



PROGRAM EVALUATION

Building upon foundations for evidence-based policy

Can government culture evolve under a promising new law?

By **Robert Hahn**^{1,2}

There is no natural constituency for evidence-based policy. It should, by rights, be the public who wants the most from their government (and their public funds). But the public, like most politicians, is often not aware of the ins and outs of evaluation methods and evidence. Think tanks and academics have long filled this gap and will likely continue to play key roles. But legislation signed into law in early 2019 could transform the way U.S. government officials design programs by introducing more scientific evidence into the process. On the basis of recommendations of the bipartisan U.S. Commission on Evidence-Based Policymaking (1), this *Foundations for*

Evidence-Based Policymaking Act of 2018 (2) (“Evidence Act”) could help usher in a cultural shift toward evidence-based policy, and in so doing shed new light on many policy challenges, including welfare, crime prevention, drug abuse, and inequality [e.g., (3)]. This act received broad support from Congress (4), suggesting that many legislators are beginning to see the value of making sure the programs that they fund deliver on their promise, and of making the policymaking process more transparent. But the act is just a foundation, not the full building. Getting from this act to more effective policy outcomes means getting departments and agencies to buy into a new culture where rigorous evaluation matters more in designing and funding programs.

The history of evidence building for policy in the United States goes back to the founders. The Constitution required a census of the population. Over the past two centuries,

Congress established several units to produce national statistics. More recently, there has been an interest in using data to try to improve the effectiveness of government programs. For example, the 1993 Governance Performance and Results Act (GPRA)—updated in 2010—was supposed to modernize agency management practices. Agencies would set goals, measure results, and report on progress. In other words, it looks a lot like an attempt at legislating an evidence base for policy. Although it does not provide for rigorous evaluation of individual policy tools, it does lay out steps for measuring results. But despite strong support for the act from Congress, agencies struggled. A lack of data and experience, difficulties with inter-departmental programs, weak leadership, and a failure to develop coherent strategies have all been cited as contributing factors to early disappointing results of the act (5, 6).

A 2017 government-wide survey of federal managers by the Government Accountability Office suggests that the absence of program evaluation is a serious problem (7). Only 40% of managers reported that an evaluation had been completed within the past 5 years of any program, operation, or project that they were involved in. Another 39% reported that they did not know if an evaluation had been completed, and 18% reported having no evaluation.

Notwithstanding such problems, the Commission on Evidence-Based Policymaking cited a number of cases where rigorous program evaluation made a difference in the implementation of specific programs. Examples included policies aimed at providing stable housing for the homeless; policies aimed at reducing drug abuse (many of which were not effective initially); and policies that provide job training (1).

PRIVACY, LEARNING, ACCESS

The new law aims to build on these successes. It will put three building blocks in place to help promote more and better program evaluation in federal executive agencies. First, it takes data privacy very seriously. Citizens want assurances that their personal information collected by the government is used appropriately. Without such assurances, there is a real risk that the public will not be willing to share data for statistical uses, which is essential for good program evaluation.

Responding to this concern, the Evidence Act requires comprehensive risk assessments related to the release of certain data, and it also codifies language directing key agencies to protect trust in data by maintaining confidentiality. It helps create a potential “win-win” opportunity that could improve privacy protections and yield new policy insights (8). For example, the Census Bureau is already

¹Smith School of Enterprise and the Environment, University of Oxford, Oxford, UK. ²Technology Policy Institute, Washington, DC, USA. Email: bob.hahn@gmail.com

using a differential privacy algorithm to help preserve privacy while making data available for research (1). In addition, there are opportunities to promote privacy and share de-identified data with researchers by using better technology and procedures.

Second, the law requires agencies to implement a learning agenda related to evidence building that would align evaluation efforts with key questions that policymakers want answered. Agencies will have to submit an annual plan for building evidence to the Office of Management and Budget (OMB), and designate senior civil servants for managing data and overseeing the generation of evidence. For example, the “Research Roadmap” developed by the Department of Housing and Urban Development identifies key research challenges relevant to the agency’s mission (1).

Third, the law would make it easier for government and academic researchers to gain access to data, and for the public to know what kinds of data are being made available. The act requires the development of inventories listing available information, the default publication of nonsensitive data in machine-readable format, and a single portal for applications to restricted sources.

The law could thus help agency staff as well as outside researchers to evaluate policy and to contribute in an informed manner to the development of more-effective interventions—all while adding safeguards to protect individual privacy. These are laudable outcomes. But embedding a culture of evidence-based policy in government—an environment where the first reaction to an idea is to ask what evidence there is—will require far more work. We need to create more and better evidence, and then ensure that it is used when policy is made. The increased availability of data to outside researchers should help with this objective. There is always the chance that a failed scheme will be exposed by academics or journalists digging through the numbers. Stronger incentives, however, will be needed to embed it more fully.

EVIDENCE ON THE USE OF EVIDENCE

Agency difficulties since the GPRA was enacted can give us insight into what the next steps should be. Agencies need capacity to conduct evaluations. It is not enough to simply designate someone as responsible, and ask for a plan. Departments and agencies need expertise and a coherent approach, and that means funding. Congress could fund new positions within agencies specifically devoted to evaluating policies, programs, and regulations. Alternatively, Congress could require that a fraction of program funding, say 0.5 to 1%, is actually

used on ongoing evaluations. It would be useful if the level of funding were linked to the expected net benefits from an agency’s evaluation activities, perhaps giving the OMB the authority to reward agencies that are performing their evaluation task more effectively. Finally, Congress should require that all federal agency program evaluations be summarized and posted at a single website (with the exception of those whose posting could adversely affect other public policy goals, such as national security). This will allow the public, researchers, and policymakers to easily see what interventions work or don’t work in particular contexts.

Agencies also need to learn from best practices, and be able to work together on the evaluation of interdepartmental programs. As a procedural matter, the newly created position of Evaluation Officer should report directly to the head of a department or agency. This would elevate the importance of this function within the agency, and could give these officers greater clout in doing work that requires cooperation among agencies. At the same time, for the evaluation officer to be effective within an agency, that person will need to have close ties with research and data divisions.

This act puts the building blocks in place to produce better evidence, but incentives may also be needed to use evidence. We need to be able to see the evidence that contributed to decisions, and we need to be able to measure performance. In other words, we need to develop evidence on the use of evidence. This issue of how evidence should be used in managing programs is a matter for Congress to decide. In addition to funding, Congress might consider a requirement that large programs undergo a rigorous evaluation of their effectiveness, and that the decision on future funding is connected to that evaluation (9, 10). This would force decision-makers to define what success looks like in concrete terms—a challenging task. Ideally, the definition of effectiveness should include the application of modern methods of program evaluation where feasible [e.g., (11, 12)], and some measure indicating that a program’s benefits will likely exceed its costs (13).

If the Evidence Act is viewed as a success in terms of its impact on the production of new knowledge or its impact on improving decisions, Congress should consider extending its scope in three ways. First, it should require independent agencies like the Federal Reserve Board and the Federal Communications Commission to be part of new legislation. There is no reason, in principle, that federal independent agencies should have been excluded. Second, a revision to the act could address the important role

that states and locales play in generating insights on the effectiveness of programs, and possibly provide funding for innovative data-sharing arrangements that could benefit citizens in other states. Many states have already taken creative approaches to assembling, sharing, and using administrative data to promote evidence-based policy, such as the Washington State Institute for Public Policy and the Michigan Interagency Council on Homelessness (14). There is also an opportunity to learn from what other countries, such as Canada, the United Kingdom, and Germany, are doing to promote evidence-based policy, data sharing, and transparency (1, 15).

Third, Congress may want to increase the transparency of agency decision-making by requiring that data and models that form the basis for major policy and regulatory decisions are identified and shared where appropriate. The Evidence Act is a good starting point for moving government toward better decision-making grounded in scientific evidence. But without further efforts, that is precisely what it will remain. ■

REFERENCES AND NOTES

1. <https://www.cep.gov/news/sept6news.html>. The recommendations received unanimous support from commission members, which may be one of the reasons Congress and the President supported the final law.
2. www.congress.gov/bills/115/congress-house-bill/4174
3. R. Chetty, N. Hendren, P. Kline, E. Saez, *Q. J. Econ.* **129**, 1553 (2014).
4. <https://www.congress.gov/bills/115th-congress/house-bill/4174/all-actions?overview=closed&q=%7B%22roll-call-vote%22%3A%22all%22%7D>
5. B. A. Radin, *Public Adm. Rev.* **58**, 307 (1998).
6. N. Hart, K. Newcomer, “Presidential Evidence Initiatives: Lessons from the Bush and Obama Administrations’ Efforts to Improve Government Performance,” Technical Paper, Bipartisan Policy Center (2018); <https://bipartisanpolicy.org/wp-content/uploads/2018/03/Presidential-Evidence-Initiatives.pdf>.
7. Government Accountability Office, Program Evaluation: Annual Agency-Wide Plans Could Enhance Leadership Support for Program Evaluations, GAO-17-743 (2017). See pp. 32–35 for a discussion of methodology.
8. J. M. Abowd, I. M. Schmutte, *Am. Econ. Rev.* **109**, 171 (2019).
9. J. Tirole, in *Economics for the Common Good* (Princeton Univ. Press, 2017), p. 373.
10. J. Baron, www.straighttalkonevidence.org/2019/01/11/fdrs-call-for-bold-persistent-experimentation-antidote-to-40-year-stagnation-in-key-areas-of-social-policy-part-two-in-a-series/.
11. J. D. Angrist, J.-S. Pischke, *Mostly Harmless Econometrics* (Princeton Univ. Press, 2009).
12. J. Pearl, D. Mackenzie, *The Book of Why* (Basic Books, 2018).
13. K. J. Arrow et al., *Science* **272**, 221 (1996).
14. K. Reynolds, K. Ramakrishnan, *Evidence-Based Policymaking at the State Level: A Guide for Governors* (2018); www.urban.org/sites/default/files/publication/99293/evidence-based_policymaking_at_the_state_level.pdf.
15. J. Lane, *Ann. Am. Acad. Pol. Soc. Sci.* **675**, 240 (2018).

ACKNOWLEDGMENTS

R.H. served as a member of the U.S. Commission on Evidence-Based Policymaking. R.H. thanks J. Baron, N. Hart, R. Metcalfe, N. Potok, and R. Shea for discussions. S. Ashworth-Hayes helped with this essay.

10.1126/science.aaw9446