfinance is quite serious. However, measures can be taken to design studies that are sufficiently powered to successfully measure the impact of microfinance on the lives of the poor.

What have we learnt?

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. 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. The initial, lofty expectations placed on

[2] OLS is a type of linear least squares method for estimating the unknown parameters in a linear regression model. It can be used to model a dependent variable in terms of its relationships to a set of explanatory variables and conduct linear statistical test.

[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.

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.

Previous access to credit matters: It is important to realize that most of the studies (with exception of Morocco and Ethiopia) were conducted in settings where access to credit was already high, which means that these studies are likely capturing marginal borrowers, that is those who borrowed after implementation of the microfinance experiment, but likely already had access to some form of credit prior to the experiment, as well. Impact estimates obtained from experimental designs that capture the effect of microfinance on marginal borrowers are likely to understate the average impacts of microfinance (Wydick 2016). Future studies of microfinance would benefit from implemention in areas where credit availability is low and households are truly credit constrained. Substituting one form of credit with another is unlikely to result in any measurable impacts unless the new forms of credit are substantially cheaper, easier to access, or more tailored to the needs of the study population.

Need for better powered studies: Future studies of microfinance should be carefully designed so that these studies are well powered to capture even modest impact of microfinance. Power can be significantly improved if studies are carefully designed to target microfinance products to those who are in most need of it and also most likely to borrow. 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. Individual randomization improves power, but it may also make it harder to identify true impacts if there are spillover effects of the treatment onto the control group. Individual randomization is also not feasible in most circumstances, either due to resource constraints, concerns about control individuals simply finding other finance options, or microfinance institutions unwilling to turn down eligible applicants.

One of the major limitations with many of the studies we discuss is that most utilize an encouragement design.[4] Power challenges are not necessarily inherent to encouragement designs.

[4] Ecouragement designs do not directly provide treatment program to individuals, but make it easier for individuals in treatment area to use the program, known as an "intention to treat" (ITT). Opening a microfinance branch in a village does not necessarily mean that everyone in the village *get* microfinance loans, but it does make individuals in the village *more likely* to take microfinance loans.

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 that the intention to treat (ITT) estimates are closer to actual impact sizes.

Need to understand what is being evaluated: Researchers also need to be clear about what exactly is being evaluated. Two otherwise identical studies, one measuring the impact of microfinance in areas with no prior exposure to microfinance and another measuring the impact of the expansion of microfinance to either neighboring areas or to marginal clients, are likely evaluating completely different questions. When a study population has other options for microfinance, it is possible that researchers are simply comparing one microfinance option against other credit options. Also, as Banerjee, Karlan, and Zinman (2015) indicates and Wydick (2016) shows, it is possible that impacts on the borrowers and/or communities that did not have pior exposure to credit may be larger than the impacts on those already being served by other credit options before the lenders in these microfinance studies start their lending intervention.

Measuring impact may still be possible where other access to credit exists because even if lack of credit is 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.

Need for rich baseline data and multiple rounds of data collection: The impacts of microfinance are likely context dependent, thus more studies that allow for estimating meaningful heterogenous effects are needed. For example, this could mean conducting a rich baseline with a relatively large sample size. Collection of data at multiple rounds also help increase precision of the estimated impacts. For economic outcomes such as profits, income, and expenditure that are measured with high variability, taking multiple measures of these outcomes at reasonable intervals and averaging these multiple measurements when estimating treatment effects could also be quite beneficial (Mckenzie 2012).

To conclude, the six studies reviewed here do not discredit the role of microcredit in poverty alleviation and improving livelihoods of poor households, nor does combining the studies' samples together definitively show impacts because of null results due to low take-up and statistical power issues. Thus, addressing this question of the impact of microfinance will require future studies that take heed of these challenges and respond with careful selection of study location and research design.

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

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.

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

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

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)*

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

Wydick, Bruce. "Microfinance on the margin: why recent impact studies may understate average treatment effects." *Journal of Development Effectiveness* 8.2 (2016): 257-265.

Annex 1

Annex 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% 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	Random selection of households	Households with at least one woman age 18–55 that have resided in the	Mexican women ages 18–60 who either have a business/ economic activity, would start one if they had	Women who met eligibility criteria and signed up to declare interest	(1) Households deemed likely borrowers, (2) random selection of households

	credit as regular borrowers under the terms above		same area for at least three years	enough money, or would consider taking credit from an institution	in receiving loan from lender	
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
Study timeframe	Baseline: May 2009 Endline: Feb-July 2010	Baseline: Jan-May 2003 Endline: March-July 2006	Baseline: 2005 Endline1: 15-18 months later (2007/08) Endline2: 3 years later (2009/10)	Baseline: April-June 2010 Endline: November -March 2012	Baseline: March 2008 Endline: October 2009	Baseline: April-May 2006 Endline: May 2008 - January 2009
Take-up of MFI loan for treatment group	76.3%	31.2%	17.8%	20.7%	57%	16.7%