

Scaling-up Proven Education Interventions

Evidence from a RCT in Kenya



In brief

- Randomized controlled trials (RCTs) have become a major methodological tool for policy evaluation in education. However, there are a number of concerns regarding scaling up the methods of a successful RCT.
- Firstly, external validity is a major concern - i.e. variation in treatment effects or the interaction between treatments and the characteristics of experimental subjects. Secondly, whether successful NGO pilot projects can be replicated by governments and scaled-up within the institutional constraints of public sector bureaucracies is a concern.
- This paper presents an RCT designed to address both of these concerns. The authors examine the prospects of scaling-up a contract teacher intervention (shown to significantly raise test scores for primary students in both Western Kenya and India) by expanding geographically across all 8 Kenyan provinces and by comparing implementation by an international NGO and the Ministry of Education.
- Our results suggest geographic heterogeneity in treatment effects does not pose a threat to external validity of earlier studies. However, institutional constraints may pose a serious problem in scaling up pilot interventions.
- **Key policy implications:**
 - Define interventions broadly: one must note that organizational issues matter.
 - Researchers should collaborate with the institutions they wish to inform as extrapolating results from an NGO program to government policy may be invalid.
 - Pay attention to the institutional context. In the Kenyan example, when scaled-up nationally, contract teachers paid at half the rate of civil service teachers led to a union backlash. Thus, the salient features of teacher contracts were compromised.

Policy Motivation

“Generalizing the results of any randomized controlled trial to assess the effect of scaling up a given intervention raises several well-known questions”

In the past decade, the methodology for policy evaluation in education has shifted to randomized controlled trials. Critics have questioned the generalizability of this methodology (e.g., Deaton 2010). Acemoglu (2010) notes: “Evidence about the effects of a particular policy change on economic outcomes is not in itself sufficient to gauge what the implications will be when such a policy is encouraged or implemented.”

Generalizing the results of any randomized controlled trial to assess the effect of scaling up a given intervention raises several well-known questions (Shadish, Campbell, and Cook 2002; Duflo 2004). The first is external validity, which reflects concerns about heterogeneity of treatment effects, or the interaction between treatments and the characteristics of the experimental subjects. A second question is whether successful NGO pilot projects can be replicated by governments -- and in particular, within the institutional constraints of public sector bureaucracies in the developing world. In Shadish, Campbell, and Cook’s (2002) terminology, this is a question of ‘construct i.e., how conditions of experimentation differ from the conditions of policy implementation.

In this paper, we present a randomized controlled trial designed to address both of these concerns: external validity, i.e., variation in treatment effects across subjects, and construct validity, focusing on how conditions of experimentation differ from the conditions of policy implementation. We examine the prospects for scaling-up a contract teacher intervention, shown to significantly raise test scores for primary students in previous randomized trials in both Western Kenya (Duflo, Dupas, and Kremer 2009) and various locations in India (Banerjee, Cole, Duflo and Linden 2007; Muralidharan and Sudararaman 2010), by (a) expanding geographically across all eight Kenyan provinces, and (b) comparing implementation by an international NGO and the Ministry of Education. The timing, salary levels, recruitment procedures and all other experimental protocols were held constant across the government and NGO arms of the evaluation.

Policy Impact

“The timing, salary levels, recruitment procedures and all other experimental protocols were held constant across the government and NGO arms of the evaluation.”

Our results provide important guidance to policy makers trying to learn from the evidence base amassed by policy evaluations based on RCTs and to bring to the fore the political economy issues involved in scaling up interventions shown to be successful in small-scale NGO trials.

Audience

This brief addresses two broad groups:

1. Researchers conducting RCTs and other applied research
2. Policy makers in developing countries and in the donor community using or commissioning RCTs for policy evaluation.

Policy Implications

Define interventions broadly

Our results demonstrate that a randomized evaluation of a contract teacher program in Kenya is not just an evaluation of a contract teacher program in Kenya. Rather, it is an evaluation of a contract teacher program implemented by an international NGO with expatriate technical assistance and careful monitoring – or, alternatively, it is a contract teacher program implemented by a government Ministry, subject to the organizational constraints of the Kenyan public sector bureaucracy. These organizational issues matter, perhaps as much or more than the minutiae of the intervention design.

“Whenever possible, researchers should strive to do policy research in collaboration with the institution whose policies they wish to inform”

Work with governments

Whenever possible, researchers should strive to do policy research in collaboration with the institution whose policies they wish to inform. Our results suggest that extrapolating results from an NGO program to government policy may be invalid. An obvious solution is to collaborate with government implementers at the pilot stage.

Pay close attention to the institutional context in which proven policy interventions are to be scaled up

In the Kenyan context, introducing contract teachers led to the ‘seesaw’ policy effect documented by Acemoglu (2010). In small-scale trials, contract teachers were effective at raising pupil performance. When scaled-up nationally, contract teachers paid at half the rate of civil service teachers led to a backlash by the teacher unions, the salient features of teacher contracts were compromised, and the effectiveness of the program was undermined.

Brief Summary of Research

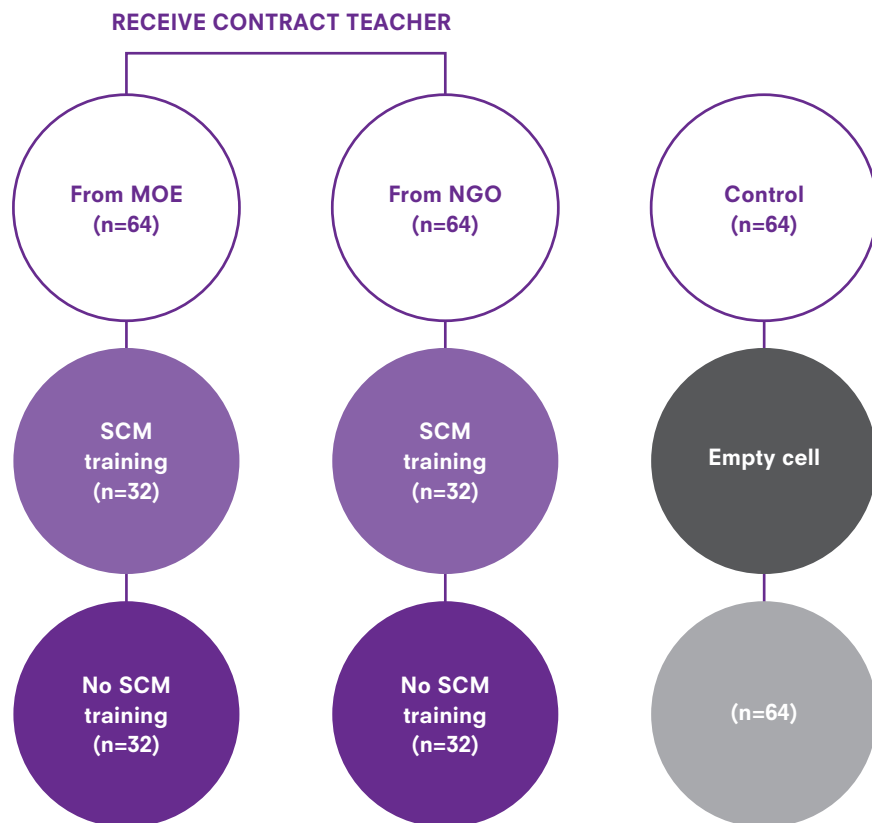
The experiment was implemented from June 2010 to October 2011 in 14 districts spanning all 8 Kenyan provinces. 24 schools were sampled from each province, yielding 192 schools in total. One contract teacher per school was randomly assigned to 128 out of 192 sampled schools. In the treatment schools, contract teachers were assigned to teacher either grade 2 or 3.

“Organizational issues matter, perhaps as much or more than the minutiae of the intervention design.”

Pupils in both treatment and control schools were tested in mathematics and English at baseline and follow-up. The impact of the contract teacher program is measured by the change in average test performance between rounds.

In addition to the contract teachers, 64 of the 128 treatment schools also received training for members of the School Management Committee (SMC) training focused on the roles and responsibilities of the parents in governing the school and holding teachers accountable. The allocation of schools between these two treatments is illustrated in Figure 1

Figure 1: Experimental Design



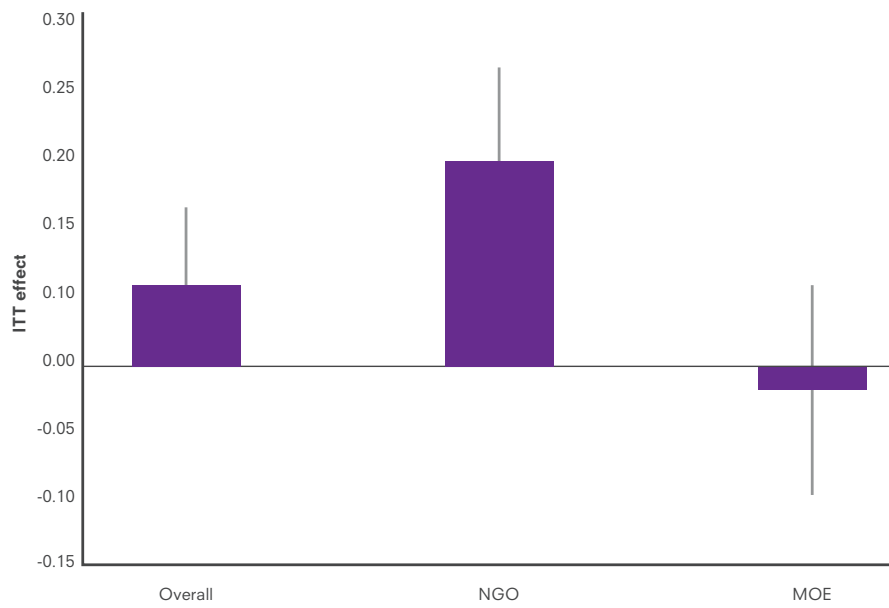
“Schools in high- and low-density areas did not experience different benefits from the program, nor did effects differ according to the initial class sizes in the school.”

Overall, comparing treatment and control schools we find a significant, positive effect of 0.12 standard deviations on combined math and English scores across all treatment arms, with few significant differences in treatment effects across baseline school characteristics. However, overall effects were entirely due to a 0.19 standard deviation increase in scores in the schools randomly assigned to NGO implementation. Effects were significantly smaller and indistinguishable from zero in schools where the program was administered by the Ministry of Education.

Figure 2 shows the intention-to-treat effect overall, and divided between the NGO treatment arm and the Ministry of Education treatment arm.

Additional analysis, not shown, explored differences in the effectiveness of the program across schools of various types and in distinct locations. Overall, few significant differences were found. Schools in high- and low-density areas did not experience different benefits from the program, nor did effects differ according to the initial class sizes in the school.

Figure 2. Effect of Contract Teachers on Pupil Performance



To summarize, our results suggest that geographic heterogeneity in treatment effects does not pose a significant threat to the external validity of earlier studies; we find few significant differences in treatment effects across baseline school characteristics. However, the results here imply that institutional constraints may pose a serious obstacle to scaling up pilot interventions in developing countries.

Dissemination

Permanent Secretary, Ministry of Education, Jogo'o House, Nairobi, Kenya
CEO, Kenya National Examination Council, Nairobi, Kenya

About the authors

Tessa Bold is Chair of Development and International Economics at Goethe University Frankfurt, Germany. Her research interests include: Applied microeconomics and microeconometrics, microeconomic foundations of macroeconomics, group and network formation, contract theory, public and organizational economics, political economy.

Mwangi Kimenyi is Senior Fellow and Director of the Africa Growth Initiative in the Global Economy and Development program of The Brookings Institution. Kimenyi is also a Research Associate with the Center for the Study of African Economies. Kimenyi's research focuses on institutions and economic development, Africa's political economy, and policies for economic growth and poverty reduction.

Germano Mwabu is Professor of Economics at the University of Nairobi. He was co-director of the AERC project on Poverty Income Distribution and Labour Markets in Sub-Saharan Africa. He has also held the position of Senior Research Fellow, Director, Public Goods Project of the World Institute for Development Economics Research at the United Nations University, Helsinki and Associate Professor of Economics at Kenyatta University.

Justin Sandefur is a research fellow at the Center for Global Development. His research focuses on the interface of law and development in sub-Saharan Africa. He has served as an adviser to the Tanzanian government to set up the country's National Panel Survey to monitor poverty dynamics and agricultural production.

The International Growth Centre
(IGC) aims to promote sustainable
growth in developing countries
by providing demand-led policy
advice based on frontier research.

Find out more about
our work on our website
www.theigc.org

For media or communications
enquiries, please contact
mail@theigc.org

Follow us on Twitter
[@the_igc](https://twitter.com/the_igc)

International Growth Centre,
London School of Economic
and Political Science,
Houghton Street,
London WC2A 2AE



International
Growth Centre

Designed by soapbox.co.uk