Randomized Control Trial as Social Observatory: A Case Study[1]

Sarah Baird[a], Joan Hamory Hicks[b], Owen Ozier[c]

[a]Department of Global Health, Milken Institute School of Public Health, George Washington University, sbaird@gwu.edu
[b]Department of Economics, University of Oklahoma, jhamoryhicks@ou.edu
[c]Development Research Group, The World Bank, oozier@worldbank.org

Keywords: deworming; RCT; longitudinal research; spillovers; case study; Busia, Kenya

Critics of randomized control trials (RCTs) in development economics argue that this methodology lends itself to 'smaller' questions with limited relevance to policy or economics. Using the seminal work of Miguel and Kremer (2004) on a school-based deworming intervention in Kenya as a case study, we argue that RCTs can spearhead policy change, serve as a laboratory to test economic theories and develop cutting-edge empirical methods, or do both. This does not happen in a vacuum, but through thoughtful design embedded in a broader research and policy agenda. Here, we describe a family of studies built on Miguel and Kremer (2004), shedding light on factors that supported the generation of evidence and insights far beyond the near-term RCT result. As in any piece of social sciences research, this descriptive evidence may not be externally valid in all settings. We nevertheless hope the lessons it offers will inspire others to examine these possibilities in their own research.

[1] We were trained by Edward Miguel and Michael Kremer, and Busia, Kenya is where we each first experienced field research. These experiences profoundly influenced our approach to development economics and subsequent research, and underpin the reflections in this article. The findings, interpretations and conclusions expressed in this paper are entirely our own and do not necessarily represent the views of the World Bank, its executive directors, or the governments of the countries they represent. The title is inspired by the work of Vijayendra Rao.

Introduction

Randomized control trials (RCTs) are critiqued for not answering the important development questions (e.g. Pritchett 2018). One group of prominent economists went so far as to say that they "[create] scenarios so abstract as to tell us little about the real world," calling RCTs "tortuous impact studies" that are "expensive to carry out, and fraught with ethical challenges" (*The Guardian* 2018). Still more see value in RCTs, but challenge their epistemological pre-eminence (Barrett and Carter 2010; Ravallion 2018). We agree that RCTs are but one tool in a large toolkit, and we agree that RCTs may never be feasible for some questions. But we argue that many claimed pitfalls of RCTs are not fundamentally inherent to RCTs. In fact, through thoughtful design and investment, RCTs can serve as laboratories to test economic theories, develop new methods, and inform policy. We use the seminal work of Miguel and Kremer (2004) on school-based deworming as a case study. We illustrate factors that could have contributed to its ramifications far beyond the first RCT result. Specifically, we focus on (i) innovative design connected to policy, (ii) secondary questions stemming from the original design, (iii) longitudinal data collection and subsequent RCTs, and (iv) policy engagement.

Miguel and Kremer (2004), hereafter MK2004, evaluated the impact of a school-based deworming intervention on the health and education of primary school-aged children in Busia, Kenya. The study involved a cluster RCT (cRCT) that randomly divided seventy-five schools into three groups of twenty-five schools each. Due to constraints of the administering NGO, the health intervention—known as the Primary School Deworming Project (PSDP)—was phased in over four years: the first group of schools received deworming treatment starting in 1998, the second in 1999, and the third in 2001. MK2004 found that the PSDP reduced worm infections and increased school participation at low cost. They also found substantial externality benefits (spillovers) in these domains for untreated children in treatment schools and children attending nearby schools.

MK2004 has been cited over 2000 times. Why so much interest in school-based deworming? From an academic perspective, attention largely comes from the careful analysis of externalities contained in the paper. While more common today, MK2004 was one of the first papers in economics to illustrate how a cRCT design permits estimation of the overall effect of a program in the presence of treatment externalities (both within and across schools).[2] Though Kremer and Miguel did not plan

[2] The original published article suggested analyzing spillovers as far as six kilometers from the school; one of the conclusions from the "worm wars" was that given the research design, between three and six kilometers was too far for precise estimation to be possible, but that nonetheless, benefits were likely to be present at shorter distances (Miguel and Kremer 2014; Ozier 2019).

for analysis of cross-school externalities at the outset of the study[3], this methodological innovation in their paper has inspired a great deal of further research.

One design answers many questions

We first look at how the spillover framework combined with the literature on 'critical' periods of human development to resolve a mystery that stemmed from MK2004. Despite showing beneficial effects of deworming on health and school attendance, the study found no measurable impact on test scores. Why might this be? A possible partial answer lies in the literature on critical and sensitive periods: interventions (or insults) in early childhood have larger and more lasting effects on physical and cognitive development than do comparable events later in childhood (Knudsen, et al., 2006). This leads to the hypothesis that deworming might have impacted test scores, had the intervention targeted younger children.

Testing the hypothesis required a second insight: the role of spillovers. As noted above, spillovers were one of the defining contributions of MK2004. That is, children did not need to receive deworming medication directly in order to benefit from mass deworming. Untreated children who attended treatment schools—and even children in nearby schools—benefited in terms of health and school attendance. This implies that while babies in these communities were too young to have received medicine at school, these babies may still have benefited from the epidemiological externalities of the original intervention. Testing this hypothesis required gathering data from these children a decade later, which is precisely what one of us did. A team measured anthropometric outcomes for many thousands of children around Busia, and gathered more extensive cognitive data on a subsample. The effects were stark: one decade after program launch, children who were babies during the first waves of the PSDP demonstrated 0.2-standard-deviation higher performance on an index of cognitive tests than did children from the same cohorts in later-dewormed communities (Ozier 2018a).

Even beyond this specific finding, MK2004 has spawned a vast and growing literature on the measurement of spillovers (see, for example, Baird et al. 2018).

[3] See Hicks, Kremer, and Miguel (2015).

Longitudinal data and more RCTs

While the MK2004 results findings were influential on their own, what came next further cemented the impact and broader value of this RCT to the economics and public health communities, and beyond. Utilizing elements of the design built into the RCT (including cost-sharing by the recipient families, and health education), a follow-up paper by the authors (Kremer and Miguel 2007) was at the forefront of the literature on experimentally estimating demand for health within development economics.

Furthermore, committed to understanding whether these childhood health improvements had longer term impacts and recognizing the dearth of longitudinal data collection efforts in less-developed country settings, Miguel initiated a follow-up on a subsample of the children in the deworming study. In 2003, he (along with Kremer and an interdisciplinary team of colleagues) launched a follow-up survey known as the Kenya Life Panel Survey (KLPS), with initial funding from the U.S. National Institutes for Health, the World Bank, and others. This funding was for two rounds of data collection: measuring impacts up to 10 years after the PSDP began, on up to 7,500 individuals. The survey was wide-ranging, collecting detailed information on education, health, marriage, fertility, migration, and attitudes of respondents. Furthermore, recognizing the potential of this unusual panel of young rural Kenyans to study other policy relevant research questions, the KLPS research team conducted a second RCT in a subset of the sample, launched just following the close of the second KLPS data collection round, to study the impacts of vocational training vouchers (Hicks et al. 2011).

Early results from these two rounds of KLPS data collection on the longer term impacts of the deworming intervention (Baird et al. 2016) laid the foundation for additional data collection rounds. Round 4 data collection is currently in the field, allowing researchers (including Baird, Hicks, Kremer, and Miguel) to study the 20-year impacts of childhood deworming. Moreover, to date, the KLPS longitudinal data has been used to answer research questions across a range of topics, and far beyond exploring the impacts of childhood deworming and provision of vocational training vouchers: the determinants of rural-urban migration; the evolution of fertility preferences; the impacts of civil unrest; the impacts of secondary school; the effects of urban migration on political attitudes; and so on (Hicks et al. 2017; Mueller et al. 2019; Jakiela and Ozier 2019; Ozier 2018b; Kramon et al. 2019). One important note is that many of these papers (and all of those listed here) are *not* exploiting RCTs, but instead leveraging the rich longitudinal data in the KLPS to answer questions that don't lend themselves to randomization. In this instance, the RCT was the impetus for a unique data collection

that, in turn, directly enabled research on a range of other topics using other methods; allowing the investment in what might narrowly be seen as a 'mere' RCT follow-up to benefit a diverse collection of research projects. While currently only a 20-year panel, the KLPS has the promise to follow in the footsteps of the seven-decade work in Palunpur, India (Himanshu, Lanjouw and Stern 2018). Public archiving of this data will also allow future generations of economists to explore the data for years to come.[4]

Policy engagement

Along with influencing a wealth of academic research, MK2004 also took hold with policymakers. Some policy interest may stem from the origins of the project. Internationaal Christelijk Steunfonds Africa (ICS), an NGO operating in what was then one of Kenya's worst-off regions—Busia—was working with schools and was interested in evaluating some health projects. Michael Kremer proposed an RCT around deworming, which was then implemented by ICS in cooperation with the district offices of the Ministries of Health and Education.

Having started with a research topic of importance to a local NGO did not guarantee the interest of national or international policy communities, however. In order that this be possible, Miguel and Kremer developed relationships with the Ministry of Health and broader Neglected Tropical Disease community in Kenya from the onset, and continued dialogue throughout the project. This dialogue contributed to the adoption of school-based deworming as national policy in Kenya, with annual deworming efforts reaching over 6 million children. In 2007, Deworm the World (https://www.evidenceaction.org/dewormtheworld-2/) was created to coordinate efforts for large-scale school-based deworming programs; programs are now active in India, Ethiopia, Vietnam, and Nigeria.[5] Policy engagement may also have hinged on the study's attention to cost-effectiveness, not only in terms of the very low cost of treatment (59 cents/person)[6], but also in quantifying benefits via the externalities discussed above (Ahuja et al. 2015). This work is now cited by the World Health Organization's (WHO) guidelines on preventative deworming wherever parasite prevalence exceeds 20 percent (WHO 2017).

[4] Public data is available at: https://dataverse.harvard.edu/dataverse/KLPS.
[5] See https://www.poverty-action.org/impact/school-based-deworming for more details.
[6] This cost includes the treatment as well as administration and delivery costs (Baird et al. 2016).

Conclusion

The innovative design of MK2004 and the authors' commitment to policy dialogue has influenced a wide variety of academic work both in Busia, Kenya and elsewhere, and shaped global policy around mass deworming. Although the cost of such long-term and large-scale data collection is non-negligible, we hope we have shown the many ways such an investment can pay off.

While findings from one site are vulnerable to external validity concerns, this article highlights the potential value of setting up a 'field site' for economics research, something more common in other disciplines. As in any piece of social sciences research, of course, even the descriptive evidence we offer may not be externally valid in all settings. There are possible confounds to consider in tracing this RCT's effects as well: for example, at least one other future Nobel prize winner (Esther Duflo) began work with MK in Kenya shortly after the PSDP began. We nevertheless hope the patterns described here will inspire others to explore the possible benefits of long-lasting involvement in their own research settings.

References:

Ahuja, A., Baird, S., Hicks J.H., Kremer, M., Miguel, E., and Powers, S. (2015), "When Should Governments Subsidize Health? The Case of Mass Deworming," *World Bank Economic Review*, Vol. 29 (suppl 1), pp. S9-S24.

Baird, S. Bohren A., McIntosh, C. and Özler, B. 2018. "Optimal Design of Experiments in the Presence of Interference," *The Review of Economics and Statistics*, vol. 100(5), pp. 844-860.

Baird, S., Hamory-Hicks, J., Kremer, M., and Miguel, E. (2016) "Worms at Work: Long Run Impacts of a Child Health Investment." *Quarterly Journal of Economics,* Vol. 131 (4), pp. 1637-1680.

Barrett, C.B. and Carter, M.R. 2010. "The Power and Pitfalls of Experiments in Development Economics: Some Non-random Reflections," *Applied Economic Perspectives and Policy*, Vol. 32 (4) pp: 515–548

Hicks, J.H., Kleemans, M., Li, N., and Miguel, E. 2017. "Reevaluating Agricultural Productivity Gaps with Longitudinal Microdata" *NBER Working Paper #23253*.

Hicks, J.H., Kremer, M., Mbiti, I, and Miguel, E. 2011. "Vocational Education Voucher Delivery and Labor Market Returns: A Randomized Evaluation Among Kenyan Youth," Report for Spanish Impact Evaluation Fund (SIEF) Phase II.

Hicks, J.H., Kremer, M., and Miguel, E. 2015. "Supplementary Appendix for 'Deworming externalities and schooling impacts in Kenya: A Comment on Aiken et al. (2015) and Davey et al. (2015),'" *International Journal of Epidemiology*, Vol. 44 (5), pp. 1593–1596.

Himanshu, Lanjouw, P., Stern, N. 2018. *How Lives Change: Palanpur, India and Development Economics*. Oxford, United Kingdom: Oxford University Press.

Jakiela, P. and Ozier, O. 2019. "The Impact of Violence on Individual Risk Preferences: Evidence from a Natural Experiment," *Review of Economics and Statistics*, 101(3), pp. 547-559.

Knudsen, E.I., Heckman, J.J., Cameron, J.L., and Shonkoff, J.P. 2006. "Economic, neurobiological, and behavioral perspectives on building America's future workforce." *Proceedings of the National Academy of Sciences* 103 (27): 10155–62.

Kramon, E., Baird, S., Hicks, J.H., and Miguel, E. 2019. "Deepening or Diminishing Ethnic Divides? The Impact of Urban Migration in Kenya," *unpublished mimeo.*

Kremer, M., and Miguel, E. 2007. "The Illusion of Sustainability," *The Quarterly Journal of Economics,* 122(3): 1007-1065.

Miguel, E., and Kremer, M. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, Vol. 72, pp. 159-21

Miguel, E. and Kremer, M. 2014. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities, Guide to Replication of Miguel and Kremer (2004)." URL: https://escholarship.org/uc/item/8db127cm

Mueller, M., Hicks, J.H., Johson-Hanks, J., and Miguel, E. 2019. "The Illusion of Stable Preferences over Major Life Decisions," *NBER Working Paper #25844*.

Ozier, O. 2018a. "Exploiting Externalities to Estimate the Long-term Effects of Early Childhood Deworming." *American Economic Journal: Applied Economics*, Vol. 10 (3), pp. 235-262.

Ozier, O. 2018b. "The Impact of Secondary Schooling in Kenya: A Regression Discontinuity Analysis," *Journal of Human Resources,* 53(1): 157-188.

Ozier, O. 2019. "Replication Redux: The Reproducibility Crisis and the Case of Deworming." forthcoming, *World Bank Research Observer*.

Pritchett, L. 2018. "The Debate Over RCTs in Development is over. We won. They lost. URL: http://www.nyudri.org/events-index/2018/2/22/lant-pritchett-talk-the-debate-about-rcts-in-development-is-over-we-won-they-lost

Ravallion, M. 2018. "Should the Randomistas (Continue to) Rule?." *CGD Working Paper 492*. Washington, DC: Center for Global Development. URL: https://www.cgdev.org/publication/should-randomistas-continue-rule

The Guardian. 2018. "Buzzwords and tortuous impact studies won't fix a broken aid system." URL: https://www.theguardian.com/global-development/2018/jul/16/buzzwords-crazes-broken-aid-system-poverty

WHO. 2017. "Guideline: preventive chemotherapy to control soil-transmitted helminth infections in at-risk population groups." Geneva: World Health Organization: Licence: CC BY-NC-SA 3.0 IGO.