

# Not Everything That Can Be Counted Is Counted

## Rethinking Development Econometrics

SNEHA MENON

There is, at present, little discourse and theory in academia about how to tackle the quick and dirty research needs of the development sector, which includes localised decision-making for programmes. There is a need to recognise qualitative approaches and move towards subjective interpretations of research results.

“Jaise maine pucha hain ... in me se kaunsa?” (Out of the options I have listed, which one?). The call operator asked the question politely, forcing the respondent to choose from the options that she had just enlisted—options that were there only for her reference. Much to my researcher colleague’s dismay, the call centre representative had unwittingly committed research anathema by “leading” her respondent. My colleague did not have sufficient funding to visit the 5,000 subjects of her impact study who were spread across the country. Nonetheless, she was expected to run a quantitative study by her supervisor and by the standards of her discipline. After all, her background in economics had ensured that she had little tolerance for any methodology that involved fewer respondents than what the sample size calculator advised. She argued that larger sample sizes would nullify the call centre’s confirmation bias, and that this methodology was still statistically superior to unrepresentative qualitative research based on five to 10 interviews.

I would like to clarify at the onset that this article is not a critique of randomised controlled trials (RCTs) or any other experimental/quasi-experimental methods of impact evaluation. That debate has had its day in the sun.<sup>1</sup>

Like my colleague, my training in economics has convinced me of the scientific grounding of RCTs, instrument variables (IV), propensity score matching (PSM), regression discontinuity design (RDD) and other textbook prescriptions, and how each of these are resistant to different forms of extraneous influence in their efforts to prove causation. But, this theory may be inapplicable to the development sector, as there is, perhaps, more noise than can be controlled for. The best-known objection to RCTs and their

ilk is that these are not externally valid; that is, their context-specific results are only useful to policymakers in the same geographic, demographic, and temporal setting as the original study. This is why there has been a flurry of studies that evaluate the impact of similar interventions. *3ie* impact, for instance, lists 4,260 impact evaluations in its repository and the number is clearly growing (Miranda et al 2016). Yet, their generalisability depends on one’s taste for variation (Vivalt 2016). The second concern is whether these would stand the test of replication. Replication can be a can of worms (Blattman 2015) and can produce striking results, as for instance in the Psychology Reproducibility Project, which could produce only half the mean effect size of the original effects (Roberts and Nosek 2015).

### Research for Development

It is worth questioning whether the considerable growth of quantitative impact evaluations is just another sign of the growing mathematisation of economics. After all, the standard economics graduate course hardly outlines how to design a research study when there is little scope for randomisation, no baseline, and scant funds and time; all of which are conditions that are commonplace in the social sector. These are subjects explored by the World Bank in an Independent Evaluation Group working paper (Bamberger 2006). Clearly, this is not a contest of methods to find the gold standard, but rather, it is an effort in prioritisation and capacity building. There is, at present, little discourse and theory in academia about how to tackle the quick and dirty research needs of the development sector, which includes decision-making about which programmes should be offered in a village, how to customise policy entitlements to different populations, and how to identify beneficiaries. These decisions could be research-driven but are often taken on the fly, based on intuition, experience, or a political agenda.

Perhaps, it is time to quantify and qualify internal and external validity, so that social enterprises can choose between multiple rounds of small surveys and a single comprehensive study of the

I am very grateful to Jean Drèze, Anita Kumar and Anirudh Tagat for their comments.

Sneha Menon ([snehamenon88@gmail.com](mailto:snehamenon88@gmail.com)) is an MPhil student in economics at the University of Oxford and is a co-founder of Insights Applied.

population of its beneficiaries. Is the presence of a third party necessary, or would it be possible for a non-governmental organisation (NGO) to train its own staff for data collection (which can be audited later) and save on the travel expenses of researchers? In my colleague's case, would she have been better off doing a few in-person interviews or having the call centre try to reach all 5,000 of the intervention beneficiaries? Ignoring the empirical trade-offs for a moment, the qualitative route is arguably more useful for building intuition as well as for understanding and validating genuine research needs.

How well is the current evaluation landscape meeting its intended users' needs?<sup>22</sup> Why do so many conferences end with a call for greater alignment of academic research and decision-making in the development sector? The idea that research should be guided by a field needs assessment is perhaps very consequentialist and may conflict with the idea of intellectual freedom and academic curiosity. But, are these really at odds with one another? The economist Jean Drèze (2002) discussed this question:

The value of scientific research can, in many circumstances, be enhanced even further if it is combined with real-world involvement and action. I see this approach as an essential complement of, not a substitute for, research of a more "detached" kind.

As we have seen, it remains a pressing question even today.

### Role of the Economist

It is interesting to assess how economists view their own role in development. One common perspective is that they see themselves as data opportunists who make the best use of an already available data set and extract useful insights. This is why we are encouraged to look for "natural experiments." But alternative approaches may start with some of the following questions: "Where are the biggest research lacunae?" and "What are practitioners' biggest research needs?" or "How to optimally use given research resources (data and/or funding and/or interns) to best help a decision-maker?"

When another researcher colleague asked his guide something along these

lines, he was promptly asked whether he wished to do a PhD after the current project or look for work, insinuating that a future PhD application would require research experience working with more "rigorous methods," and so thinking along the lines of research needs would be fruitless. Why is it that the ability to identify real research needs is not counted as a PhD-worthy skill?

While many papers outline the need for further research, and there are evidence gap maps too,<sup>3</sup> there are few formal mechanisms to match researchers to practitioners in India, particularly economists, who are instead required to only utilise data sets. Is it too idealistic to envision the pairing of consenting economics PhD candidates with consenting practitioners?

At the theoretical end, new discourses have emerged on the development of iterative methodologies for programme design and evaluation (Pritchett et al 2012). The development consulting sector has also produced toolkits for customising monitoring and evaluation (M&E)<sup>4</sup> and for developing lean data surveys.<sup>5</sup> But, mixed or qualitative methods are yet to find their way into mainstream economics syllabi. The words "social audit"<sup>6</sup> do not come up in empirical methodologies. Given that the unspoken social science pecking order is based on "quantitativeness," economics takes pride in its place as the top discipline and has a strong methodological influence on other disciplines, like sociology and political science. Another worrying trend is seen in the paper acceptance standards of economics journals, where few Indian journals like *Economic & Political Weekly* continue to publish qualitative studies.

### What Is the Solution?

I am not arguing either for or against the idea that the findings of studies with larger samples and more controls come closer to the truth. However, it is, in the palette of greys that is development economics literature that I see more space for qualitative shades. Specifically, I enlist some steps worth considering:

(i) Those who wish to do research of consequence on policy implementation should be encouraged to source their

research question from the very practitioners they wish to influence. I would go on to say that research questions should be developed by the very people on whom the research will have an impact. Certainly, unemployed youths, on being offered a training course by a social enterprise, would be able to assess the usefulness of an auctions-based study to check their "ability to pay" for these courses.

(ii) While knowing how to calculate the Coefficient of Relative Risk Aversion (CRRA) is useful, it is equally important for economics students to learn how to create a survey under certain constraints. For instance, they should learn how to manage the trade-off between including more respondents per village and more villages for a particular study. Perhaps, this could be achieved by using census data to measure heterogeneity or by taking into account that travelling to more villages would involve higher costs.

(iii) Studies need to be judged not just by the stars in the tables,<sup>7</sup> but by other measures, such as whether the findings were revalidated by qualitative checks post-analysis, or if the study was designed in a particular way to assist a practitioner.

(iv) We need a spectrum of formal research that would fill the gap between academic journals and government-commissioned reports. There are very few Indian research forums that are frequented by both academics and practitioners, apart from opinion pieces in newspapers. Forums where NGOs can learn about the experiences of other organisations with a particular programme, or where researchers across the country aggregate and cross-validate their research findings, are imperative.

Our objective should be to overcome our aversion to subjectivity in the interpretation of results. After all, a degree of subjectivity is commonplace in "statistically airtight" studies too. What else could explain the lengthy arguments on endogeneity and reverse causality, to which walls of economics seminar rooms have been witness? Why not, then, come to terms with the presence of ambiguity? As New School economist Sanjay Reddy puts it, "Judgment will necessarily be

involved in such an exercise. This is not an embarrassment but rather the very condition of confronting reality” (2012: 66).

## NOTES

- 1 See Cohen-Setton et al (2014) for a good summary. For more on the debate, see Deaton (2009), Pritchett and Sandefur (2013), and Blattman’s (2013–17) blog for criticism; and for a good defence, see Imbens (2009), Banerjee and Duflo (2008), and Newman’s (2012–14) blog.
- 2 This question is addressed at length in Shah et al (2015).
- 3 3ieimpact, “Evidence Gap Maps,” <http://www.3ieimpact.org/en/evaluation/evidence-gap-maps/>.
- 4 Innovations for Poverty Action, “Goldilocks: Right Fit M&E,” <http://www.poverty-action.org/goldilocks>.
- 5 Acumen, “Lean Data Addresses the Unique Measurement Needs of Social Enterprises,” <http://acumen.org/ideas/lean-data/>.
- 6 Loosely speaking, this is the process of verifying the stated/promised outcomes of a policy.
- 7 Stars are often used to signify the statistical significant findings in an analysis table.

## REFERENCES

Bamberger, Michael (2006): “Conducting Quality Impact Evaluations Under Budget, Time and

Data Constraints,” Independent Evaluation Group (IEG) Working Paper Series, World Bank, Washington DC.

Banerjee, Abhijit V and Esther Duflo (2008): *Economics: The Experimental Approach to Development*, Massachusetts: The National Bureau of Economics Research.

Blattman, Chris (2015): “Worm Wars,” Chris Blattman website, <https://chrisblattman.com/tag/worm-wars/>.

— (2013–17): “Randomized Trials,” Chris Blattman website, <http://chrisblattman.com/tag/randomized-trials/>.

Cohen-Setton, Jérémie, Emmanuel Letouzé and Adrien Lorenceau (2014): “The Popularity of Randomized Control Trials,” *Bruegel*, <http://bruegel.org/2014/01/blogs-review-the-popularity-of-randomized-control-trials>.

Deaton, Angus (2009): “Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development,” Princeton: Research Program in Development Studies, Center for Health and Wellbeing, Princeton University.

Drèze, Jean (2002): “On Research and Action,” *Economic & Political Weekly*, Vol 37, No 9.

Imbens, Guido W (2009): “Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua,” Department of Economics, Harvard University, Cambridge, Massachusetts.

Miranda, Jorge, Shayda Sabet and Annette N Brown (2016): “Impact Evaluations Still on the

Rise,” 3ieimpact blog, <http://blogs.3ieimpact.org/is-impact-evaluation-still-on-the-rise/>.

Newman, Kirsty (2012–14): “Randomized Controlled Trials,” *Kirsty Evidence*, <https://kirstyevidence.wordpress.com/tag/randomized-controlled-trial/>.

Pritchett, Lant and Justin Sandefur (2013): “Context Matters for Size: Why External Validity Claims and Development Practice Don’t Mix. Lant Pritchett, Justin Sandefur,” Working Paper 336, Centre for Global Development.

Pritchett, Lant, Salimah Samji and Jeffrey Hammer (2012): *It’s All About MeE: Using Structured Experiential Learning (“e”) to Crawl the Design Space*, Princeton: University of Princeton University.

Reddy, Sanjay (2012): “Randomise This! On Poor Economics,” *Review of Agrarian Studies*, Vol 2, No 2, p 66.

Roberts, Russ and Brian Nosek (2015): “Brian Nosek on the Reproducibility Project,” *Ecotalk*, [http://www.ecotalk.org/archives/2015/11/brian\\_nosek\\_on.html](http://www.ecotalk.org/archives/2015/11/brian_nosek_on.html).

Shah, Neil Buddy, Paul Wang, Andrew Fraker and Daniel Gastfriend (2015): “Evaluations with Impact: Decision-focused Impact Evaluation as a Practical Policymaking Tool,” Working Paper 25, 3ie International Initiative for Impact Evaluation, London.

Vivald, Eva (2016): *How Much Can We Generalize from Impact Evaluations?* Stanford: Stanford University.