

Using Evidence as the Driver of Policy Change: The Next Steps in Supporting Innovation, Continuous Improvement, and Accountability

Testimony of Gordon L. Berlin, President of MDRC, Before the Senate Finance Committee

May 10, 2016

Good morning, Chairman Hatch, Ranking Member Wyden, and members of the Committee. I appreciate the opportunity to participate in this hearing on the role of evidence-building in the future design and evolution of the nation's education and social welfare policies and programs.

My name is Gordon Berlin, and I am President of MDRC, a nonprofit, nonpartisan education and social policy research organization that is dedicated to learning what works to improve policies and programs that affect the poor. Founded in 1974, MDRC evaluates existing programs and develops new solutions to some of the nation's most pressing social problems, using rigorous random assignment research designs or near equivalents to assess their impact.

The federal government spends billions of dollars on policies and programs designed to improve the human condition; ameliorate poverty; increase employment, earnings, and income; invest in education to build human capital; and ensure America's competitive position in a technologically advancing world. But to make a real difference, to ensure a return on investment for both taxpayers and the beneficiaries of these programs, we have to do things that actually work.

Over the last decade and a half, during a period defined in the public consciousness by political partisanship, the legislative and executive branches have quietly forged a bipartisan consensus around the need to build evidence of effectiveness that would ensure high rates of return on investment for the nation's social programs. The establishment by Congress of the new Evidence-Based Policymaking Commission is only the most recent example of this consensus.

In my testimony today, I offer a brief history of the federal government's fruitful investment in evidence-building; then concentrate on several of the most recent efforts, including supporting pay-for-success initiatives; and conclude by identifying obstacles to effective evidence-building, offering potential solutions and suggesting a framework for exploiting the opportunities that lie ahead.

A Brief History of Evidence-Building

The history of evidence-building has evolved through four phases: in the 1980s and 1990s, federal waivers made the states the engine of innovation; in the early 2000s, the Office of Management and Budget drove the federal agencies to focus on developing rigorous evidence, and the Congress created the Institute of Education Sciences; in the late 2000s, the federal government built partnerships with the nonprofit and philanthropic communities and defined tiered standards of evidence; and today Congress has taken the lead, building on the earlier

phases to create new evidence-building frameworks and pay-for-success initiatives that seek to support innovation and stronger results.

Putting States in the Lead: The Waiver Era. The earliest roots of the evidence evolution can be traced to the 1980s and 1990s when the nation turned to the states as laboratories of innovation. Using the Social Security Act’s 1115 waiver authority, the federal government allowed states to experiment with new approaches to welfare policy. In return, states were expected to subject their programs to rigorous evaluation, nearly always using random assignment, to demonstrate that the reforms had the intended effects and were cost neutral. The federal Administration for Children and Families and its Office of Planning, Research and Evaluation (OPRE) played a key role in managing the waiver process and organizing that research — as well as in setting the agenda in subsequent years for advancing our understanding of the contributions job search, training, and community work experience could make. The remarkable body of evidence built during this period provided the backbone for passage of the 1988 Family Support Act and the 1996 welfare reform law. No one understands the results of that work better than the members of this Committee: together with a strong economy, the resulting welfare-to-work programs and related supports embodied in federal law led to an unprecedented increase in employment rates among welfare recipients and a sharp decline in the nation’s welfare rolls. I turn later in this testimony to some of the challenges associated with the waiver authority in health and welfare policy broadly and the current state of the Temporary Assistance for Needy Families program specifically.

Creating Capacity: The Era of Investment in Federal Research Centers. With OPRE’s rich body of research evidence on welfare as one example, the Office of Management and Budget in the Bush Administration began encouraging all federal agencies to use randomized controlled trials to evaluate the effectiveness of the programs they funded. In 2002, Congress passed the Education Sciences Reform Act, which created the Institute of Education Sciences (IES), the independent statistics, research, and evaluation arm of the U.S. Department of Education. IES has been universally recognized for raising the standards of evidence in education research. In 2008, recognizing that states had little incentive to allocate dollars to programs with the strongest evidence, President Bush and Congress appropriated \$10 million to support states that committed to scale up evidence-based home visiting programs. This bipartisan action laid the groundwork for the next era in evidence-based policy.

Partnering with the Nonprofit Sector: The Era of Tiered Evidence. During the third phase in the late 2000s, Congress capitalized on innovation and creativity in the nonprofit sector in legislation that created the Investing in Innovation Fund (\$650 million) at the U.S. Department of Education, the Workforce Innovation Fund (\$125 million) at the U.S. Department of Labor, the Teen Pregnancy Prevention Program (\$110 million) at the U.S. Department of Health and Human Services (DHHS), and the Social Innovation Fund (\$50 million) at the Corporation for National and Community Service. Each of these funding streams employed “tiered” evidence structures that allotted the largest dollar amounts to the programs with the strongest evidence, each encouraged new program development, each emphasized taking what works to scale, and, importantly, each required that additional evidence be generated as programs expanded. Programs funded by these streams include such effective interventions as the Success for All school model, the Reading Partners volunteer tutoring program, the Youth Villages YVLifeSet

program for young people transitioning out of foster care, and the transitional employment program for reentering prisoners run by the Center for Employment Opportunities.

Perhaps the largest evidence-based program of the era is the Maternal, Infant, and Early Childhood Home Visiting (MIECHV) program, which provided \$1.5 billion in funding for home visiting programs over five years. In many respects, MIECHV, which was largely crafted by this Committee, was a model of building and using evidence to shape policy. The authorizing legislation reflected the considerable amount of research that had already been conducted on home visiting in two ways. First, it required DHHS to establish transparent criteria to determine which models of home visiting had prior evidence of effectiveness. It then required states to spend the majority of their MIECHV funds on those models.

Recognizing that prior studies had also shown gaps in where home visiting was effective, the legislation also encouraged states to innovate by allowing part of the funds to be used for promising approaches to home visiting. For example, home visiting programs have historically not had large effects on maternal mental health, so states could use these funds to pay for new services to try to ameliorate maternal depression. At the same time, the authorizing legislation recognized a need for continued learning about the effects of home visiting. To that end, it specified that 3 percent of the budget should be spent on technical assistance and research activities. Finally, the legislation also required a national evaluation of the MIECHV program, noting that random assignment designs were to be used “to the maximum extent feasible.”

In sum, there are several elements of MIECHV worth emphasizing:

- Prior evidence was used to influence how federal funds could be spent, making it more likely that the funds would make a difference for families.
- The legislation recognized that there were areas where home visiting was not as effective as desired and offered states funds to test innovative approaches.
- Funds were set aside for research to make sure that learning continued under MIECHV and could influence future realizations of home visiting.

Together these initiatives laid the groundwork for the modern era of evidence-building that Congress has now embarked upon.

Today: Using Evidence as a Driver of Change. Building on past efforts, this Congress is actively using evidence as an instrument of change — setting aside a small fraction of funding for evidence-building, charging the executive branch with using tiered strategies that privilege those programs and models with the best evidence with the highest levels of funding, and requiring that states look to evidence-based models as they expand federally funded programs. In addition, recognizing that context matters, these set-asides include funding for innovation in the private and nonprofit sectors and also in states and localities.

Indeed, social impact bond financing — a strategy that brings together private-sector capital, nonprofit know-how, and government urgency in a partnership designed to take promising new approaches to scale — is an important part of the Workforce Innovation and Opportunity Act of 2014 (WIOA) and the Every Student Succeeds Act of 2015 (ESSA).

WIOA places new emphasis on evidence-based decision-making by asking states and the U.S. Department of Labor (DOL) to identify, evaluate, and share promising and proven practices. In addition, DOL is charged with establishing a robust program of research to fill crucial knowledge gaps, including learning what works to help disconnected youth return to school and enroll in postsecondary training and conducting regular and rigorous evaluation of programs for dislocated workers and of the Job Corps. States are required to evaluate their programs annually using rigorous methods, and the law sets aside dollars for governors to encourage experimentation at the state and local level. Surprisingly, the law doesn't employ the tiered evidence strategies pioneered in other bills. The law also created new authority for local workforce investment boards to use up to 10 percent of funds across all three formula grant programs — Adult, Youth, and Dislocated Workers — to support pay-for-performance or social impact bond projects. As we begin to learn more about what works, emphasis should turn to ensuring that this evidence is used in the design and delivery of local workforce programs:

- A tiered evidence strategy could be employed to signal the value of developing new approaches, testing promising programs at some scale, and expanding programs with the strongest evidence.
- State and local workforce agencies should work with the Department of Labor to build the data systems necessary to support both better performance management and evidence-building.
- As a complement to the mandatory sanctions that current law imposes when states fail to meet performance goals, create funding streams that reward performance improvement and the use of evidence-based programming.
- To encourage innovation and evidence-building, governors should use a portion of WIOA discretionary dollars to co-fund with DOL a series of workforce innovation laboratories to experiment with new approaches.

The Every Student Succeeds Act (ESSA) allows states to use some of their Title I, Part D, and Title IV, Part A, funds to support pay-for-performance initiatives as long as they also conduct a rigorous third-party evaluation to determine if the project actually worked. ESSA also provides a clear definition of evidence modeled after the Investing in Innovation (i3) tiers of evidence. It requires low-performing schools to use evidence-based interventions and sets aside 7 percent of Title I funds so that states can pay for them. ESSA also clearly values the development and evaluation of innovations in education by codifying a tiered approach in the Education Innovation Research Grants Program, which replaces the i3 program, and by allowing the Institute of Education Sciences to pool money set aside for evaluations of individual programs. But several key features in the new law could also be improved:

- School districts should be offered incentives to choose the interventions that have the strongest evidence.
- The law's definition of evidence includes correlational studies, which are weaker than experiments, and could end up supporting unsubstantiated interventions.
- Substantial discretion is left to the states in gauging evidence quality but without a strong framework or support; therefore, application of evidence is likely to be uneven and too easily overridden by political and other considerations.

- A clearer connection should be established between the evidence built by IES and summarized in its What Works Clearinghouse and what constitutes reliable evidence at the state level, where officials are charged with deciding what “evidence-based” will mean.

In addition, the separate Social Impact Partnership Act (S. 1089) would authorize \$300 million to support states and localities in the use of social impact bonds (SIBs) to tackle social and public health challenges — of which \$10 million is allocated for feasibility studies, \$45 million for evaluations, and the remaining funds to support deals and pay for administrative costs. The purpose is to spur state and local experimentation, build better evidence about what works, and foster the bundling of philanthropic, private investment, and public funds to scale effective social programs operated by nonprofit organizations and local governments. To promote wider use of this new vehicle, federal funds can be used to support project costs — essentially allowing the Treasury Department to be a financing party in SIB deals, providing dollars to state and local entities so that they may in turn pay back investors if an initiative is found to be effective by a rigorous third-party evaluation. While the bill admirably specifies a comprehensive list of acceptable project outcomes, that list does not include pre-k or college outcomes, two promising areas for pay-for-success investing.

Lessons from Experience on Social Impact Bonds

Social impact bonds — or pay-for-success initiatives — are attractive across the political spectrum as they bring together the concepts of lending for a social purpose, return on investment, and payment based on performance. In the simplest terms, a SIB uses private funds to pay for a social, educational, or health program, and the government repays investors (plus a return) only if the program achieves prespecified results. A SIB can be a valuable new tool in the evidence-building movement. A financing mechanism for testing and scaling innovative new approaches to complex social problems, it provides important flexibility to state and local authorities, who might not otherwise have the resources to experiment with new approaches. This is a valuable complement to the broader agency-based federal role that should be retained as the centerpiece of future evidence-based policymaking.

At MDRC, we have first-hand experience with social impact bonds, having served as the intermediary at the center of the first SIB initiative in the United States, which financed the Adolescent Behavioral Learning Experience (ABLE) program at Rikers Island in New York City. It was designed to learn whether a cognitive behavioral therapy program delivered to all 16- to 18-year-old adolescents awaiting trial at Rikers Island jail would equip them with the social and decision-making skills to help take responsibility for their actions, avoid rash and impulsive reactions, make more constructive life choices, and ultimately avoid a costly return to Rikers.

The Rikers SIB sought to answer two distinct questions, each of which should be judged independently. The first was: did the program reduce recidivism? The short answer is no; ABLE was discontinued after three years when results from an independent evaluation by the Vera Institute of Justice demonstrated that it was not meeting its goal of reducing recidivism among 16- to 18-year-olds. The second question was: Is a social impact bond a viable mechanism for

financing, testing, and scaling a social program innovation? Here the answer is clearly yes. Government got to try something it would not otherwise have attempted, while investors assumed all of the risk.

But, as Rikers proved, SIB funding does not guarantee programmatic success. A SIB-supported program that turns out not to work can result in losses for investors, no cost savings for the government, no opportunity for long-term growth for nonprofit service deliverers, and limited benefits for participants. Yet while the intervention may not have worked, the SIB structure did succeed: the government did not pay for ineffective services, but learning what didn't work may have helped point the way to what might.

As early adopters like us gain real-world operating experience, the reality of SIBs is turning out to be more nuanced than either proponents or detractors have promised. As I argue in [*Learning from Experience: A Guide to Social Impact Investing*](#), emerging lessons from Rikers and other deals reveal both SIBs' value to government entities and also the reality that this value will only be realized if the tensions inherent in structuring the terms of a SIB deal can be addressed squarely. These include:

- **The balance of risk and reward.** The returns governments are able to pay may not be proportional to the risks some lenders are able to take.
- **The focus on narrowly defined government savings.** Many deals still depend on the possibility of government savings. But insisting on government savings can unnecessarily rule out projects that might otherwise offer valuable and readily quantifiable returns to participants and society at large.
- **The role of evidence.** SIBs reduce risk for government entities by promising that they will have to pay *only* for successful interventions. To fulfill that promise, a SIB must include independent, rigorous evaluation of a program's effectiveness over the status quo — a requirement that poses a new form of investor risk. Unfortunately, efforts to mitigate that risk have driven some SIB deals to rely instead on simple outcome measures, which may misleadingly provide only the illusion of benefits and savings to government entities.

What does that mean for Congress moving forward on pay-for-success? The SIB vehicle is a valuable tool for stimulating experimentation, innovation, and the scaling of promising approaches at the state and local level. But it is not a substitute for a comprehensive agency-directed research agenda, which is essential if evidence is to be harnessed effectively to improve performance. The federal government can enhance the value of SIBs as a “learning what works” tool by:

- Stating clearly that the goal is to build evidence about what works.
- Signaling its support for a benefit-cost perspective that places equal weight on quantifiable government budget savings, benefits for participants, and benefits for society as a whole.
- Supporting the evaluation costs of SIB projects only when they use rigorous research designs capable of reliably attributing causality, requiring those studies to collect a range of relevant outcome measures (not just the single measure that payback would be based upon), and insisting that those measures be quantifiable measures of success.

- Stimulating deals by supporting state and local pay-for-success projects in the form of backend payments as specified in the Social Impact Partnership Act, WIOA, and the Social Innovation Fund Classic competitions, but giving broad latitude to dealmakers (public agencies, investors, and nonprofit service providers) by not being overly prescriptive about the requirements governing a SIB deal — in essence, allowing the private market aspects of SIBs to work, while still insisting on rigorous evidence and payment terms that are fair to both taxpayers and investors.
- Sustaining what works by requiring, to the extent possible, that state and local agencies continue supporting effective programs after a SIB deal ends.
- Insuring transparency by requiring that public agencies using federal funds must release publicly the final results of evaluations, the payment terms of deals, and final benefit/cost calculations, including the lost savings that would result if the government did not continue to support effective initiatives.

Looking Forward: The Future of Evidence-Building

The federal investment in evidence over the past two decades has yielded great benefits in our understanding of what works in education, workforce development, health and human services, and housing policy. As this evidence base continues to grow, however, we must address a number of challenges to make the most of what we are learning by: (1) creating a culture of continuous improvement with incentives for using research evidence to make programs even more effective (2) balancing the role of agency-directed research with pay-for-success innovation in the states and localities, (3) making better use of federal waiver authority, (4) improving access to data for more efficient evaluations, while protecting confidentiality, and (5) clarifying federal authority to conduct research.

Continuous Improvement Instead of Up-or-Down Judgments. Historically, we have failed to build a reliable record of what has worked and what has not, dooming ourselves to constantly repeating past mistakes and relearning the same lessons over and over again. With its actions in the last decade or so, Congress has taken decisive steps to build that record and to drive policy and funding with evidence.

But to be successful, consensus will also be needed around means and ends. Is the goal simply to fund what works and end what does not? Or is it to invest in continuously improving the nation's social programs? Those who are skeptical of the efficacy of social programs question whether we will really act on evidence when programs do not work. On the other side, supporters of social programs worry that evidence will be used as a blunt instrument to justify diminishing the nation's investment in ameliorating social problems. Both sides have a point.

To take just one example: in the early 1990s, the national evaluation of the Job Training Partnership Act's youth employment programs found few effects for youth, and Congress moved decisively to defund all of the Act's youth training and employment programs. The discontinuation of federal summer youth employment funding closed down one of the nation's largest employers of young people; funding for training programs was also terminated, and even the Job Corps experienced reductions. But no research and development investments were made in their stead to learn what might work better. Twenty years later, the nation is in the midst of the

worst youth unemployment crisis in recent history. Fortunately, Congress has built a learning agenda about youth employment in the recently passed WIOA, but many years and lives were wasted in the interim.

Eventually, we do have to say no when efforts to enhance a given approach repeatedly fail to produce evidence of effectiveness. But that doesn't mean doing nothing when the problems remain pressing. In short, the evidence-building agenda should be framed around long-term goals, and the tactics should focus on the means to that end. Like a business committed to becoming a dynamic learning organization, government should commit to continuous improvement and focus on making its programs more effective over time. The Edna McConnell Clark Foundation, one of the first (and largest) Social Innovation Fund intermediaries, offers an excellent example of building evidence and using it to continuously improve the programmatic and organizational capacity of its youth-serving grantees.

Bolstering Agency Research Capacity and Supporting Pay-for-Success Innovation. In recent legislation, support for evidence-building, especially as it relates to innovation and scaling, has taken two forms: (1) traditional research and demonstration funding that charges the research arms of federal agencies with building bodies of evidence about what works and (2) social impact bonds or pay-for-success structures that attract private funding for innovation.

While both vehicles will build evidence, they approach the task differently, with each having advantages and disadvantages. Simply put, the agency research model is preferable when the goal is to learn what works for a broad and persistent problem like the transition from foster care, where one wants to develop a reliable body of evidence, one study building on the last. Agencies should be charged with creating a portfolio of research about a problem's underlying causes and to test a range of possible solutions, always answering three questions: what difference did the program make, how did it do so, and why? To support that work, Congress could authorize federal agencies to set aside at least 1 percent of existing program funds for evaluation, a solution that is budget-neutral. And because context matters, agencies should also be encouraged to pay attention to the systems in which programs operate — something federal agencies too seldom do. It is not enough to learn what works; introducing what works into broader systems and maintaining quality at scale is the next frontier in evidence-based policymaking.

The SIB model is best used when the goal is to encourage local experimentation and innovation and as a means of ensuring local buy-in. Its principal disadvantages are the narrow focus on a single outcome measure to determine success and the difficulty of packaging a deal that adequately balances risk against reward and that satisfies both investors and government payers.

The Role of Waivers. Unlike the story of the welfare waivers in the 1980s and 1990s, the use of waiver authority in health and health care offers a cautionary tale. As documented by the Government Accountability Office in 1995, 2002, and 2008, states have used 1115 waiver authority to experiment with new service delivery models, eligibility rules, and other strategies with the goal of improving outcomes and patient experiences and lowering costs in the nation's Medicaid program. While the waiver process unleashed an exceptional amount of state experimentation, little evidence was generated because there was no requirement that rigorous evaluations be conducted of the innovations. The Center for Medicare and Medicaid Innovation

created as part of the Affordable Care Act is now hard at work attempting to more rigorously build that evidence base.

Ironically, we now face the very same problem in welfare policy. The waiver structure that spurred experimentation and nourished evidence-building no longer exists, even though we face a next generation of problems — incomes have stagnated, child poverty rates remain alarmingly high, and the fraction of families with neither earnings nor government-provided benefits has increased. We need to refine our understanding about how to promote job advancement and retention; how to understand the role of public employment during severe economic downturns and in urban, rural, and Tribal areas where unemployment is perennially high; how to serve those with severe, persistent, and multiple employment barriers; and how best to engage with noncustodial fathers who owe child support and are unemployed. Initially, the block grant structure of TANF and the surpluses states enjoyed as a welcome by-product of the remarkable economic expansion of the late 1990s fueled a new round of state-led innovation. But the economic slowdown and states' reluctance to commit state funding to new endeavors have limited the amount of experimentation in a number of critical areas. An annual set-aside of program dollars that states could apply for on a matching basis to pay for pilot tests of new ideas — tests that include requirements for rigorous, independent evaluations — is needed.

Improving Access to Federal Data for More Efficient Evaluations. In evaluating the effectiveness of social programs, researchers need ready access to administrative data. Research firms that are funded by federal agencies to evaluate programs often rely on data collected by states from employers on employment and earnings, for example, data that the states already report to the federal government for certain child support enforcement and other purposes. These data are housed in accessible form at the federal level within the National Directory of New Hires (NDNH) database. However, research contractors face severe restrictions in accessing this essential database for assessing whether federally supported programs actually work. Instead, they are forced to get the *very same data* directly from the states, at great cost to the federal government and at considerable burden in duplicative reporting for the states. If the NDNH database were made more widely available to evaluators (with appropriate privacy safeguards), it would enable Congress and the federal agencies to assess the impact that social programs have on jobs and earnings at much less cost and burden to the federal government and the states.

Similar opportunities exist for building robust data systems from the wealth of data about individuals' experiences with the health care, public assistance, criminal justice, child welfare, school, and college systems. Integrated data systems would save time and money and reduce paperwork burdens in the conduct of evaluations, while providing comprehensive information about program effectiveness over time. The American Opportunity Study being overseen by a committee of the National Academies of Sciences, Engineering, and Medicine and housed at the U.S. Census Bureau in a secure data environment would take an important step in that direction. Private industry is far ahead of the public sector — exploiting big data to understand customer desires, track trends, and assess performance, so that low-performing business units can learn from high-performing ones. The public sector has the data to build comparable integrated systems to track progress longitudinally and to enhance program performance, yet federal and states agencies (and their contractors) cannot regularly access and share data for evaluation purposes. We can do better. Agencies and their contractors need clear authority to access and

link administrative data for evaluation purposes when it is housed at the federal level and similarly clear authority when it is housed at the state level. Government efficiency hangs in the balance.

Protecting Confidentiality. At all levels and branches of government, there is a tug of war between those who are focused on improving program effectiveness and those who are concerned with protecting privacy. Staff responsible for managing data are rightly charged with keeping it secure and protecting privacy but too seldom with developing protocols for sharing it securely with other agencies and evaluators. Although the stakes are high and the opportunity significant, the program office that houses the data often have little or no interaction with the same agency's evaluation office. If these two objectives — measuring program effectiveness and safeguarding privacy — remain mutually exclusive, continued paralysis is the inevitable result.

The Family Education Rights and Privacy Act (FERPA) illustrates the challenge. Congress is considering amending FERPA because of concerns over threats to the privacy of student data, and meanwhile state legislatures have stepped in. Just last year 47 state legislatures introduced over 180 bills to address student data protection issues, a reaction originally prompted by public outcry over educational technology vendors and their use of children's information for advertising and commercial gain. Unfortunately, education researchers from academia and other nonprofit institutions have gotten swept up in the furor. Under current federal law, education agencies can share data with researchers only for research projects to benefit students and improve instruction — and only under extremely strict privacy conditions. But some are suggesting that Congress should significantly scale back even that authority. Without access to student data, little education research could be conducted at all. The bottom line is that it's essential to continue to protect the security and privacy of student data, but we must be careful to not unintentionally end the analysis of student data for its original purpose: assessing and improving education.

To strike the right balance between protecting privacy and confidentiality while still allowing access for research, Congress could start by specifying required levels of encryption and protection using the highest standards established by the National Institute of Standards and Technology (NIST). The NIST standards are appropriate for research data that must be kept confidential to protect the privacy and well-being of study participants.

Clarifying Federal Authority to Conduct Research. Inconsistencies in federal authority to conduct independent research and evaluation as well as procurement and contracting rules pose additional hurdles for efficient evidence-building. To guide policy, research must be independent, objective, and reliable. However, the authorizing legislation establishing agency research departments does not always set forth these requirements — for example, the law governing the Institute of Education Sciences does while the one over the Office of Planning, Research and Evaluation at DHHS does not. When authority is clear, agencies and their contractors have less difficulty accessing data, recruiting sites, establishing data-sharing agreements, and getting local buy-in. For example, in our experience studying home visiting programs, we found states and localities willing and ready partners in a random assignment research design when legislation instructed the federal agency to make program funding

contingent upon participation in the evaluation. In a complementary home visiting study that was not explicitly described by Congress, however, site recruitment proved difficult.

Procurement and process obstacles to cost-effective evaluations should also be addressed. While its goals are laudable, the Paperwork Reduction Act's requirements for clearance by the Office of Management and Budget and for filing two public notices for every survey involving more than 10 people add time and money to fielding studies. Rules designed to improve government procurement of goods and services can create havoc in evaluation research — for instance, an 18-month follow-up survey may span fiscal years and the pace at which the survey will be fielded is difficult to predict, but the funding of that survey can't cross fiscal years. Research, demonstration, and evaluation projects are not readily severable. Given the need for long follow-up periods, studies may spread over five or more years. Under current procurement rules, contract requirements and dollar amounts for each year must be specified separately. This has dramatically increased costs to administer and manage research.

Conclusion

The bipartisan commitment to building evidence by Congress has created a new generation of innovation in education and social policy — better programs, more effective policies, more responsible use of taxpayer dollars. But there is more to be done. It is not enough to identify isolated examples of what works. We need to build a culture of continuous improvement and accountability, always striving to learn more and to more effectively direct funding at all levels of government to those programs with the strongest evidence.

Thank you for the opportunity to testify today.