# Econometric debate on randomized controlled trials (RCTs)

So far, 2009 has been a fertile time for methodological debates in econometrics. One hot debate touches directly on randomized control trials (RCTs), a methodology often used in impact evaluations of development interventions. On one side, renowned Princeton development economist Angus Deaton argues that randomized experiments are overhyped and that other methods of learning about impacts provide guidance which is often more closely related to theory. On the other side, Guido Imbens reminds readers of the reasons why randomized experiments have gained a wide following. Other methods rely on assumptions that often make them not quite fully convincing.

Randomized experiments are a big deal in international development research and impact evaluation because they provide the promise of obtaining clean measures of the causal impact of interventions. Angus Deaton, however, gives three counter-arguments. First, evaluating a particular intervention is useful for its promoters and beneficiaries, but might not yield much theoretical perspective on why the project worked. Second, experiments provide an estimate of the *average* impact of the intervention, though it might not be the quantity of most interest. The median impact, or the distribution of impacts, can be as interesting, or even more so. Third, the analysis of data produced by experiments often involve advanced statistics and hence is subject to the same pitfalls as structural models, such as small sample biases or specification search (a.k.a. "data mining").

Guido Imbens replies in a new working paper. His first argument is that having several methodologies at our disposal is better than restricting our sources of evidence to only one particular method of analysis. More information, from more research projects employing various methodologies, he argues, will always be more useful than less information from a limited number of methodologies. Analysts and policymakers who are provided a bigger picture can balance the advantages and disadvantages of the various pieces of evidence, both in terms of internal validity (the confidence that the intervention led to the impact measured) and external validity (the ability to generalize the findings from one study to other settings).

Second, Imbens disagrees with Deaton's refusal to grant randomized experiments a particular place in the "hierarchy of evidence." Imbens forcefully argues that randomized experiments are at the top of that hierarchy by virtue of their claims to internal validity. Clearly, randomized experiments cannot answer all interesting questions in international development and in economics. But the kind of work being done by FAI researchers shows that they can be rigorous and theoretically interesting (see here, here, and here).

The debate seems to come down to the weight put on the internal validity of causal estimates versus their intrinsic interest and wider applicability. When push comes to shove, we suspect that Deaton would be happy to see more RCTs in general. What he doesn't like is arguments that say "if it's not an RCT, the evidence doesn't count." Fair enough. We think Imbens would agree.

May 14, 2009