Why the debates that the recent microcredit RCTs are meant to close are going to be endless

We have been handed down a fresh round of evidence on the impact of microcredit from six randomized control trials (RCTs) undertaken by eminent economists (see summary paper [here](#)). The results are what most reasonable people would expect: that microcredit is useful for many but by no means transformational. Perhaps the more novel finding is that the impact on average is zero, not negative.

So does this finally settle the debate then, some forty years after Muhammad Yunus´s first (non-random) experimentation with microcredit? Alas, I don't think so, though it will certainly affect how arguments are pitched henceforth. There are four reasons why evidence from RCTs alone will not close the debate.

First, RCTs are pure empiricism, devoid of any theory. Give a bit of something to some, withhold it from others, and compare. So when the results come out, all one can do —indeed, what the dozens of online commenters on these studies have done— is to interpret them by retrofitting them into one´s own prior theories and mental models. How did the impact happen? Why the differences between men and women, young and old, richer and poorer? Would this work elsewhere in the same way, or if the interest rate were lower? But all we have learned is that six concrete programs are not impactful, on average. The rest is inference and theorizing. We are back to extrapolating on the basis of isolated (now n+6) data points.

Second, RCTs measure impacts on a number of variables: consumption, investment, schooling, female empowerment, job creation, etc. Results are often a mixed bag, so we back up into the problem of how to total up these impact categories to get a net-net impact. There´s now the inevitable suggestion that we should be measuring how people´s happiness is affected, as a sort of grand bottom line. How can we hope to measure ultimate impact if we lack even a basic definition of what we mean by impact?

Third, there´s the question of what we are supposed to make of averaged impacts. We know that debt is not right for everyone, so why would or should debt be held to an *average* standard of goodness? Surely debt is good for some people, in some circumstances, at certain times. What´s the point of diluting —and possibly cancelling— the benefits to these people statistically? The issue is one of targeting interventions, not passing across-the-board judgments on them.

Finally, for all the sophisticated statistical tools deployed, it all hinges on the quality of the data collected. Given that RCTs are fundamentally an empirical exercise, it is odd that in the dozens of online commentaries on the six studies there hasn´t been much discussion of the empirics itself. The research papers are often not written for people like me to understand (have you tried reading them?), so I am relegated to being a *user* of the abstract, introduction, and conclusions sections. But what I can

read —what we can all process— is the survey questionnaire. I would urge all those who feel they need to have an opinion on these studies to start by reading the questionnaires. Take [this one](), for example. In your mind, is this is fit for the purpose of evaluating the impact of microcredit? Does this meet the [lean research standards proposed by Kim Wilson]()? Try to have the questions answered tonight by your partner or spouse. Then try to imagine what kind of data you´d get if you were going around asking these questions of perfect strangers, showing up at their doorstep unannounced.

The suggestion that such empirical studies can settle the impact question of microcredit is just as naïve as the prior suggestion that the best way to get poor people out of poverty is to get them into recurrent cycles of debt. I agree entirely with the findings from these RCTs. Was anyone really hopeful that the six RCTs would turn out otherwise?