

# Building Momentum for Evidence-Based Policymaking in State and Local Governments

Hannah Myers and Rohit Naimpally

*J-PAL North America, Massachusetts Institute of Technology*

[hmyers@mit.edu](mailto:hmyers@mit.edu)

[rnaimpally@povertyactionlab.org](mailto:rnaimpally@povertyactionlab.org)

## Abstract

What if policymakers allocated government funding based on scientific evidence of what works, instead of anecdote, status quo, or political belief? Despite increasing pressure to use “big data” to inform decisions, it’s often challenging for policymakers to disaggregate the impacts of a specific program from broader economic and societal conditions. By using data they already collect and applying the same scientific tool that transformed modern medicine—randomized evaluations—to social policy, researchers and policymakers can work together to cut through opinion and build a foundation of rigorous evidence in its place. However, reorienting government decision-making to identify and fund effective programs is often slow and complex work. Using examples from the United States, we offer six concrete steps policymakers can take to systematize the effective use of data and ensure the greatest return on taxpayer dollars.

*Keywords* – randomized controlled trials; impact evaluation; government; social services; research; policy

## Evidence-Based Policymaking

What if policymakers allocated government funding based on scientific evidence of what works, instead of anecdote, status quo, or political belief? In most major policy debates, compelling arguments can be made on both sides without necessarily being informed by empirical evidence. Despite increasing pressure to use “big data” to inform decisions, it’s often challenging for policymakers to disaggregate the impacts of a specific program from broader economic and societal conditions. By using data they already collect and applying the same scientific tool that transformed modern medicine—randomized evaluations—to social policy, researchers and

policymakers can work together to cut through opinion and build a foundation of rigorous evidence in its place.

While bipartisan consensus is rare in the current US political climate, Republican House Speaker Paul Ryan, Democratic Senator Patty Murray, and President Obama came together in 2016 to enact legislation creating a federal [Commission on Evidence-Based Policymaking](#). The commission aims to develop a strategy for increasing the availability and use of data to build evidence about government programs. State and local governments—which are [collectively responsible for spending \\$2.5 trillion each year \(about 40 percent of total government spending\)](#)—are also joining in. In Washington State, for example, state human services departments [track and report](#) the percentage of funding allocated for evidence-based and/or research-based programs.

Reorienting government decision-making to identify and fund programs that work can be slow and challenging, but can make a real difference in people’s lives. We offer six concrete steps policymakers can take to use data effectively and ensure the greatest return on taxpayer dollars.

### 1. Facilitate access to administrative data for research and impact evaluation.

Hospitals, governments, school systems, and other institutions gather a wealth of information on individuals as part of their regular operations. This data can be an excellent source of information for research when equipped with safeguards for privacy, and can reduce research costs, create more possibilities for long-term follow-up, and improve the accuracy of findings.

Several randomized evaluations using administrative data have already transformed what we know about social policy — and helped policymakers make better decisions. Economists [Raj Chetty](#), [Nathaniel Hendren](#), and [Lawrence Katz](#) tracked the long-term outcomes of families who left high-poverty areas in the mid-1990s through the Moving to Opportunity housing choice voucher program. By matching income records from nearly 20 years later, they found that young children from families that moved [earned significantly higher incomes, attained more education, and became single parents at lower rates compared to their peers who stayed](#). Citing this research, the [Department of Housing & Urban Development](#) is considering overhauling the formula that has been in use for four decades to calculate rental assistance to see if it can [increase opportunities for families](#) to move to low-poverty areas.

In health care, economists [Katherine Baicker](#) and [Amy Finkelstein](#) used a randomized evaluation to track the behavior of individuals who gained access to Medicaid in a lottery system in Oregon in 2008. They learned that Medicaid [led people to use more healthcare services](#) across the board—preventative care, hospitalizations, emergency department visits, doctor office visits, and prescription drug use—while [reducing financial strain and rates of depression](#). This study provided hard facts and continues to inform the politically charged Medicaid expansion debate.

It is hard to overstate the transformational potential these data can have for researchers seeking to determine what works—and what does not—in social policy. Changing the default access so data sets are available instead of unavailable, unless sensitive data could be disclosed or there are privacy concerns, would open the door to myriad more important studies and policy insights.

## **2. Require that agencies link administrative datasets.**

Government programs often affect people's lives along a variety of dimensions. [A housing program can have a profound effect on health](#), an education program can affect students' job prospects, or a program for substance use disorder treatment can influence the likelihood that patients will get into trouble with the law. Linking administrative records from different government agencies enabled the influential analysis of the housing voucher lottery discussed above. Researchers conducting similar randomized evaluations need to match individuals at the start of a program (e.g. Housing Authority data) with their respective outcomes (e.g. income records).

These outcomes can be hard to detect if data are confined to silos within agencies that do not work together.

Policymakers should insist that agencies think collaboratively about how to link data; South Carolina, for example, has established [an integrated data system](#) to make it easier for government and evaluators to study the impact of programs.

One such study looks at the impact of the [Nurse-Family Partnership](#) (NFP), a nurse-home visiting program for low-income, first-time mothers. Through the program, specially trained nurses provide holistic support from early pregnancy through the child's second birthday. Using linked administrative data from multiple agencies to understand health and wellbeing comprehensively, researchers will measure NFP's short- and long-run impact on a wide range of health, economic, and other outcomes for mothers and their children.

Governments should also flag sensitive data. Currently, researchers may request data that is only tangentially relevant to their analysis, without realizing that it captures sensitive information that could unnecessarily jeopardize the entire request. Knowing which data are sensitive can direct researchers toward questions they can feasibly answer.

## **3. Add requirements and support for rigorous evaluation into existing funding streams.**

When authorizing pilot programs, lawmakers should encourage agencies to roll them out in a way that allows agencies and scholars to compare the effect of the programs on those who receive services from them against a statistically identical population that receives only pre-existing services. As one example, federal law [previously mandated](#) that Washington, D.C.'s school voucher program be administered through a lottery and evaluated with the highest level of rigor—in this case a randomized controlled trial. Though it has since been removed, evaluations that were made possible through this requirement contributed to our understanding of how school vouchers work—a critical policy question in the United States at present.

Policymakers should seek out technical assistance for agencies to become better producers and consumers of evidence and to create a culture of evaluation across a jurisdiction's executive branch.

## **4. When allocating scarce resources to oversubscribed programs, consider determining eligibility by lottery rather than first-come-first-served.**

Lotteries can be the fairest and most transparent way to select individuals off a wait list. When allocating slots on a first-come-first-served basis, for example, intangible factors such as an applicant's motivation, connections, or resources can influence who was able to sign up the soonest and thus receive services. A lottery design ensures that those who receive services and those who do not are on average similar as groups on a wide range of characteristics that can and cannot be measured. This is not only a more fair allocation, but it also allows researchers to measure the causal impact of a policy.

Because lotteries are widely accepted as a fair way to allocate scarce resources, governments sometimes use lotteries even when there is no planned research. For example, the state of Oregon chose to expand Medicaid to previously ineligible applicants by lottery in 2008. This approach allowed Katherine Baicker and Amy Finkelstein to later follow up on applicants and accurately measure the impact of the policy on beneficiaries' financial circumstances and physical and mental health, as discussed above.

However, there are real-world challenges that can sometimes make [randomized designs challenging](#). In some cases, program implementers have enough resources and may feel obligated to distribute a program to everyone in the study area. If everyone receives the program at once, there are no individuals to construct a comparison group for a randomized evaluation.

Researchers Robert Fairlie (University of California, Santa Cruz) and Jonathan Robinson (University of California, Santa Cruz) faced this challenge when they [conducted an evaluation](#) to determine the impact of home computers on academic achievement. Under the program, students who did not have a home computer were eligible to receive one. To ensure a fair distribution of computers while preserving a comparison group, the program was phased in across schools over time. A first group of schools was randomly selected to receive computers at the beginning of the school year, with the remaining schools receiving the computers at the end of the school year. An endline survey was administered at the end of the school year, just before the second group of schools received their computers. By comparing the outcomes of students in the first group to those in the second group of schools, the researchers were able to measure the impact of receiving the computers on academic achievement.

This study used two phases, but it is also possible to use multiple phases, depending on several factors that shape the context of the evaluation (e.g., the length of the intervention, the number of randomization units, the particular outcomes being studied). Additionally, if a

randomized evaluation occurs as a program expands its capacity, there may be a need to train additional staff or troubleshoot logistical challenges that come with scale. In these cases, instead of scaling up and implementing the program all at once, the phase-in design provides extra time to scale up carefully while maintaining fidelity to the program model.

In other cases, it can be challenging for service providers to give a program exclusively to individuals in one group, and leave individuals in the comparison group without services. This can be especially challenging when those receiving the program and those not receiving it are in close proximity. In these situations, assigning program and comparison groups to different service providers, or randomizing at a higher level, i.e. at the clinic level instead of by individual, can address these concerns.

## **5. Institutionalize best practices by creating independent evaluation offices.**

Independent evaluation offices are tasked with identifying opportunities for randomized evaluations and other rigorous research, linking administrative datasets across state agencies to facilitate these evaluations, and applying existing evidence to improve the efficacy of state programs. For example, the Washington state legislature established [the Washington State Institute for Public Policy](#) to carry out policy-relevant research and to use this evidence to advise legislators. In a similar model with a focus on behavioral science, [Philadelphia](#) and [Washington, D.C.](#), have created behavioral "nudge" units, along the lines of [the White House Social and Behavioral Sciences Team](#), to apply research from behavioral science to improve the efficacy of their programs.

Similarly, states should take a page from Congress' book by establishing state-level commissions on evidence-based policymaking. These would be charged with carrying out systematic reviews of existing data and evaluation infrastructure and finding better ways to institutionalize the generation and use of evidence in government policy.

## **6. Consider using results-based funding structures, or “Pay for Success” programs, that embed evaluation into policy design.**

Pay for Success (PFS) ties payments for service delivery to measurable outcomes, which are typically assessed by an independent evaluator. This approach to contracting is often both politically attractive—as taxpayers only pay for programs that meet pre-specified objectives—and creates an opportunity for rigorous research on what works in

social services from healthcare to criminal justice to education.

The first PFS, or social impact bond, project launched in the United Kingdom in 2010, and since then the concept has gained traction both in the UK and abroad. In 2016, the UK Government launched a series of funds for the development of PFS projects and significant legislation to support PFS projects has been introduced with bipartisan support in the U.S. Senate.

The South Carolina Department of Health and Human Services (SCDHHS) is leading an innovative expansion of the Nurse Family Partnership, discussed above, through a Medicaid waiver coupled with a PFS contract. The PFS contract structure has enabled a new randomized controlled trial to build on previous research, aiming to understand the impact and cost-effectiveness of the program in a modern context, for a different population, and operating at larger scale.

By implementing these kinds of measures, government policymakers can build the infrastructure to design and fund programs that work—a win-win for policymakers and for the citizens they serve.

## Acknowledgements

We would like to thank Mary Ann Bates, Julia Chabrier, Laura Feeney, Quentin Palfrey, Vincent Quan, and Emma Rackstraw for their help creating this paper.

We would also like to thank the Laura and John Arnold Foundation, the Robert Wood Johnson Foundation, and the Sloan Foundation for their generous support of J-PAL North America to make this work possible.

## References

Baicker, K., Finkelstein, A., Song J., & Taubman, S. (2014). The Impact of Medicaid on Labor Force Activity and Program Participation: Evidence from the Oregon Health Insurance Experiment. *American Economic Review Papers and Proceedings*, 104(5), 322-28.

Baicker, K., Taubman, S., Allen, H., Bernstein, M., Gruber, J., Newhouse, J., Schneider, E., Wright, B., Zaslavsky, A., and Finkelstein, A. (2013). The Oregon Experiment—Effect of Medicaid on Clinical Outcomes. *The New England Journal of Medicine*, 368(18), 1713-1722.

Chetty, R., Hendren, N., & Katz, L. (2016). The Effects of Exposure to Better Neighborhoods on Children: New Evidence

from the Moving to Opportunity Experiment. *American Economic Review*, 106(4), 855-902.

Fairlie, R. & Robinson, J. (2013). Experimental Evidence on the Effects of Home Computers on Academic Achievement among Schoolchildren. *American Economic Journal: Applied Economics*, 5(3), 211-240.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J., Allen, H., & Baicker, K. (2012). The Oregon Health Insurance Experiment: Evidence from the First Year. *Quarterly Journal of Economics*, 127(3), 1057-1106.

Heard, K., O'Toole, E., Naimpally, R., & Bressler, L. (2017). Real-World Challenges to Randomization and Their Solutions. *J-PAL North America Report*: Massachusetts Institute of Technology.

Ludwig, J., Duncan, G., Gennetian, L., Katz, L., Kessler, R., Kling, J., & Sanbonmatsu, L. (2013). Long-Term

Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity. *American Economic Review Papers & Proceedings*, 103(3), 226–231.

Ludwig, J., Sanbonmatsu, L., Gennetian, L., Adam, E., Duncan, G., Katz, L., Kessler, R., Kling, J., Tessler Lindau, S., Whitaker, R., & McDade, T. (2011). Neighborhoods, Obesity, and Diabetes — A Randomized Social Experiment. *The New England Journal of Medicine*, 365 (16), 1509–1519.

Taubman, S., Allen, H., Wright, B., Baicker, K., & Finkelstein, A. (2014). Medicaid Increases Emergency Department Use: Evidence from Oregon's Health Insurance Experiment. *Science*, 343(6168), 263-268.