6. BUILDING AND USING EVIDENCE TO IMPROVE GOVERNMENT EFFECTIVENESS

An effective and efficient Federal government requires evidence—evidence about where needs are greatest, what works and what does not work, where and how programs could be improved, and evidence about how programs of yesterday may no longer be suited for today or prepare us for tomorrow. Strong evidence about policies and programs should be acted upon, suggestive evidence should be considered, and where evidence is weak it should be built to enable better decisions in the future. Agencies should integrate quality evidence and rigorous evaluation into budget, management, and policy decisions through a broad set of activities. Doing so requires the infrastructure and capacity to credibly build and use evidence and develop a culture of learning and continuous improvement. With a strong evidence infrastructure and culture agencies constantly (1) ask and answer questions that help them find, implement, and sustain effective programs and practices, (2) identify and improve or eliminate ineffective programs and practices, (3) test promising programs and practices to see if they are effective and can be replicated, and (4) find lower cost ways to achieve better results.

Building a Portfolio of Evidence

Government agencies should use a range of evidence types and analytical and management tools to learn what works and what does not, for whom and under what circumstances, and how to improve results. Evidence refers to facts or information indicating whether a belief or proposition is true or valid. Evidence can be quantitative or qualitative and may come from a variety of sources, including performance measurement, program evaluations, statistical series, retrospective reviews, data analytics, and other science and research. A portfolio of evidence may include:

- Impact evaluations, including randomized control trials and rigorous quasi-experimental designs, which can answer questions about a program's impact relative to a counterfactual—i.e. whether the outcome was achieved because of the program or due to some other factor.
- Process or implementation evaluations that can answer questions about whether a program is implemented as designed and whether the program structure is sound.
- Performance monitoring and measurement that can answer questions about program efficiency, outputs, and outcomes, but not about causal impact.
- Statistics and other forms of research and analysis that can provide insight into trends, strategies, and underlying processes.

There are multiple ways to assess policies and programs. The best approach or method depends on the specific information that is needed to answer key policy, programmatic, or operational questions, and on practical and methodological considerations. While many forms of evidence are complementary, some evidence that is useful for one purpose may not be useful for another. For example, performance measures are an essential resource for agencies to understand ongoing, real-time program performance so they can use that information to build a culture of continuous improvement, but they often do not answer certain key questions, including the effects of programs. Evaluations provide context for the performance measures and help us better understand what can and cannot be learned from them. In particular, rigorous impact evaluations, especially randomized experiments, can provide the most credible information on the impact of the program on outcomes, isolated from the effects of other factors. Combining performance and evaluation information, and using the results of one to inform the design of the other, can be very powerful in understanding program performance and ensuring that a program is maximizing performance and impact on an ongoing basis.

One example of building evidence to improve government effectiveness in the FY 2018 Budget is at the Department of Education, which is refocusing and expanding its signature tiered evidence program, Education Innovation and Research (EIR), to provide grants to implement and evaluate innovative approaches to supporting private school choice. The President's Budget requests $370 million for EIR, with $250 million reserved for building evidence on the effectiveness of private school choice programs. In another example from the Budget, the Administration is requesting that Congress give the government's disability programs authority to mandate participation in demonstration projects. With this authority the Administration proposes to conduct an aggressive set of rigorous experiments to improve the labor force participation of people with disabilities.

Developing a Learning Agenda

Agencies are encouraged to adopt a "learning agenda" in which they collaboratively identify the critical questions that, when answered, will help their programs to be more effective, and to plan to answer those questions using the most appropriate tools. An agency learning agenda will:

- Identify the most important questions that need to be answered in order to improve program implementation and performance. These questions should reflect the priorities and needs of Administration and agency leadership, policy and program offices, program partners at state and local levels, researchers

and additional stakeholders, as well as legislative requirements and Congressional interests.

- Strategically prioritize these questions given the level of current understanding, available resources, feasibility, and other considerations to determine which studies or analyses will help the agency make the most informed decisions.
- Identify the most appropriate tools and methods (e.g. evaluations, research, analytics, and/or performance measures) to answer each question.
- Conduct studies, evaluations, and analyses using the most rigorous methods that are feasible and most appropriate.
- Disseminate findings in ways that are accessible and useful to Administration and agency leadership, policy and program offices, state and local partners, practitioners, and other key stakeholders—including integrating results into performance measurement and strategic planning.
- Act on the results by using the information for policy decisions and continuous program improvement.

Implementing a learning agenda approach creates an environment that encourages individuals, offices, and teams to reflect on and learn from their experience and from others. It requires a planned approach to learning in the context of evidence-based decision-making and improving program performance through evaluation and analysis. A learning agenda should be flexible and also reinforce and maximize efforts throughout the life of a program. Once integrated into agency processes, the agenda can help staff and partners learn rapidly to enable iterative course corrections and improvements.

Building an Evidence Infrastructure

Optimal development and use of evidence is made possible by an integrated infrastructure. A strong evidence infrastructure requires a variety of capacities, and developing and supporting the use of evidence and evaluation in decision-making requires coordination between those managing the operations of a program, including administrative data collection and maintenance, and those responsible for using data and evaluation to understand program effectiveness. It requires strong leadership from multiple levels of an agency—policy officials, program administrators, performance managers, strategic planning, policy and budget staff, evaluators, and statistical staff—to ensure that data and evidence are developed, analyzed, understood, and appropriately acted upon. To build the capacity to generate and use evidence, agencies should:

- Ensure that staff with appropriate analytic skills and backgrounds are hired, supported, and effectively deployed.
- Safeguard the ability of Federal principal statistical agencies to objectively design, collect, process, edit, compile, store, analyze, release, and disseminate data.
- Build or support independent evaluation offices to conduct rigorous, independent evaluations.
- Invest in improving administrative data infrastructure, access, and quality, including collecting better quality data from entities receiving federal funding.
- Make better use of existing administrative data to build evidence.
- Utilize new tools and methods such as rapid-cycle iterative evaluation and approaches that utilize behavioral science.
- Expand the building and use of evidence in grant programs.
- Partner with other agencies to share data or jointly design and fund studies.

Centralized or chief evaluation offices play an important role in an evidence infrastructure that can develop and sustain agency capacity to build and use evidence. A recent Government Accountability Office (GAO) report found that Federal agencies with a centralized evaluation authority reported greater evaluation coverage of their performance goals and were more likely to use evaluation results in decision making[1]. Centralized or chief evaluation offices are often essential for ensuring that key evidence and evaluation principles are reflected in practice. The establishment of a centralized evaluation office and an official, public evaluation policy that reflects these principles is a particularly strong and mutually reinforcing combination. A centralized office allows the agency to credibly establish the independence and transparency of its evaluation activity, develop the specialized expertise required to implement rigorous evaluations, and have a centralized entity responsible for coordinating and disseminating research findings.

The Federal evidence infrastructure plays a critical role in supporting State and local efforts to build and use evidence. For example, the Department of Education (ED) has supported a suite of resources that helps States and districts find and develop evidence-based education interventions that work for them, while strongly protecting student privacy. The What Works Clearinghouse's (WWC) Find What Works tool allows educators and policymakers to find education programs and interventions shown to work in a particular context. The Regional Educational Laboratories serve as the primary dissemination partner for the WWC while also helping States and localities build and use evidence to improve student outcomes. Where existing evidence is weak or nonexistent, States and districts can use ED's new "RCT-YES" and Rapid Cycle Evaluation Coach tools to rigorously evaluate innovative, locally tailored educational practices and also use the new CostOut tool to estimate an intervention's costs and cost-effectiveness. ED also provides more inten-

[1] Government Accountability Office Publication No. 15-25, "*Program Evaluation: Some Agencies Reported that Networking, Hiring, and Involving Program Staff Help Build Capacity*," November 2014.

sive support at low cost through Research Collaborations Grants, which funds partnerships between research institutions and State or local education agencies to promote evidence-building on topics that have important implications for student outcomes, and through Low-Cost, Short Duration Evaluations of Education Interventions Grants, which support rigorous evaluations of education interventions that State or local education agencies believe will provide meaningful improvements in student outcomes within a short period of time. Since protecting student privacy is an essential feature of all education research, ED's Privacy Technical Assistance Center provides timely information and updated guidance on privacy, confidentiality, and security practices through a variety of resources, including training materials and opportunities to receive direct assistance with improving the privacy, security, and confidentiality of longitudinal data systems.

Making Better Use of Administrative Data to Build Evidence

Making better use of the administrative data—the data government already collects—is an especially promising strategy for building evidence. Administrative data are data collected by government entities for program administration, transparency, regulatory, or law enforcement purposes. Administrative data, especially when linked across programs or to survey data, can often make both performance measurement and rigorous program evaluations more informative, less costly, and less burdensome to data providers. Federal and state administrative data include rich information on labor market outcomes, health care, criminal justice, housing, and other important topics, but they are often greatly underutilized in evaluating program effects as well as in day-to-day performance measurement and for informing the public about how society and the economy are faring. Given this, a critical part of an evidence infrastructure is helping agencies make better use of administrative data while ensuring individual privacy and data security.

In recent years, Federal agencies have steadily made progress improving the use of administrative data for evidence building. Some agencies are creating capacity to support research and evaluation in a particular policy area, but most Federal agencies could make greater use of administrative data to build evidence or allow those outside government to do so. In addition, many agencies have data that would be useful to other agencies, other levels of government, or outside researchers and citizens to help understand and improve programs. Yet not all agencies have the technological infrastructure, legislative authority, or expertise needed to utilize, share, or link data themselves, nor does it make sense to duplicate these capacities at every agency.

Federal statistical agencies already play a leading role in bringing together data from multiple sources while protecting privacy, confidentiality, and data security. Statistical agencies use data to create a wide variety of statistical products that can be securely accessed by researchers inside and outside of government to conduct a broad array of policy- and program-relevant analyses. High-capacity statistical agencies have partnered with other Federal agencies to link and analyze administrative and survey data for evidence building purposes. For example, the work of the Census Bureau's Center for Administrative Records Research and Applications (CARRA) builds on the Bureau's existing strengths by developing a comprehensive infrastructure to prepare and share administrative data. The Census Bureau's infrastructure links a variety of different data sets, allowing pilot projects to measure outcomes such as mobility, earnings, and employment. Current pilots are measuring labor market outcomes for individuals with former military service and those who obtained manufacturing credentials, and the Census Bureau continues to enhance its secure infrastructure for processing and linking data sets to support evidence-building pilots. Partnerships such as these build on the critical capacities that statistical agencies already have in order to make better use of existing data without creating unnecessary duplication.

Using a Portfolio of Evidence

The credible use of evidence in decision-making requires an understanding of what conclusions can and, equally important, cannot be drawn from the information. Evidence should be rigorous, relevant, transparent, independent, and generated in an ethical manner. Evidence has varying degrees of credibility, and the strongest evidence generally comes from a portfolio of high-quality evidence rather than a single study or data point, i.e., from multiple sources and/or multiple studies covering different aspects and nuances of the topic. Whenever possible, critical decisions should be made based on a body of evidence that has been generated about a particular topic or intervention. One example is the Reemployment Services and Eligibility Assessments (RESEA) program at the Department of Labor. The program was originally created in 2005 and was aimed at reducing improper payments in the Unemployment Insurance (UI) program. Initial research of this program suggested that it was effective at reducing State's UI benefit costs, often in excess of the program's cost. A 2011 random assignment evaluation again showed the program's cost-effectiveness, particularly in Nevada, which was providing more intensive reemployment services and reducing UI benefit costs at a higher rate than the other states studied, more than offsetting the additional program costs. A follow-up evaluation of the Nevada program demonstrated that the intensive reemployment services were helping participants get back to work faster and at higher wages than the control group of UI claimants. As a result of this research, Congress increased appropriations for the program, ultimately approving an expanded national program more closely resembling Nevada's. The FY 2018 Budget proposes to continue this expansion of the RESEA program by proposing mandatory funding to provide these services to the one-half of UI claimants profiled as most likely to exhaust benefits before returning to employment.

Conclusion

There has been meaningful progress in recent years toward building and using evidence for better government, and a bipartisan consensus has emerged regarding the need for further progress. This is especially the case when considering the potential for using existing administrative data for research and evaluation. The bipartisan Commission on Evidence-Based Policymaking is considering how data, research, and evaluation are currently used to build evidence and improve public programs and policies, and how to strengthen evidence-building to inform program and policy design and implementation. The Commission will present its recommendations this Fall, and the Administration looks forward to working with Congress to increase the production and use of evidence throughout the government and for public use. More and better use of evidence would allow us to determine where needs are greatest, and what programs are and are not working and why, in order to develop a more effective and efficient Federal government. Using evidence to improve government is what taxpayers expect—smart and careful use of limited resources to best address national priorities.