Patterns of randomization, geographic concentration and funding in the recent development economics literature

Jessica Leight

February 1, 2022

1 Introduction

Over the last twenty years, randomized controlled trials have grown rapidly in economics, leading to a robust debate about their methodological and philosophical advantages and disadvantages. Authors including the three recent Nobel laureates recognized for their contribution to the development of randomized controlled trials have documented the substantial growth of RCTs in the years roughly 1990 to 2015 (Banerjee et al., 2016), and recent papers have noted the rapid growth in the absolute number of RCTs sponsored by institutions such as the Abdul Latif Jameel Poverty Action Lab (Ravallion, 2020). This shift has also generated a countervailing phenomenon of objections and claims that the discipline is now only focused on randomized controlled trials, or has an insufficient focus on broader institutional questions or important policy questions that are not easily answered using field trials (Deaton, 2010; Rao, 2020).

At the same time, there has been relatively little systematic documentation of the funding sources of RCTs — typically, multiyear projects that are funded by grants from research or programmatic funders to universities or research organizations — and the geographic focus of both RCT and non-RCT papers. While several published papers have documented the overrepresentation of the U.S. (and to a lesser extent Europe) in the economic literature more broadly (Das et al., 2013; Fontana et al., 2019; Hirvonen, 2020), the relative representation of different developing countries has not been much examined.

The objective of this paper is to report several key stylized facts about research in development economics as published over the last two years in a cross-section of economics journals. I focus on three key dimensions: the percentage of development economics

papers that report field randomized controlled trials; the representation of various developing countries in the published literature (both RCTs and non-RCTs); and the number and type of funding sources reported by published RCTs. The economics journals examined include what are colloquially known as top five journals (the American Economic Review, the Quarterly Journal of Economics, Econometrica, the Journal of Political Economy, and the Review of Economic Studies); other top general interest journals (including the Review of Economics and Statistics, the Economic Journal, the Journal of the European Economic Association, and all four American Economic Journal field journals); and a set of top field journals in development economics (the Journal of Development Economics, Economic Development and Cultural Change, and the World Bank Economic Review). I focus on papers published over roughly the past two years (from October 2019 to October 2021), yielding a full sample of 611 papers.

The analysis highlights three key stylized facts. First, randomized controlled trials are far from constituting the majority of published work in development economics, accounting for only 19% of papers in the pooled sample. This percentage is, however, almost twice as high in the sample of papers published in top five journals (36%). Second, the composition of countries represented in development economics research is heavily weighted toward China and India, the most commonly observed countries (constituting 14% and 12% of all papers in the sample) and larger middle-income countries in Latin America (Mexico and Brazil are the next most observed countries, constituting 6% and 3% of published papers, respectively). By contrast, 21% of published papers focus on the entire continent of sub-Saharan Africa.

Third, the composition of research funders is varied across different journal tiers. Within the sample of 119 papers reporting randomized controlled trials, 106 report funding (and many report multiple funding sources). The two most commonly reported sources in the pooled sample are the International Growth Centre (cited by 21% of the published RCTs), and the World Bank, cited by 20%. In the sample of papers published in top five journals, the pattern is similar but much more pronounced: among the 22 papers reporting randomized controlled trials, 50% of them are funded by the International Growth Center, 27% by the Abdul Latif Jameel Poverty Action Lab, and 23% by the Economic and Social Research Council in the U.K.

2   Methodology

The sample of papers constituted is drawn from a cross-section of journals publishing work in development economics. For the general interest journals, I conducted a search in the database EconLit using a search strategy designed to identify paper published focusing

on lower and middle income countries; the search strategy can be found in Appendix A. These papers were additionally manually screened to identify some that did not meet the search criteria. In this screening process, a small number of papers (20 in total) were screened out: these were generally papers analyzing the effects of developing countries' economic shocks on developed countries; papers in the international trade literature; and papers that use historic data from Europe or other contexts. For the field journals, I included all papers published in the specified date range, assuming that that any paper published in the specified journals is recognized as part of the development economics literature.

I then manually screened the abstracts (and, if necessary, the paper introductions) in order to code papers as randomized controlled trials, and to identify the site of the research. There was rarely ambiguity as to whether a given paper was a randomized controlled trial; lab experiments and survey experiments in which the only outcomes were measurement outcomes or variables observed in the context of an artificial experiment were excluded. Analyses of baseline or descriptive data from a randomized trial were also excluded. However, papers analyzing a randomized mechanism of assignment implemented by a government were included, irrespective of whether the mechanism was explicitly designed as part of a randomized evaluation.

Similarly, the research site was typically unambiguously identified in the abstract of the papers analyzed or, failing that, the introduction. Papers that were reporting only theory were coded as theory, and papers reporting cross-country analyses or multicountry analyses were coded as multicountry. In some cases, papers reported on analyses of data from two discrete sites (e.g., a randomized controlled trial conducted at two sites), in which case both countries were recorded.

Finally, for the papers coded as randomized controlled trials, I extracted the information provided by the authors about funding. This information is typically provided in an acknowledgment footnote on the title page, but in the cases where information was not provided in the footnote, I did searches of the key words ("funding" and "acknowledge") to check whether funding information was reported in other parts of the paper. I did not consult any other sources such as authors' websites or trial registries, focusing only on the information reported in the text of the paper. The acknowledgment text was extracted in its entirety, and then coded to correspond to the funders highlighted. In many cases, this includes acknowledgment of multiple sources, all of which are counted, and funding sources are aggregated for all funders that are reported by at least two separate papers.

Table 1 shows a table of the funders tracked. In some cases, these funders are very easily defined (i.e., USAID). In some cases, to facilitate aggregation, I coded multiple funding sources in a category, and have specified below the specific institutions that com-

prise this category; this strategy was used primarily to aggregate across various sources of government research funding provided by Australia, Europe, China, and the U.S. Importantly, a number of these organizations (particularly the IGC, 3ie, J-PAL, and various mechanisms at the World Bank) also have their own funders; thus a particular evaluation may be funded by J-PAL through a mechanism that is itself funded by other foundations or bilateral donors. This information is very rarely reported by the authors (and may or may not in fact be available to them). Accordingly, I have not attempted to trace back sources of ultimate funding, but have focused on the organizations actively providing grants.

It is important to note that acknowledgments very rarely report the amount of the grant provided or how this compares to the total cost of the research. Accordingly, each contribution is effectively weighted equally in this exercise, and all funding is coded as a binary variable. The only exception is when a paper reports multiple funding sources within a category I have aggregated (i.e., if grants from both the NSF and the NIH are reported, this is recorded as value two in the other U.S. government sources category).

3   Results

First, I calculate the share of published papers that are randomized controlled trials across journal sources. In the full set of 611 papers, 61 are from top five journals, 112 from other general interest journals, and 438 from field journals. As noted above, in the pooled sample 19% of published papers are RCTs; this is 36% in top five journals, and roughly comparable to the mean (15% and 18%) in the top general interest and field journal sample, respectively.

Second, I identify the five most represented countries across each category of journal and calculate the share of published papers constituted by multicountry and theory papers. The patterns are relatively consistent and are summarized in Table 2: as noted above, research focused on China and India is dominant (with China the most common country in all samples other than the top five, where India is the most common). The other countries represented are drawn from Asia (Indonesia, Pakistan, and Bangladesh) and middle-income countries in Latin America (Mexico and Brazil), with only one country observed from sub-Saharan Africa (Uganda, in the field journal sample). Theory papers are relatively rare, but multicountry papers constitute around 15% of the sample except in the top five journals.

In the randomized controlled trial sample, India and Bangladesh again are observed to be relatively dominant, but interestingly these are followed by three sub-Saharan African countries (Uganda, Kenya and Nigeria). This constitutes suggestive evidence

that the sample of countries represented in randomized controlled trial papers may be geographically broader. Three additional sub-Saharan countries are also observed in the next five countries most commonly represented in the RCT sample— Malawi, Zambia, and Ghana — in addition to China and Mexico, robustly represented in the non-RCT sample as well.

To further contextualize the patterns of publishing concentration vis-à-vis the size and poverty level of the countries analyzed, I link the sample to World Bank data reporting the population and the poverty ratio for specified poverty lines, focusing on the international poverty line of $1.90 a day in 2011 purchasing parity power dollars (commonly known as the dollar a day standard) and the poverty line of $5.50 a day in 2011 PPP terms, used by the World Bank to denote upper middle income countries. The key patterns are summarized graphically in Figure 1.

Figures 1a and 1b show the numbers of papers published in the sample normalized by the number of people living below the $1.90 poverty line as reported in the most recent year of data available; the population numbers are calculated per 100,000,000 poor people to render the graphs more tractable. In Figure 1a, we see the 11 countries that are most represented in the sample of published papers. It is evident that Mexico and China are relatively overrepresented compared to the size of their poor population, while India and Nigeria are substantially underrepresented, at 23 and 15 papers per one hundred million poor residents, respectively. Figure 1b then shows the ten countries observed in the sample of published papers that have the highest ratio of number of papers published to extremely poor people: here, the ratio tops out at around 30,000 for Bosnia and Herzegovina, and remains high for a range of countries largely in Eastern Europe, Central Asia, and Latin America that are dramatically overrepresented relative to the number of poor people.

To provide a more comprehensive summary, Figures 1c and 1d show scatterplots of the relationship between the number of papers published and the number of poor people living below the $1.90 poverty line (in Figure 1c) or the $5.50 poverty line (in Figure 1d)/ (The x-scale is logarithmic.) The sample is restricted to countries represented in at least eight papers (the 75th percentile) for tractability. In Figure 1c, there is essentially no evidence of a positive relationship, and three notable outliers: both Mexico and China are characterized by a relatively large number of papers published despite relatively few poor people, and India has an extremely large number of poor people and a relatively high number of papers. In Figure 1d, the relationship is somewhat more positive, suggesting that publication intensity is more correlated with the number of moderately poor people than the number of extremely poor people in a given country.

Third, I aggregate information about the representation of funders, focusing now more

narrowly on the sample of 119 randomized controlled trials. As previously noted, 106 of these RCT-reporting papers report funding. The median and mean number of sources reported is two, with a maximum of ten. Figure 2 then summarizes the representation of funders within the sample; the funders are ranked from least to most represented in the full sample, and the corresponding bars show representation in the top five sample. In general, as noted above the most represented funders are the IGC, the World Bank, and other U.S. funding sources (primarily the NIH and NSF). Among the more minimally represented funders are the Weiss Fund (a smaller fund that targets research by early-career development economists at select institutions), the CGIAR, Australian government sources, the Bill and Melinda Gates Foundation, the Inter-American Development Bank, and the Economic and Social Research Council. These funders are all reported by less than 5% of the published RCT papers.

In the top five sample, a number of funders represented drop out entirely, and the dominance of the IGC, J-PAL and the ESRC is clear, followed by USAID and FCDO. Given that the IGC is majority funded by the FCDO, if these categories are summed then the FCDO provided funding to nearly 30% of published RCTs, and nearly 70% of RCTs published in top five journals during this period.

4   Discussion and conclusion

Our findings here generally suggest that while RCTs are far from a majority of published development economics papers, they do constitute nearly 40% of published papers in the top five journals. Banerjee et al. (2016) similarly noted that the majority of all papers are not RCTs, and documented some facts around the share of development economics papers that were RCTs in the top five journals in 2015 (31%), at the New England University Development Conference (between 16% and 24% in 2012–2015) and in the BREAD workshop (around 40–50%). The data here suggests that since that date, representation of RCTs in top five journals has continued to increase, though again, even in these forums RCTs do not constitute a majority of published development economics research. The rates observed in the general sample of published economics papers are substantially lower, and more comparable to the rates previously reported for papers presented at NEUDC.

The findings on the geographic concentration of development economics research are novel.[1]   They suggest that even within the sample of developing countries, a number

---

[1](Peters et al., 2016) and (Peters et al., 2018) also identified a sample of RCT papers published in top general interest journals between 2019 and 2014, primarily to analyze their treatment of questions related to external validity, and provide a geographic breakdown. However, this analysis is inclusive of RCTs conducted in the U.S. in (Peters et al., 2016), and does not include non-RCT development

of countries are overrepresented relative to their population of people living in poverty. Among the countries that are most commonly the focus of development economics research, Mexico, China and Colombia are relatively highly represented; sub-Saharan Africa is notably underrepresented. Interestingly, in the sample of RCT papers, the pattern is rather different: while India (one of the two most commonly represented in the full sample) is still dominant, three sub-Saharan African countries (Uganda, Kenya, and Nigeria) are now among the most five represented countries. Some researchers have already argued that one of the benefits of the rise of RCTs is that it catalyzes more collaborations with local research institutions (based in developing countries) and more cross-disciplinary collaborations, though both are still relatively rare (Naritomi et al., 2020). Another observation suggested by the evidence here is that RCTs encourage data collection in a broader spectrum of relatively poor countries (particularly in sub-Saharan Africa), rather than a focus on countries that have substantial administrative data or existing surveys.

Similarly, the patterns observed of concentration in funding of RCTs are novel, particularly for the most highly published papers. The International Growth Centre and the Foreign Commonwealth and Development Office (formerly DFID) are clearly dominant in the top five sample, while the World Bank and U.S. research funders other than USAID show substantial representation in the broader sample. Despite its clearly dominant role in the disbursement of development aid, USAID is not particularly well-represented in this sample. Even in the pooled sample, more papers acknowledge J-PAL (a much smaller organization, albeit one purely focused on research), and similarly more papers acknowledge other U.S. government sources such as the NIH and NSF (despite the fact that these have a very broad disciplinary remit and do not primarily fund social science research). The findings are suggestive of a high level of variation in the ability of funders to identify research that generates highly-placed publications, though of course funders' portfolios vary in size, and their funding objectives may similarly vary and encompass a range of objectives in addition to publication.

This note is not intended to suggest that the current allocation of research across RCTs and non-RCTs and across countries is suboptimal, or to draw any conclusions about the relative salience of different funders. It is informative to scrutinize these patterns in order to enable more reflection on the part of the discipline about questions that are examined, and what the returns are to different modalities of research funding in different spheres.
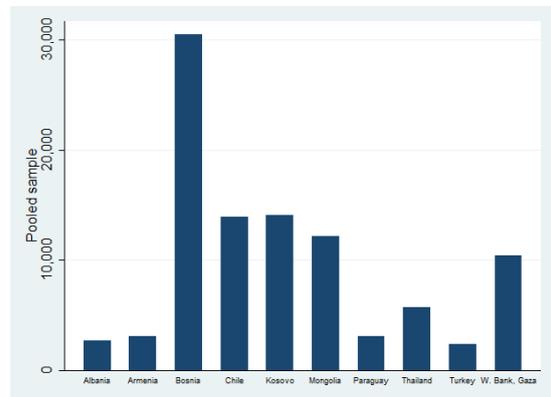
---

economics papers.

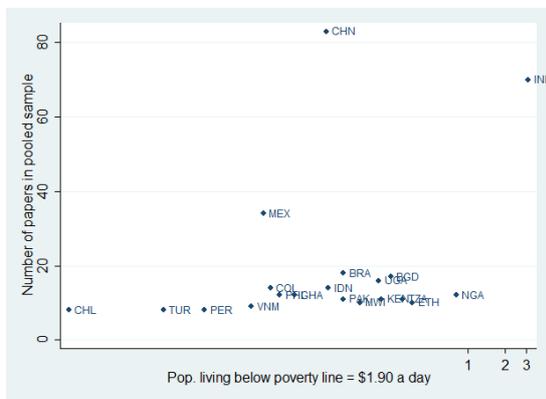Figure 1: Representation of countries in paper sample and population living in poverty

(a) Normalized number of papers, most commonly observed countries
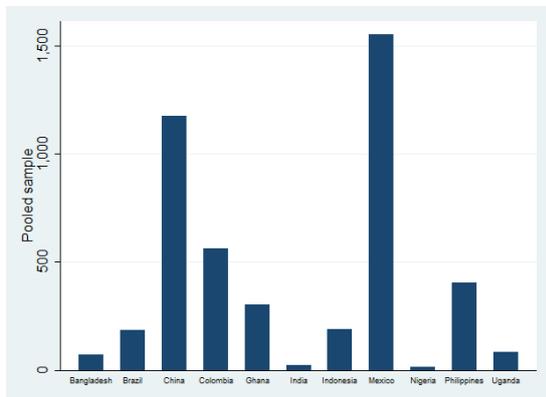
(b) Normalized number of papers, most highly represented countries

(c) Relationship between number of papers and pop living in poverty

(d) Relationship between number of papers and pop living in less severe poverty
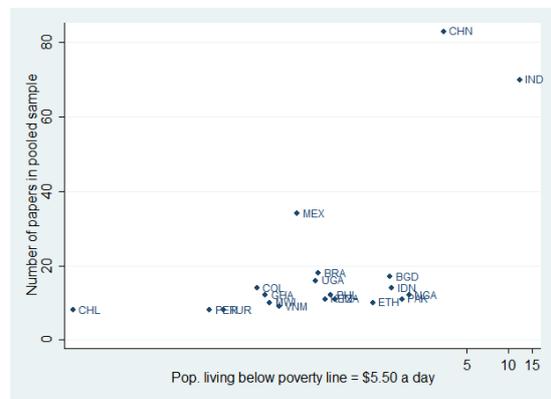
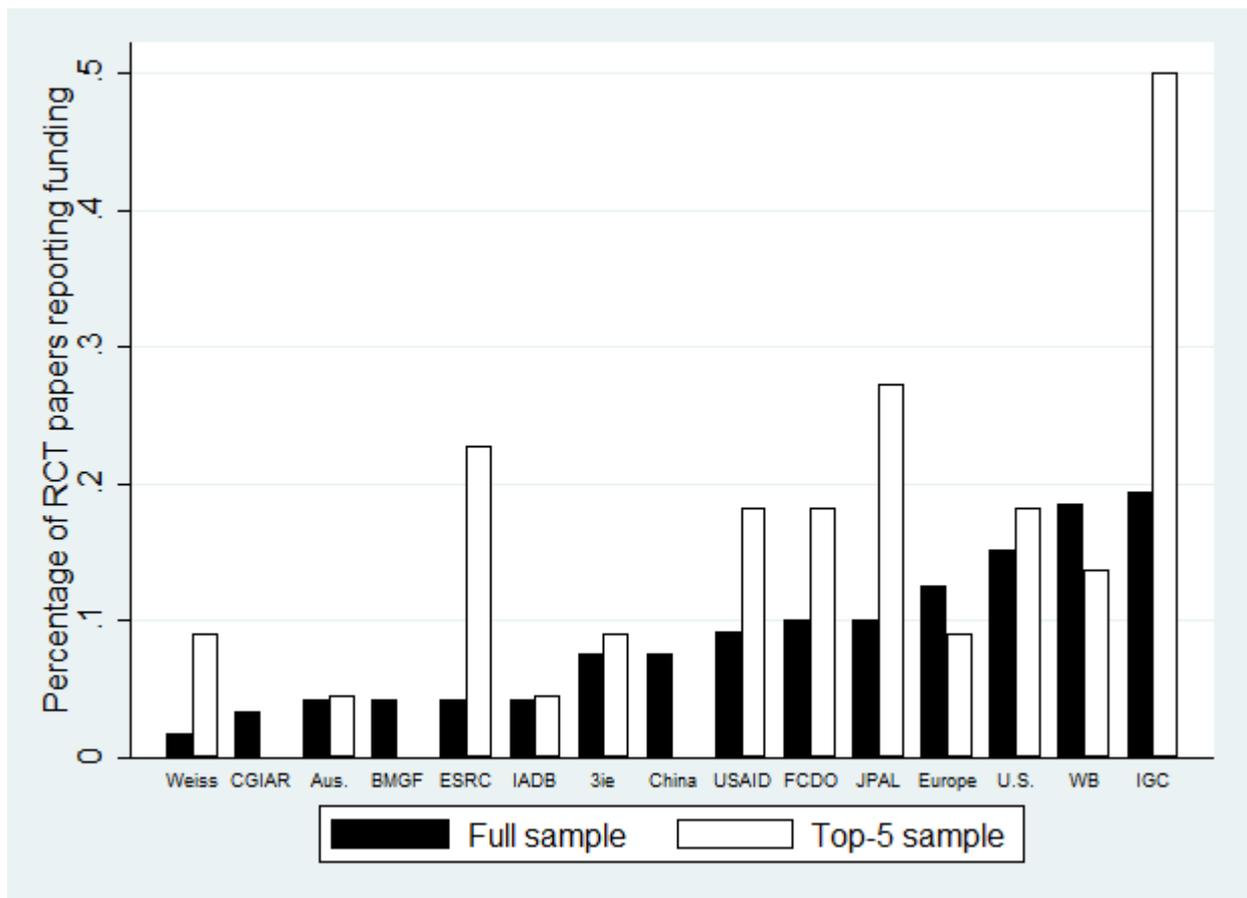Figure 2: Funding sources for published randomized controlled trials

Table 1: Overview of funders

Funder	Additional notes
Australian government sources	Encompasses the Australian Departments of Foreign Affairs and Trade, the Australian Agency for International Development (now part of DFAT), and the Australian Research council.
Bill and Melinda Gates Foundation (BMGF)	
China government sources	Encompasses the National Natural Science Foundation of China, the Research Grants Council of the Hong Kong Special Autonomous Region, the Fundamental Research Funds for the Central Universities, the 111 Project, the Ministry of Industry and Information Technology, and the National Key Research and Development Program of China
CGIAR	
Economic and Social Research Council (ESRC)	
Europe government sources	Encompasses the Swedish International Development Cooperation Agency; the Government of the Netherlands; the Research Council of Norway; the European Commission; Irish Aid; the Spanish Impact Evaluation Fund; the French National Research Agency (ANR); the Ministerio de Economía y Competitividad of the Spanish Government; and the European Research Council
Foreign and Commonwealth Development Office (FCDO)	Formerly the Department for International Development (DFID)
Inter-American Development Bank (IADB)	
International Growth Centre (IGC)	IGC is substantially funded by FCDO, formerly DFID.
International Initiative for Impact Evaluation (3ie)	3ie is funded by a range of outside funders.
Abdul Latif Jameel Poverty Action Lab (J-PAL)	J-PAL is funded by a range of outside funders.
United States Agency for International Development (USAID)	In some cases researchers acknowledged university research initiatives that were themselves entirely funded by USAID; in this case, the funding is attributed to USAID.

Table 1: Overview of funders, cont.

Funder	Additional notes
Other US government sources	Encompasses the National Science Foundation, the National Institutes of Health, the Millennium Challenge Corporation, and the Department of Labor
Weiss Fund for Research in Development Economics	Supports research by students and faculty at BU, Columbia, Harvard, MIT, Northwestern, Princeton, Stanford, UC - Berkeley, UCSD, U Chicago, and Yale
World Bank (WB)	The World Bank is funded by a range of outside funders.

Table 2: Countries represented by paper sample

(1) Top 5	(2) Top general interest	(3) Field journal	(4) Pooled sample	(5) Pooled sample - RCTs
India (20%)	China (17%)	China (13%)	China (14%)	India (11%)
Indonesia (8%)	India (16%)	India (9%)	India (12%)	Bangladesh (8%)
Brazil (8%)	Mexico (8%)	Mexico (5%)	Mexico (6%)	Uganda(7%)
Pakistan (7%)	Brazil (6%)	Bangladesh (3%)	Brazil (3%)	Kenya (6%)
China (7%)	Bangladesh (4%)	Uganda (3%)	Bangladesh (3%)	Nigeria (6%)
Theory (3%)	Theory (1%)	Theory (4%)	Theory (3%)	
Multicountry (8%)	Multicountry (16%)	Multicountry (15%)	Multicountry (15%)	

References

Banerjee, Abhijit Vinayak, Esther Duflo, and Michael Kremer, "The influence of randomized controlled trials on development economics research and on development policy," *The State of Economics, The State of the World*, 2016, pp. 482–488.

Das, Jishnu, Quy-Toan Do, Karen Shaines, and Sowmya Srikant, "US and them: The geography of academic research," *Journal of Development Economics*, 2013, *105*, 112–130.

Deaton, Angus, "Instruments, randomization, and learning about development," *Journal of economic literature*, 2010, *48* (2), 424–55.

Fontana, Magda, Fabio Montobbio, and Paolo Racca, "Topics and geographical diffusion of knowledge in top economic journals," *Economic Inquiry*, 2019, *57* (4), 1771–1797.

Hirvonen, Kalle, "This is US: Geography of evidence in top health economics journals," *Health Economics*, oct 2020, *29* (10), 1316–1323.

Naritomi, Joana, Sandra Sequeira, Jonathan Weigel, and Diana Weinhold, "RCTs as an opportunity to promote interdisciplinary, inclusive, and diverse quantitative development research," *World Development*, 2020, *127*, 104832.

Peters, Jörg, Jörg Langbein, and Gareth Roberts, "Policy evaluation, randomized controlled trials, and external validity—A systematic review," *Economics Letters*, 2016, *147*, 51–54.

Peters, Jörg, Jörg Langbein, and Gareth Roberts, "Generalization in the Tropics – Development Policy, Randomized Controlled Trials, and External Validity," *The World Bank Research Observer*, 03 2018, *33* (1), 34–64.

Rao, Vijayendra, "Evidence-based development needs a diversity of tools, with a bottom-up process of "embedded" dialogue," *World Development*, 2020, *127*, 104823.

Ravallion, Martin, "Highly prized experiments," *World Development*, 2020, *127*, 104824.