World Development 127 (2020) 104817

Contents lists available at ScienceDirect

World Development

journal homepage: www.elsevier.com/locate/worlddev

The challenges of scaling effective interventions: A path forward for research and policy

C. Austin Davis [a], Ahmed Mushfiq Mobarak [b,*]

[a] American University, United States
[b] Yale University, United States

ARTICLE INFO

Article history:
Accepted 1 December 2019

Keywords:
Complexities of scaling anti-poverty interventions
General equilibrium consequences

ABSTRACT

With minimal statistical or theoretical assumptions, randomized controlled trials (RCTs) provide a necessary input for poverty analysis: credibly estimated causal relationships. But complexities arise when moving from RCT research results to anti-poverty policy, with unintended consequences. RCT evidence by itself offers an incomplete prediction of the effects of policy, due to heterogenous effects, spillovers and general equilibrium changes, macroeconomic and welfare effects, political economy reactions, and implementation challenges, when programs are scaled. We suggest strategies for tightening the link between development research and anti-poverty policy, for example, by changing the practice of RCTs to be more ambitious about what is randomized, and to combine the analysis of experimental data with other rigorous methods that go beyond estimating treatment effects. We describe our efforts to encourage and coordinate this type of work via a new research initiative.

Questions like "Why do some people or some countries remain poor?" and "What policies or interventions will improve their circumstances?" have been central concerns of development economics. Ideally, we like to offer credible answers to these questions by estimating causal relationships. Randomized controlled trials (RCTs) are a useful addition to our analytical toolkit because of their promise to reveal causal relationships with few statistical or theoretical assumptions. RCT results are also easy to explain to various audiences, including policymakers.

These virtues explain the influence of RCTs on researchers and development actors. This year's Nobel laureates were deservedly honored building research infrastructure that improved the work of hundreds of researchers and encouraged evidence-based policy. But, as with other tools, RCTs have limitations as a guide to policy. While RCTs overcome selection problems with minimal assumptions (no small task!), that's only part of the way to providing general, forward-looking policy guidance.

Heterogenous treatment effects are one complication for taking policy lessons from RCTs. In their simplest form, RCTs estimate the average treatment effect of an intervention on a particular sample in a particular environment. These estimates offer an imperfect prediction for other implementations of the intervention if the effects depend on characteristics of individuals in the sample or features of the environment. Moreover, the distribution of treatment effects may be more important than the average; for instance, a policymaker might prioritize interventions that make no one worse off or that benefit some constituency.

Each intervention embodies many implementation choices which may impact efficacy. Many business training programs have been motivated by the observation that developing-country microenterprises generally do not engage in basic practices like accounting or advertising. Unfortunately, RCTs find that business training does not systematically enhance profits (see McKenzie and Woodruff (2014) and Blattman and Ralston (2015) for reviews). Implementers of business training curriculum make myriad choices, including the content, format, schedule, instructors, and students. Would some other set of choices yield better results?

Then there are complexities that arise when an intervention is implemented at the scale of a government policy. Spillovers are one such complexity. Disease eradication is a canonical example, with large-scale administration of vaccines, bed nets, or deworming medication offering protection even to untreated individuals. But spillovers need not be positive; at scale, general equilibrium effects might undermine the benefits of an asset transfer program by lowering the price of the asset or its output. Spillover effects may have a macroeconomic flavor, involving dynamic processes or other types of agents. For instance, large-scale microfinance

* Corresponding author at: Yale University, 165 Whitney Avenue, New Haven, CT 06520-8200, United States.

E-mail addresses: cadavis@american.edu (C.A. Davis), ahmed.mobarak@yale.edu (A.M. Mobarak).

https://doi.org/10.1016/j.worlddev.2019.104817
0305-750X/© 2019 Published by Elsevier Ltd.

can affect aggregate productivity and investment by firms (Buera, Kaboski, & Shin, 2012).

Finally, moving from research activities to large-scale policy can provoke political reactions. Ineffective leaders may remain in office by claiming credit for successful programs. Externally funded interventions may crowd out government support for substitutes or crowd in support for complements. Implementation challenges reemerge here, as politicians and bureaucrats often face different incentives and constraints than the organizations that partner with researchers.

Broadly, we see four fruitful strategies for addressing these challenges, building on the rigorous methods that are now standard in the field: i) being ambitious about what gets randomized, ii) creative, theory-driven experimental design, iii) combining randomization with other tools like structural or statistical models to explore the complexities that may emerge at larger scale, and iv) centralized, coordinated research initiatives.

Ambitious experiments can measure the effects of scale and bring the benefits of RCTs to important questions that are difficult to study (Muralidharan & Niehaus, 2017). Through a partnership with the government of Mexico, Cunha, de Giorgi, and Jayachandran (2019) study how the medium of social benefits affects market-level prices; the design of this experiment allows them to measure the general equilibrium effects of delivering social benefits in the form of cash versus in-kind transfers.

Innovative experimental designs, by addressing why or how an intervention works, can enhance external validity and guide the design of future interventions. Karlan and Zinman (2009) use a clever two-stage randomization to distinguish the effects of two types of asymmetric information, hidden information and hidden action, in credit markets. The difference matters for specific policy design: Should we pursue screening technologies and information sharing, or legal and contracting reforms?

Combining randomization with structural or statistical models uses experimentally identified parameters to extrapolate beyond the experimental setting or measured outcomes. Gechter, Samii, Dehejia, and Pop-Eleches (2019) confront the problem of using prior experimental results to predict the effects of future policy implementations on a new set of participants. Two recent papers develop new techniques to aggregate experimental results on microcredit, finding that about three quarters of participants saw no benefit at all (Meager, 2019a, 2019b). The tools of welfare economics, when paired with RCT data, can credibly answer thorny questions like "as a migrant, how unpleasant is it to be separated from your family?" (Lagakos, Mobarak, & Waugh, 2019)

We can tighten the link between research and policy not only through the innovations of individual studies, but also by coordinating multiple studies. A series of microfinance evaluations published as a special edition of American Economic Journal: Applied Economics, illustrate the potential of this approach. But further gains may be realized through closer coordination and centralized management. This approach crowds in research thanks to comparative advantage, economies of scale, and data harmonization. With project administration handled centrally, researchers can specialize in identifying priority research questions, generating research designs, and conducting and communicating the analysis. The large fixed costs of RCTs can be amortized over multiple projects, attacking the same problem from many angles. Finally, the policy relevance of research is enhanced by providing a comprehensive, coherent base of evidence about a particular problem or intervention.

The Yale Research Initiative on Innovation and Scale (Y-RISE, http://yrise.yale.edu) was founded to support this kind of next-generation research that bridges the gap between RCT evidence and at-scale policy effects. One way Y-RISE encourages scale-up research is by making it cheaper, providing start-up grants and research assistance. Y-RISE has assembled several networks of researchers with the goals of shifting researcher rewards, identifying areas where the network can expand the frontier of knowledge, and fostering collaboration.

Inspired by a body of work around seasonal poverty, Y-RISE is developing the type of coordinated, centrally managed research initiatives described above. For example, we are exploring different solutions to seasonal deprivation in different geographies, including seasonally timed loans during lean periods in southern Africa (Fink, Jack, & Masiye, 2018) and migration subsidies in South Asia (Bryan, Chowdhury, & Mobarak, 2014). Within the migration arm, we aim to inform policymakers about a range of direct and indirect, unanticipated issues that may arise with large-scale migration, beyond the direct effects of subsidies on recipient behavior. These include: i) spillover effects on non-recipients (Akram, Chowdhury, & Mobarak, 2018), ii) unintended consequences on health, relationships and domestic violence (Mobarak & Reimão, 2020), and iii) more ambitious questions about the aggregate productivity consequences of mobility. Responsible evidence-based policymaking requires engaging with the full range of evidence produced using a combination of research methods.

More generally, Y-RISE aspires to build on the rigorous, scientific approach to poverty that earned this year's Nobel. Bringing together researchers, implementers, policymakers, and funders, we will combine RCTs with other research methods to investigate the complexities associated with scaling promising anti-poverty interventions. It is our hope that these efforts will generate the kind of nuanced and holistic evidence that can have a positive effect on policymaking and human wellbeing.

References

Akram, A. A., Chowdhury, S., & Mobarak, A. M. (2018). Effects of Emigration on Rural Labor Markets. Working Paper.

Blattman, C., & Ralston, L. (2015). Generating Employment in Poor and Fragile States: Evidence from Labor Market and Entrepreneurship Programs. Working Paper. doi: 10.2139/ssrn.26.

Bryan, G., Chowdhury, S., & Mobarak, A. M. (2014). Under-investment in a profitable technology: The case of seasonal migration in Bangladesh. *Econometrica, 82*(5), 1671–1748.

Buera, F.J., Kaboski, J. P., & Shin, Y. (2012). The Macroeconomics of Microfinance. NBER Working Paper 17905. doi: 10.3386/w17905.

Cunha, J. M., de Giorgi, G., & Jayachandran, S. (2019). The price effects of cash versus in-kind transfers. *Review of Economic Studies, 86*, 240–281. https://doi.org/10.1093/restud/rdy018.

Fink, G., Jack, B. K., & Masiye, F. (2018). Seasonal Liquidity, Rural Labor Markets and Agricultural Production. NBER Working Paper 24564. doi: 10.3386/w24564.

Gechter, M., Samii, C., Dehejia, R., & Pop-Eleches, C. (2019). Evaluating Ex Ante Counterfactual Predictions Using Ex Post Causal Inference. Working Paper.

Karlan, D., & Zinman, J. (2009). Observing unobservables: Identifying information asymmetries with a consumer credit field experiment. *Econometrica, 77*(6), 1993–2008.

Lagakos, D., Mobarak, A. M., & Waugh, M. E. (2019). The welfare effects of encouraging rural-urban migration. Working Paper.

McKenzie, D., & Woodruff, C. (2014). What are we learning from business training evaluations around the developing world?. *World Bank Research Observer, 29*(1), 48–82.

Meager, R. (2019a). Aggregating Distributional Treatment Effects: A Bayesian Hierarchical Analysis of the Microcredit Literature. Working Paper.

Meager, R. (2019b). Understanding the average impact of microcredit expansions: A Bayesian hierarchical analysis of seven randmoized experiments. *American Economic Journal: Applied Economics, 11*(1), 57–91. https://doi.org/10.1257/app.20170299.

Mobarak, A. M., & Reimão, M. (2020). Seasonal migration and seasonal poverty in Asia. *Asian Development Review, 37*(1), 1–42. https://doi.org/10.1162/adev_a_00139.

Muralidharan, K., & Niehaus, P. (2017). Experimentation at scale. *Journal of Economic Perspectives, 31*(4), 103–124. https://doi.org/10.1257/jep.31.4.103.