Translating Rigorous Evidence into Policies That Benefit the Poor

Quentin Palfrey

J-PAL North America

Abstract

There has been an encouraging level of momentum and interest among policymakers and researchers behind the need for more rigorous and credible evidence to inform public policy decisions. The Abdul Latif Jameel Poverty Action Lab (J-PAL) at MIT has made significant progress in meeting this demand by spurring new research, translating and synthesizing existing bodies of research, empowering policymakers to generate and consume evidence, convening researchers and decision makers, and creating public goods that encourage best practices in rigorous, transparent research. Chief among these best practices are the use of randomized evaluation to rigorously test what works for the poor and the use of administrative data to make such scientific, low-cost research possible. By bringing together academics, policymakers, and practitioners in the field around this set of best practices, we can generate rigorous evidence, gain important policy insights, and translate those lessons into concrete policy action.

Keywords – poverty; evidence; RCT; randomized evaluation; administrative data

1 Growing Demand for Evidence in Policymaking

Impact evaluations are part of a broader, expanding agenda of evidence-based policy making. This growing global trend is marked by a shift in focus from inputs to measurably improving outcomes and results, particularly for the world's most vulnerable citizens. From the Sustainable Development Goals to pay-for-performance incentives for public service providers, this global trend is reshaping how citizens and policymakers alike think about policy priorities, and in turn shaping how public policies are being carried out. Not only is the focus on results being used to set and track national and international targets, but results are increasingly being used by, and required of, governments to enhance accountability, inform budget allocations, and guide policy decisions.

This trend has become particular pronounced in the United States over the last decade. The federal government is institutionalizing evidence-based policymaking in many different ways. Just this year, a bipartisan group in Congress created the Commission on Evidence-Based Policymaking (CEP), housed in the Executive Branch at the U.S. Census Bureau, to improve how the federal government uses survey and administrative data. CEP will examine how best to centralize and make available the enormous amount of survey and administrative data currently collected by various arms of the Federal government. In addition, the commission will look into how to incorporate randomized evaluation into existing and new federal programs.[1]

As an additional extension of this growing emphasis on evidence in the Federal government, in 2015 the White House Office of Science and Technology Policy assembled the Social and Behavioral Sciences Team (SBST)—a cross-agency group of experts in applied behavioral science that translates findings and methods from the social and behavioral sciences into improvements in Federal policies and programs. Already due to SBST projects, more students are going to college and better managing their student loans, more Veterans are taking advantage of education and career counseling benefits, and more families are securing health insurance coverage.[2]

[1] Martinez, Shelly Wilkie. 2016. "Commission on Evidence-Based Policymaking." U.S. Census Bureau. http://www.amstat.org/misc/CEP_OverviewJSM2016.pdf.

[2] Executive Office of the President. 2015. "Social and Behavioral Sciences Team Annual Report." https://www.whitehouse.gov/sites/default/files/microsites/ostp/sbst_2015_annual_report_final_9_14_15.pdf.

2 Best Practices in Responding to and Encouraging Demand for Policy-Relevant Evidence

2.1 *Randomized Evaluation*

In this context in which policymakers and their constituents are demanding results and accountability from public programs, impact evaluation can provide robust and credible evidence on whether a particular social policy achieved its intended outcomes. Simply put, an impact evaluation assesses the changes in the well-being of individuals that can be attributed to a particular project, program, or policy. This focus on attribution is the hallmark of impact evaluations; correspondingly, the central challenge in carrying out effective impact evaluations is to identify the causal relationship between the program and the outcomes of interest.

To be able to estimate the causal effect or impact of a program on outcomes, any method chosen must estimate the so-called counterfactual: what the outcome would have been for program participants if they had not participated in the program. Since it is impossible to directly observe this, methods of impact evaluation aim to create a comparison group of non-participants who closely resemble the participants. All else equal, the more closely the comparison group mirrors the participants before the start of the program, the more confident we can be that any observed differences in outcomes after the program are due to the program itself.

Random assignment is generally recognized as the most rigorous way to create a valid comparison group. In a randomized evaluation, participants are randomly assigned to either a treatment group that receives the program or to a comparison group that does not receive the program. With a large enough sample, random assignment creates two groups that are equivalent, on average, at the start of the program. This allows the evaluator to attribute any subsequent differences between the treatment and control groups to the program rather than to other factors. Other evaluation methods typically require stronger assumptions about the equivalence of the two groups at the start of the program in order to draw any causal inferences about the effectiveness of the program. Consequently, results from randomized evaluations can be easier to communicate and less subject to confusion or controversy than results from other methods. Because of these advantages, the Abdul Latif Jameel Poverty Action Lab (J-PAL) at the Massachusetts Institute of Technology (MIT) focuses its efforts on generating evidence based on randomized evaluations.

2.2 *Administrative Data*

In addition to randomly assigning treatment and comparison groups, evaluations can take advantage of data that are already gathered for operational purposes, such as criminal justice or health records. Compared to new data collected through surveys, using these administrative data—when handled properly, with appropriate privacy safeguards and other precautions—can substantially reduce evaluation costs, allow for faster turn-around of results, enable long-term follow-up, and can improve the accuracy of study findings.

Data gathered by government agencies, school systems, police departments, and hospitals for purposes other than evaluation has powered some of the most important and policy-relevant research produced over the last several years, including in some of the evaluations discussed in the next section. In the 21st century, these data are now often collected and stored digitally, and accessing this data for research purposes can open up a world of possibilities. Governments are making strides to facilitate access to this valuable data.

3 Application of Best Practices

One example for the implementation of these best practices is the body of work focused on improving the outcomes of at-risk young men living in violent environments around the world. Violence and related social ills are costly to governments and civil society in high- and low-income countries alike, but disproportionately impact the poor. In many environments, poor young men with limited economic opportunities drive high rates of crime and violence, and may be targets for mobilization into destructive activities. Poor citizens are also more likely to be victims of violence.

A large body of largely observational evidence in psychology research in the United States has shown that cognitive behavior therapy (CBT), a therapeutic approach to improving a wide range of harmful beliefs and behaviors, can be an effective way to reduce violence and criminality among children and adolescents. To understand the potential effectiveness of CBT among men in different contexts, J-PAL affiliated researchers conducted randomized evaluations of CBT interventions among criminally-engaged men in post-war Liberia and among young men of color on the South Side of Chicago.

In Liberia, researchers recruited criminally-engaged men–almost 40 percent of whom were former soldiers and one-quarter of whom were homeless–and randomly assigned half to eight weeks of group cognitive behavioral therapy, teaching self-control skills and a noncriminal self-image. Relative to the men in the comparison group, cognitive behavioral therapy increased self-control and noncriminal

values, and acts of crime and violence fell between 20 and 50 percent.[3]

In another study conducted in Chicago, affiliates of J-PAL and the University of Chicago Crime Lab investigated whether CBT could be applied in violence-prone cities in the United States.

In 2013, Chicago had more known gang members than any other city and had more illegal guns recovered than any other city, creating a particularly difficult and violent environment for young men.[4] In 2012, Chicago's murder rate was four times the national average.[5]

The Becoming a Man (BAM) intervention, a CBT-based approach to group therapy, seeks to teach Chicago youth to stop, look, and listen and to avoid maladaptive automatic responses to everyday situations. In this study, program participants were young men in Chicago Public Schools, many of whom were at-risk for dropout or violent behavior and about one-third of whom had been previously arrested. BAM sessions focused on examining automatic thought processes for 27 weeks for 1 hour per week as part of the school day.

Researchers linked administrative datasets from the Chicago Public Schools and the Illinois State Police to measure educational and crime outcomes of BAM participants relative to a comparison group. They found that CBT improved schooling outcomes and reduced violent-crime arrests by 44 percent in one study and reduced overall arrests by 31 percent in a second study.[6]

Based on the success of these CBT programs, both programs have been scaled up in their respective contexts. BAM, in particular, has received attention and support from private funders, the local government in Chicago, and even from the federal government.

4 Innovating Further to Make 'What Works' Work Better

Although BAM improved some schooling outcomes, it is reasonable to assume that these at-risk students were already so far behind that it might be hard to catch up. Researchers and policymakers alike wanted to know whether other innovative programs could improve educational outcomes even more.

Researchers decided to once again apply evidence from international contexts to the U.S. context; tutoring programs in Indian and Kenya designed to assess the academic level of students and teach at that level have shown large impacts for students who fall behind.[7]

Researchers implemented a 1-hour-per-day individualized math tutoring component in Chicago Public Schools to directly address the mismatch between student preparation and the in-class instruction provided by their schools. This simple intervention improved student math test scores and grades and reduced failure. Students learned an additional one to two years of math in a single school year, and the program narrowed the black-white test score gap by one-third. It also reduced course failures overall by one-quarter, suggesting that there were achievement improvements beyond the tutored subject.[8]

This teaching at the right level intervention provides yet another important case study in applying the core concepts of effective interventions from one context to other contexts and in pushing the boundaries of one intervention to make it even more effective. By disseminating the lessons from evaluations like this, we learn how to make policies aimed at improving the lives of the poor even more effective.

5 Looking Ahead

There is so much more that researchers, policymakers, and practitioners can do—both on their own and together—to create and utilize rigorous, policy-relevant evidence on what works in poverty alleviation. Researchers must ensure that existing policy lessons from rigorous research are disseminated clearly and widely to help policymakers better align scarce resources with evidence. J-PAL helps to do this by creating publications aimed at policymakers sharing the policy lessons of J-PAL affiliated research.[9]

In addition, new and further collaboration between researchers and policymakers can be mutually beneficial by responding to the most important questions facing policymakers and practitioners while providing new and exciting opportunities for rigorous academic work. Some collaboration has been facilitated by the increasingly popular Pay for Success (PFS) model, first launched in the United Kingdom in 2012 and now underway in various forms in nearly 40 states. PFS financing is a new approach

---

[3] Blattman, Christopher, Julian C. Jamison, and Margaret Sheridan. 2015. "Reducing Crime and Violence: Experimental Evidence on Adult Noncognitive Investments in Liberia." Working Paper.

[4] Heller, Sara B., Anuj K. Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold A. Pollack. Forthcoming. "Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago" *Quarterly Journal of Economics.*

[5] Federal Bureau of Investigation. Uniform Crime Reporting Statistics. http://www.ucrdatatool.gov/index.cfm.

[6] Heller et al. (forthcoming).

[7] Banerjee, Abhijit V., Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *Quarterly Journal of Economics* 122(3): 1235-64; Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review* 101(5): 1739-74.

[8] Cook, Philip J., Kenneth Dodge, George Farkas, Roland G. Fryer Jr, Jonathan Guryan, Jens Ludwig, Susan Mayer, Harold Pollack, and Laurence Steinberg. 2015. "Not Too Late: Improving Academic Outcomes for Disadvantaged Youth." Institute for Policy Research Working Paper Series WP-15-01.

[9] J-PAL "Evaluations." https://www.povertyactionlab.org/evaluations.

to scale social services that relies on an independent evaluator to generate evidence about the impact of the services delivered. This approach leverages public-private partnerships to improve outcomes for vulnerable populations.[10] PFS and approaches like it also require more research partnerships in order to generate rigorous evidence of impact.

J-PAL North America's State & Local Innovation Initiative and Health Care Delivery Initiative directly connect researchers to innovative practitioners and policymakers to facilitate randomized evaluations that respond directly to policymakers' priorities. J-PAL has also created public resources to help researchers become aware of and gain access to administrative data[11] and to help policymakers apply randomized evaluation in their own jurisdictions.[12] J-PAL has also helped to build capacity within the policy community to both create and consume rigorous evidence through training and technical support.

There is certainly more to do to institutionalize these practices at all levels of government and further spread the best practices of randomized evaluation and the use of administrative data. We hope that others in the evidence-based policymaking community will help to amplify and build upon these efforts.

References

Banerjee, Abhijit V., Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *Quarterly Journal of Economics* 122(3): 1235-64

Blattman, Christopher, Julian C. Jamison, and Margaret Sheridan. 2015. "Reducing Crime and Violence: Experimental Evidence on Adult Noncognitive Investments in Liberia." Working Paper.

Cook, Philip J., Kenneth Dodge, George Farkas, Roland G. Fryer Jr, Jonathan Guryan, Jens Ludwig, Susan Mayer, Harold Pollack, and Laurence Steinberg. 2015. "Not Too Late: Improving Academic Outcomes for Disadvantaged Youth." Institute for Policy Research Working Paper Series WP-15-01.

Duflo, Esther, Pascaline Dupas, and Michael Kremer. 2011. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." *American Economic Review* 101(5): 1739-74.

Executive Office of the President. 2015. "Social and Behavioral Sciences Team Annual Report." https://www.whitehouse.gov/sites/default/files/microsites/ostp/sbst_2015_annual_report_final_9_14_15.pdf.

Federal Bureau of Investigation. Uniform Crime Reporting Statistics. http://www.ucrdatatool.gov/index.cfm.

Heller, Sara B., Anuj K. Shah, Jonathan Guryan, Jens Ludwig, Sendhil Mullainathan, and Harold A. Pollack. Forthcoming. "Thinking, Fast and Slow? Some Field Experiments to Reduce Crime and Dropout in Chicago" Quarterly Journal of Economics.

J-PAL "Evaluations." https://www.povertyactionlab.org/evaluations.

J-PAL. "Research Resources." https://www.povertyactionlab.org/research-resources.

J-PAL North America. "Using Administrative Data for Randomized Evaluations." https://www.povertyactionlab.org/node/21362.

Martinez, Shelly Wilkie. 2016. "Commission on Evidence-Based Policymaking." U.S. Census Bureau. http://www.amstat.org/misc/CEP_OverviewJSM2016.pdf.

Pay for Success Learning Hub. "Pay for Success U.S. Activity." Accessed August 30, 2016. http://www.payforsuccess.org/pay-success-deals-united-states.

[10] Pay for Success Learning Hub. "Pay for Success U.S. Activity." Accessed August 30, 2016. http://www.payforsuccess.org/pay-success-deals-united-states.

[11] J-PAL North America. "Using Administrative Data for Randomized Evaluations." https://www.povertyactionlab.org/node/21362.

[12] J-PAL. "Research Resources." https://www.povertyactionlab.org/research-resources.