

Comment: Martin Ravallion

Randomized Trials and Development Policy

Measure what is important, don't make important what you can measure.

—Robert McNamara, president of the World Bank, 1968–1981

Randomized controlled trials (RCTs) are on the menu of options for development impact evaluation. That is not news, for it has been true for at least 40 years.¹ What has changed over the past 10–15 years is the academic popularity of RCTs. The chapter by Banerjee, Duflo, and Kremer (BDK) describes and reflects on the expanding use of RCTs in development economics. The authors have been at the forefront of this change.

In theory, the idea of an RCT is simple enough. Access to the program is randomly assigned to some units, with others set aside as a control group. The impact is then estimated by the difference in the sample mean outcomes between treated and control groups. This converges toward the true mean impact in the population as the sample sizes increase.

In practice, RCTs are rarely perfect, their internal validity is rarely assured, and their external validity is often questionable. As argued by Deaton and Cartwright (2018), these limitations do not appear to be well understood among practitioners. It does not help that prominent advocates often make unguarded claims that exaggerate the virtues of RCTs. For example, it is clearly not true that “any difference between treatment and control units reflects the impact of the treatment,” as BDK say, because there is always some experimental error (including, of course, sampling error).

1. The earliest development RCT that I know of was done in 1978 by the World Bank and was published in 1981, namely, Jamison et al. (1981).

The concerns go deeper. Not even the theoretical rationale for randomization is as clear as advocates claim. Indeed, quite generally, there exists a deterministic (nonrandom) assignment of treatment status (based on continuous covariates) that minimizes the expected error variance, as shown by Kasy (2016). This holds for a given sample size. Comparing methods, it makes more sense to fix the budget for the evaluation than to fix the sample size. RCTs can be costly. With a given budget, RCTs will often have lower sample sizes than are possible with observational studies (OSs). An OS can then turn out to be closer to the truth in practice, even if it comes with a bias (Ravallion 2018).

Has the new popularity of RCTs in development research helped inform development policy making? That is not the only reason we might do RCTs; another is to better understand how an economy works—to identify key structural parameters. However, policy making is an important reason. BDK clearly agree. Indeed, that is explicitly the goal of the premier institution for promoting RCTs in development, namely, the Abdul Latif Jameel Poverty Action Lab (J-PAL), founded by two of the authors. On the bio page of Banerjee and Duflo (2011), it is said that “J-PAL’s mission is to reduce poverty by ensuring that policy is based on scientific evidence.” (“Scientific evidence” can be taken as code for RCTs.) J-PAL and other advocates of RCTs have framed their task as that of figuring out what works and what does not, to scale up the former and scale down the latter. Is that what is happening now?

To inform antipoverty policy making, researchers ideally should be filling the gaps between what we know about the effectiveness of policies and what policy makers need to know. As economists, we should first ask ourselves: Why do such gaps exist? Imperfect information plays a role. Here the problem is that development practitioners cannot easily assess the quality and expected benefits of an evaluation, to weigh against the costs. Compared to the complex econometric methods used in some OSs, the simplicity of an RCT helps practitioners understand what is being done. However (as already noted), that understanding is not always as deep as it needs to be for practitioners to properly assess the lessons from an RCT, including its limits.

There are also important externalities. The benefits of an evaluation are rarely confined to that specific project but instead spill over to future projects. These external benefits are probably greater for OSs than for RCTs,

for which external validity has been a recurrent concern (see, for example, Pritchett and Sandefur 2015). In addition, current project managers cannot be expected to take proper account of the external benefits to other projects when deciding how much to spend on their own project's evaluation. Thus there may well be an underinvestment in OSs, which generate more externalities, relative to RCTs.

Knowledge gaps also stem from misalignments of evaluative effort. One aspect is that development evaluators often ignore the scope for fungibility. Recipients (governmental or not) can reallocate their own efforts in response to new funding, such as development aid. As a consequence of such fungibility, donors are often implicitly supporting something else and evaluating the wrong program from the point of view of assessing their impact. Then evaluative efforts are not aligned well with development efforts. This applies as much to RCTs as to OSs.

Methodological preferences on the part of evaluators can reinforce such misalignments, and here the emphasis on RCTs may well be hurting our progress in addressing important knowledge gaps. There are both *output* and *substitution* effects of the RCT boom. The output effect is obvious, as documented by BDK. The substitution effect relates to the methods used. There has been a marked increase in the share of journal articles on development economics that use RCTs. But that is not where a methodological substitution is worrying; instead, it is in policy evaluation. We have seen a marked switch in favor of RCTs in institutions such as the World Bank. The Bank's own Independent Evaluation Group reports that more than 80 percent of the impact evaluations starting in 2007–2010 used randomization, compared with 57 percent in 2005–2006 and only 19 percent in prior years (World Bank 2012).

A problem in overall policy evaluation stems from the fact that randomization is clearly only feasible for a nonrandom subset of policies and settings, so we lose our ability to comprehensively address our knowledge gaps (Ravallion 2009, 2018). For example, it is rarely feasible to randomize the location of medium- to large-scale infrastructure projects and sectoral and economy-wide reforms, which are core activities in almost any poor country's development strategy. Indeed, the very idea of randomized assignment is antithetical to the goals of many development programs, which typically aim to reach certain types of people or places. Governments will (hopefully) be able to do better in reaching poor people than a random

assignment would. Randomization also tends to be better suited to relatively simple programs, with clearly identified participants and nonparticipants, rather short time horizons, and little scope for the costs or benefits to spill over to nonparticipants.

There are both supply and demand sides to this misalignment. On the supply side, the reality today is that graduate students and their teachers are wandering around looking for something they can randomly assign. If randomization is not feasible for the question being posed, then researchers are often drawn to ask other questions. Governments in the developing world are having a harder time finding someone to help evaluate those public programs for which randomization is not a feasible option.

The potential biases go further. On the demand side, governments (and development agencies) are largely free to choose what gets evaluated. Even when they agree to an RCT, they can choose those programs for which they do not care what the verdict will be. Other programs will not get evaluated in equilibrium. (And, as noted, they may include what was really being funded by aid.) The risks are plain.

If we are really concerned about obtaining reliable estimates of the impact of the portfolio of development policies, we should choose a representative sample from that portfolio and then find the best method for each of the selected programs/policies, with randomization as only one of a number of options. That is not what is happening now.

References

- Banerjee, Abhijit V., and Esther Duflo. 2011. *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*. New York: Public Affairs.
- Deaton, Angus, and Nancy Cartwright. 2018. "Understanding and Misunderstanding Randomized Controlled Trials." *Social Science and Medicine* 210: 2–21.
- Jamison, Dean, Barbara Searle, Klaus Galda, and Stephen P. Heyneman. 1981. "Improving Elementary Mathematics Education in Nicaragua: An Experimental Study of the Impact of Textbooks and Radio on Achievement." *Journal of Educational Psychology* 73 (4): 556–567.
- Kasy, Maximilian. 2016. "Why Experimenters Might Not Always Want to Randomize, and What They Should Do Instead." *Political Analysis* 24: 324–338.
- Pritchett, Lant, and Justin Sandefur. 2015. "Learning from Experiments When Context Matters." *American Economic Review: Papers and Proceedings* 105 (5): 471–475.

Ravallion, Martin. 2009. "Should the Randomistas Rule?" *Economists' Voice* 6 (2): 1–5.

Ravallion, Martin. 2018. "Should the Randomistas (Continue to) Rule?" Working Paper 492, Center for Global Development, Washington, DC.

World Bank. 2012. *World Bank Group Impact Evaluations: Relevance and Effectiveness*. Independent Evaluation Group, World Bank, Washington, DC.

This is a section of [doi:10.7551/mitpress/11130.001.0001](https://doi.org/10.7551/mitpress/11130.001.0001)

The State of Economics, the State of the World

Edited by: Kaushik Basu, David Rosenblatt,
Claudia Sepúlveda

Citation:

The State of Economics, the State of the World

Edited by: Kaushik Basu, David Rosenblatt, Claudia Sepúlveda

DOI: 10.7551/mitpress/11130.001.0001

ISBN (electronic): 9780262353472

Publisher: The MIT Press

Published: 2020



The MIT Press



This work is available under the Creative Commons Attribution—NonCommercial—NoDerivatives 3.0 IGO license (CC BY-NC-ND 3.0 IGO) <http://creativecommons.org/licenses/by-nc-nd/3.0/igo>.

Some rights reserved

The findings, interpretations, and conclusions expressed in this work are those of the authors and do not necessarily reflect the views of The World Bank, its Board of Executive Directors, or the governments they represent. The World Bank does not guarantee the accuracy, completeness, or currency of the data included in this work and does not assume responsibility for any errors, omissions, or discrepancies in the information, or liability with respect to the use of or failure to use the information, methods, processes, or conclusions set forth. The boundaries, colors, denominations, and other information shown on any map in this work do not imply any judgment on the part of The World Bank concerning the legal status of any territory or the endorsement or acceptance of such boundaries.

Nothing herein shall constitute or be construed or considered to be a limitation upon or waiver of the privileges and immunities of The World Bank, all of which are specifically reserved.

Attribution—Please cite the work as follows: The World Bank. 2019. *The state of economics, the state of the world* / edited by Kaushik Basu, Claudia Sepulveda, and David Rosenblatt. Published by MIT Press. © World Bank. License: Creative Commons Attribution—NonCommercial—NoDerivatives 3.0 IGO (CC BY-NC-ND 3.0 IGO).

Third-party content—The World Bank does not necessarily own each component of the content contained within the work. The World Bank therefore does not warrant that the use of any third-party-owned individual component or part contained in the work will not infringe on the rights of those third parties. The risk of claims resulting from such infringement rests solely with you. If you wish to re-use a component of the work, it is your responsibility to determine whether permission is needed for that re-use and to obtain permission from the copyright owner. Examples of components can include, but are not limited to, tables, figures, or images.

All queries on rights and licenses should be addressed to the Publishing and Knowledge Division, The World Bank, 1818 H Street NW, Washington, DC 20433, USA; fax: 202-522-2625; e-mail: pubrights@worldbank.org.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Names: Basu, Kaushik, editor. | Sepúlveda, Claudia Paz, 1969– editor. | Rosenblatt, David, editor.

Title: *The state of economics, the state of the world* / edited by Kaushik Basu, Claudia Sepulveda, and David Rosenblatt.

Description: Cambridge, MA : MIT Press, [2019] | Includes bibliographical references and index.

Identifiers: LCCN 2018046336 | ISBN 9780262039994 (hardcover : alk. paper)

Subjects: LCSH: Economic development. | Information technology—Economic aspects. | Monetary policy. | Social change.

Classification: LCC HD82 .S8223 2019 | DDC 330.1—dc23

LC record available at <https://lcn.loc.gov/2018046336>

10 9 8 7 6 5 4 3 2 1