

8. PROGRAM EVALUATION

The Administration believes that the Federal Government should use taxpayer dollars efficiently and effectively. Central to that is a culture where agencies constantly ask, and try to answer, questions that help them find, sustain, and spread effective programs and practices; find and fix or eliminate ineffective ones; test promising programs and practices to see if they can be replicated; and find lower-cost ways to achieve a positive impact. The Federal fiscal situation necessitates doing more with less, not only to reduce budget deficits, but to build confidence that Americans are receiving maximum value for their hard-earned tax dollars. It is therefore critical to apply an evidence-based approach to government management that utilizes rigorous methods appropriate to the situation, learns from experience, and is open to experimentation.

One of the challenges to evidence-based policy-making is that it is sometimes hard to say whether a program is working well or not. Historically, evaluations have been an afterthought when programs are designed, and once a program has been in place for a while, building a constituency for rigorous evaluation is hard. The Administration is committed to addressing this problem.

This Administration is strongly encouraging appropriately rigorous evaluations to determine the impact of programs and practices on outcomes, complementing the performance measurement and management practices described in chapter 7, “Delivering High-Performance Government”, in this volume. In many policy debates, stakeholders come to the table with deep disagreements about the effectiveness or ineffectiveness of particular interventions. Evaluations that are sufficiently rigorous, relatively straightforward, and free from political interference are especially valuable in such circumstances.

Evaluations do what performance measurement, alone, cannot. Evaluations determine whether programs produce outcomes superior to alternative policy choices, or not putting into place a policy at all. This is in contrast with performance measurement, which tracks progress toward intended program outcomes, but does not compare outcomes to alternative programs or the status quo. If a particular job training approach has a high job placement rate, is it because it is effective or because it attracts those easiest to place in jobs? An evaluation would compare the employment of participants in the job training program to comparable individuals who did not participate in the program in order to isolate the effects of the training from other factors. Evaluations can answer a wide-range of germane questions such as whether workers are safer in facilities that are inspected more frequently, whether one option for turning around a low-performing school is more effective than another, whether outcomes for

families are substantially improved in neighborhoods that receive intensive services, and whether no-fee debit cards increase savings among the unbanked.

Evaluation is one component of the evidence infrastructure that plays a role in a wide range of decision-making. The best government programs embrace a culture where performance measurement and evaluation are regularly used and complement one another. Agencies use performance measurement to detect practices that hold the most promise for improving performance and those with the greatest problems. Descriptive evidence of program recipients helps managers better target their resources. Regression analyses of administrative data shed light on how to better match recipients with appropriate services. Rigorous evaluations using experimental or quasi-experimental methods identify the effects of programs in situations where doing so is difficult using other methods; and rigorous qualitative evidence complements what can be learned from empirical evidence and provides greater insight into the contexts where programs and practices are implemented more and less successfully.

Continuing its emphasis on rigorous program evaluations initiated in the President’s 2011 Budget, the Administration is proposing new evaluation funding for 2012 for 19 evaluations that have the potential for strong study designs and that address important actionable questions or strengthen agency capacity to support such strong evaluations.

Agencies that submitted proposals were expected to demonstrate that their funding priorities are based upon credible empirical evidence—or that they have a plan to collect that evidence—and to identify impediments to rigorous program evaluation in their statutes or regulations so that these might be addressed going forward.

The evaluation initiative included an extensive review process, with proposals reviewed by program examiners and evaluation experts at the Office of Management and Budget (OMB) and the Council of Economic Advisers (CEA). In some cases agencies then had a series of meetings with OMB and the CEA to sharpen their proposals. Going forward, OMB and the CEA plan to continue to work with these agencies on implementing strong research designs that answer important questions.

While the evaluation proposals include a broad range of domestically and internationally focused agencies, each shares the characteristics of rigor and presenting an actionable choice based on results. The accompanying table presents the evaluation activities proposed for funding as part of the 2012 evaluation initiative. These activities include a series of evaluations assessing the effectiveness of different strategies for improving college enrollment, persistence, and completion, capacity-

Table 8–1. FUNDED PROGRAM EVALUATION INITIATIVE PROPOSALS

| Agency | Description |
|--|---|
| Department of Education | Evaluation of providing high schools with financial aid submissions data |
| Department of Education | Evaluation of integrating FAFSA and tax form preparation |
| Department of Education | Evaluation of college “bridge programs” for adult learners |
| Department of Education | Evaluation of early college placement testing and counseling |
| Department of Education | Evaluation of call centers to increase community college retention |
| Department of Education | Evaluation of Promise Neighborhoods |
| Department of Health and Human Services | Impacts of Medicaid expansions in Affordable Care Act |
| Department of Health and Human Services | Evaluation of health homes for enrollees with chronic conditions |
| Department of Health and Human Services | Falls prevention demonstration and evaluation |
| Department of Health and Human Services | Enhancing quality in early childhood programs |
| Department of Labor | Evaluation of TAA Community College and Career Training Grants |
| Department of Labor | Capacity building |
| Millennium Challenge Corporation | Gender-specific impacts of MCA Benin Access to Land project |
| United States Agency for International Development | Evaluation of Rwanda Integrated Improved Livelihoods program |
| United States Agency for International Development | Evaluation of Haiti Integrated Watershed Management program |
| United States Agency for International Development | Capacity building for evaluation consultancies |
| Department of the Treasury | Research studies to explore new and improved uses of IRS data |
| National Aeronautics and Space Administration | Evaluation of Applied Sciences program |
| Office of Personnel Management | Evaluation of Federal Government telework and Results Only Work Environment pilot |

building for the United States Agency for International Development (USAID) that should help make rigorous evaluation a more routine aspect of their international development assistance efforts, and an analysis of ways to make the Federal workforce more efficient.

The evaluations proposed in this initiative encompass only a fraction of the evaluations performed by the Federal Government. For example, the Recovery Act launched a number of evaluations across the Federal Government on such topics as the effects of different rent formulas on housing assistance recipients, the effects of smart grid meters on residential electricity usage, and the effects of extended unemployment insurance benefit programs on employment outcomes. In addition, the Administration is placing additional focus on agency evaluation budgets to ensure that those dollars are producing high quality evidence that informs key decisions.

New funding for rigorous evaluations is only part of the Administration’s efforts to reinvigorate evaluation activities across the Federal Government. Additional effort is being placed on building agency capacity for doing good evaluations. Whether that is supporting an agency in standing up a central evaluation office, empowering existing evaluation offices, institutionalizing policies that lead to strong evaluations, or hiring evaluation experts into key administrative positions, this Administration strives to build a robust evaluation infrastructure.

In addition, an inter-agency working group is promoting stronger evaluation across the Federal Government by (a) helping build agency evaluation capacity and creating effective evaluation networks that draw on the best expertise inside and outside the Federal Government, (b) sharing best practices from agencies with strong, independent evaluation offices and making research

expertise available to agencies that need assistance in selecting appropriate research designs in different contexts, (c) devising new and creative strategies for using data and evaluation to drive continuous improvement in program policy and practice, and (d) developing Government-wide guidance on program evaluation practices with sufficient flexibility for agencies to adopt practices suited to their specific needs.

OMB is working with agencies to make information readily available online about all Federal evaluations focused on program impacts that are planned or already underway. This effort, analogous to that of the Department of Health and Human Services (HHS) clinical trial registry and results data bank (ClinicalTrials.gov), will promote increased transparency and accountability, and allow experts inside and outside the Government to engage early in the development of program evaluations.

For several new grant-based initiatives, the Administration is using a three-tiered approach to evidence-based funding. First, money is proposed to promote the adoption of programs and practices that strong evidence suggests will lead to significant improvement in results. Second, for programs with some but not as much supportive evidence, additional resources are proposed with the condition that the programs will be rigorously evaluated going forward. Over time, the Administration anticipates that some second-tier programs will move to the first tier, but only if they prove more promising and cost-effective than other programs. Third, agencies are encouraged to innovate and test ideas with strong potential—ideas supported by preliminary research findings or reasonable hypotheses.

A good example of this approach—in which new or expanded programs have evaluation “baked into their

DNA”—is the Department of Education’s Invest in Innovation Fund (i3). The i3 fund invests in high-impact, potentially transformative education interventions—ranging from new ideas with huge potential to those that have proven their effectiveness and are ready to be scaled up. Whether applicants to i3 are eligible for funding to develop, validate, or scale up their program, and therefore how much funding they are eligible to receive, will depend on the strength of the existing evidence of the program’s effectiveness, the magnitude of the impact the evidence demonstrates the program is likely to have, and the program’s readiness for scaling up.

This three-tiered structure will provide objective criteria to inform decisions about programs and practices in which to invest and create the right incentives for the future. Organizations will know that to be considered for significant funding, they must provide credible evaluation results that show promise, and, before that evidence is available, be ready to subject their models to analysis. As more models move into the top tier, it will create pressure on all the top-tier models to compete to improve their effectiveness to continue to receive support. For example, the Administration has chosen to invest in many of those areas, but has made a concerted effort to increase investments in early childhood education and home-visiting programs that are backed by strong evidence—because rigorous evidence suggests that investments in those areas have especially high returns.

Rigorous evaluation will be a central component of several cross-agency initiatives designed to identify more cost-effective approaches to achieving positive outcomes for disadvantaged populations. These populations

are often eligible for multiple services and benefits administered by separate Federal and State agencies, which are poorly coordinated and governed by rules that stifle effective collaboration and innovation. In 2012, the Departments of Labor and Education will support joint pilots to test interventions and systemic reforms with the potential to improve education and employment outcomes at lower cost to taxpayers. The Social Security Administration and the Department of Education will launch a joint initiative to test interventions that improve outcomes for children with disabilities and their families, which may yield substantial long-term savings if these children leave the Supplemental Security Income program. OMB’s Partnership Fund for Program Integrity Innovation will test promising solutions developed collaboratively by Federal and State agencies to improve payment accuracy, improve administrative efficiency, and enhance service delivery in overlapping benefit programs. Rigorous evaluation of these cross-agency pilots will help determine which strategies lead to better results at lower cost, allowing Federal and State governments to identify the most promising strategies that warrant expansion.

The President has made it very clear that policy decisions should be driven by evidence—evidence about what works and what does not and evidence that identifies the greatest needs and challenges. By instilling a culture of learning into Federal programs, the Administration will build knowledge so that spending decisions are based not only on good intentions, but also on strong evidence that yield the highest social returns on carefully targeted investments.

