

Statement of Jon Baron
President, Coalition for Evidence-Based Policy

Chairman Reichert, Ranking Member Doggett, and Members of the Ways and Means Subcommittee on Human Resources:

I appreciate the opportunity to testify on the evaluation of efforts to help families support their children and escape poverty. As brief background, the Coalition for Evidence-Based Policy is a nonprofit, nonpartisan organization, established in 2001. We work with federal officials to increase the effectiveness of government social spending through rigorous evidence about “what works,” and the core ideas we have advanced have helped shape evidence-based reforms enacted into law and policy during both the Bush and Obama Administrations. We are not affiliated with any programs or program models, and have no financial interest in any of the policy ideas we support, so we serve as a neutral, independent resource to policy officials on evidence-based programs. Our work is funded primarily by national philanthropic foundations, including the MacArthur Foundation and the William T. Grant Foundation.

My testimony will first discuss how evidence-based program reforms can greatly increase the effectiveness of government social spending in improving people’s lives. Then I will offer some suggested next steps that the Subcommittee might consider to advance such reforms in the programs within its jurisdiction.

I. Rigorous randomized studies have the ability to identify program reforms (“interventions”) that increase the effectiveness of social spending while actually reducing its cost.

It is often assumed that the only way to increase government’s impact on social problems such as poverty and educational failure is to spend more money – an assumption that conflicts with the current national interest in reducing the deficit. Largely overlooked, however, are clear examples from welfare and other areas where rigorous randomized trials – widely considered the strongest method for evaluating program impact – have identified program reforms that produced important improvements in people’s lives, while simultaneously *reducing* government spending. These examples suggest that a systematic government effort to build a body of such proven reforms, and disseminate them widely in federal social programs, could improve life outcomes for millions of Americans without adding to – or while even reducing – the deficit.

Examples from programs within the subcommittee’s jurisdiction:

- **Welfare-to-work strategies shown to produce sizable increases in participants’ employment and earnings, and reductions in their use of government assistance.** In the 1980s and 1990s, government, foundations, and leading researchers sponsored or carried out a large number of randomized controlled trials of state and local welfare reforms. Three major reform efforts – two in California, one in Oregon – were found especially effective. Focused on moving welfare recipients quickly into the workforce through short-term job-search assistance and training (as opposed to longer-term remedial education), the initiatives produced gains in participants’ employment and earnings of 20-50%. Remarkably, they also produced net savings to the government, in reduced welfare and food stamps, of \$1,700 to \$6,000 per person.¹

These findings helped build political consensus for the strong work requirements in the 1996 welfare reform act, and shape many of the work-first state-level reforms that followed. The scientific rigor of the findings were critical to their policy impact.²

- **Subsidized Guardianship – an innovation in the foster care system shown to increase children’s placement in a permanent home while reducing foster care spending.** In the late 1990s, the Department of Health and Human Services (HHS) granted Illinois a waiver from federal law to implement Subsidized Guardianship – an alternative to foster care in which the state pays a subsidy to a child’s relative or foster parent to become the child’s legal guardian. The approach is similar to adoption but, unlike adoption, does not require the termination of parental rights and so can be easier to achieve from a legal standpoint.

As a condition of the waiver, Illinois evaluated Subsidized Guardianship in a large randomized trial.³ Over a nine-year period, the study found that the program (i) increased children’s placement in a permanent home by 8%; (ii) reduced average days in foster care by 16%; and (iii) produced net government savings of \$2,300 per child (because subsidizing a guardian is administratively less costly than foster care). Based on these findings and successful replication trials in other states, CBO scored savings of \$800 million for federal legislation, enacted in 2008, to expand Subsidized Guardianship nationally.⁴

- **Reemployment and Eligibility Assessments – an innovation in the Unemployment Insurance (UI) system shown to produce UI savings while increasing workforce earnings.** In 2009, the Department of Labor launched a four-state randomized trial of the Reemployment and Eligibility Assessment (REA) program for UI claimants.⁵ The program includes a mandatory in-person review of the claimant’s eligibility for UI, and personalized job-search and other reemployment assistance. Over a 12-18 month period, the study found: (i) \$180 in net government savings per claimant from reduced UI payments; (ii) especially large savings in Nevada – \$604 per claimant – possibly due to distinctive features of Nevada’s REA program that could be replicated elsewhere; and (iii) an increase in job earnings of \$2,600 (18%) per claimant in Nevada – the one site that obtained a reliable estimate of the effect on earnings. (The study also found a smaller – 5% – increase in earnings in Florida over a 12-month period, but the study’s analysis suggests this finding may not be reliable.⁶)

These results suggest that nationwide implementation of REA for all eligible UI claimants could produce \$1.5 billion in net government savings per year,⁷ while increasing workers’ earnings. If the larger Nevada effects could be reproduced nationally, the savings might be as high as \$5 billion per year,⁸ and the increase in workers’ earnings could be substantial.

- **Nurse-Family Partnership – a home visiting program for low-income, first-time mothers shown to produce major improvement in participants’ lives, while reducing their use of public assistance.** The Nurse-Family Partnership (NFP) is one of the main program models funded by HHS’s Maternal, Infant, and Early Childhood Home Visiting program. NFP has been shown in three well-conducted randomized trials to produce major improvements in participants’ life outcomes, such as: (i) 20-50% reductions in child abuse/neglect and injuries; (ii) 10-20% reductions in mothers’ subsequent births during their late teens and early twenties; and (iii) sizable improvements in cognitive and educational outcomes for children of the most at-risk mothers. In addition to these benefits, newly-published reports from the ongoing trial in Memphis, Tennessee show, 12 years after the women gave birth, a \$1,113 reduction in annual government spending per woman on welfare, food stamps, and Medicaid during the 12 years. As a result, the total discounted government savings over the 12 years (\$13,350) more than offset the program’s cost (\$12,493).⁹

II. To identify enough of these reforms to generate broad-based gains in government effectiveness requires strategic trial-and-error – i.e., rigorously testing many promising interventions to identify the few that are effective.

Rigorous evaluations, by measuring programs' true effect on objectively important outcomes such as workforce earnings, college attendance, teen pregnancy, and child maltreatment, are able to distinguish those that produce sizable effects from those that do not. Such studies have identified a few interventions that are truly effective – such as those described above – but these are exceptions that have emerged from testing a much larger pool. Most, including those thought promising based on initial studies, are found to produce few or no effects – underscoring the need to test many. This is true not only in social spending, but in other fields where rigorous evaluations have been carried out. For example:

- **Education:** Of the 90 interventions evaluated in randomized trials commissioned by the Institute of Education Sciences (IES) since 2002, approximately 90% were found to have weak or no positive effects.¹⁰
- **Employment/training:** Of the 13 interventions evaluated in Department of Labor randomized trials that have reported results since 1992, about 75% were found to have found weak or no positive effects.¹¹
- **Medicine:** Reviews have found that 50-80% of positive results in initial (“phase II”) clinical studies are overturned in subsequent, more definitive randomized trials (“phase III”).¹²
- **Business:** Of 13,000 randomized trials of new products/strategies conducted by Google and Microsoft, 80-90% have reportedly found no significant effects.¹³

III. The current pace of rigorous testing is far too slow to build a meaningful number of proven-effective interventions to address our major social problems. Of the vast array of ongoing and newly-initiated program activities in federal, state, and local social spending, only a small fraction are ever evaluated in a credible way to see if they work. For example, based on our careful monitoring of the literature, the federal government commissions randomized evaluations of only 1-2 dozen such program activities each year.

IV. The end goal – a sizable body of proven social interventions – is of critical importance. In recent decades, the U.S. has failed to make significant progress in key areas such as –

- **Poverty:** The U.S. poverty rate – now at 15% – reached its low in 1973. It has shown little change (whether by official or alternative National Academy measures) since the 1970s.¹⁴
- **K-12 education:** Reading and math achievement of 17-year-olds – the end product of our K-12 education system – is virtually unchanged over the past 40 years, according to official measures,¹⁵ despite a 90% increase in public spending per student (adjusted for inflation).¹⁶
- **Well-being of low to moderate income Americans:** The average yearly income of the bottom 40% of U.S. households, now at \$20,221, has changed little since the early 1970s.¹⁷

V. We therefore urge the Subcommittee to lead a bipartisan reinvention of U.S. social spending based on evidence about “what works,” through steps such as the following:

A. Authorize and encourage the agencies to make maximum use of waivers from federal law and regulation to build credible evidence.

- 1. “Waiver-evaluations” were deployed with great success in 1980s/90s welfare reform, making a critical contribution to the body of welfare-to-work evidence discussed above.** Specifically, in the years leading up to the 1996 welfare reform act – through both Republican and Democratic Administrations – OMB and HHS had in place a “demonstration waiver” policy, under which HHS waived provisions of federal law and regulation to allow states to test new welfare reform strategies, but only if the states agreed to evaluate their reforms in rigorous (usually randomized) studies.

This policy directly resulted in more than 20 large-scale randomized controlled trials that tested an important and diverse set of reforms, and thereby helped build the influential body of welfare-to-work evidence discussed above. These reforms that were tested include, for example, mandatory job search and employment activities (e.g., Vermont); employment subsidies for welfare recipients who left welfare for full-time work (e.g., New York, Minnesota); time limits on welfare (e.g., Florida, Connecticut); “family cap” policies designed to discourage additional births among women on welfare (e.g., Arkansas, New Jersey); and various combinations of the above reforms.

- 2. We encourage the Subcommittee to advance a similar waiver-evaluation concept approach across the broad range of programs within its jurisdiction, designed to:**
 - a. Stimulate a robust array of state/local program innovations, aimed at (i) producing budget savings while improving program effectiveness, or (ii) improving participant outcomes without added cost; and**
 - b. Require rigorous – preferably randomized – evaluations to determine which of these innovations really work.**

For some programs, this would require legislation to expand the program’s waiver authority and/or tie that authority to a requirement for rigorous evaluations wherever feasible. Other programs already have sufficient authority, and the Subcommittee could encourage them to use it more widely and strategically to stimulate state/local innovation and evidence-building. We would be pleased to work with the Subcommittee, if helpful, to explore how the waiver-evaluation concept might be operationalized across various programs within its jurisdiction.

B. Authorize and encourage agencies to allow greater researcher access to administrative data, with appropriate privacy protections, so as to facilitate low-cost rigorous evaluations.

- 1. Researchers have shown it is often possible to conduct randomized trials at low cost, by measuring study outcomes with administrative data already collected for other purposes.** In a development that could revolutionize social policy and practice, researchers have shown it is often possible to conduct scientifically-rigorous evaluations of program effectiveness at low cost, addressing a major obstacle to their widespread use. Costs are reduced by measuring study outcomes using administrative data already collected for other purposes, such as child maltreatment rates, employment and earnings, student test scores, criminal arrests, receipt of government assistance, and health care expenditures. This eliminates what is typically the largest cost component of a rigorous study: locating each individual in the program and control group at various points in time after program completion, and administering interviews or tests to obtain their outcome data.

2. **It is now possible to rigorously test hundreds of social program models and strategies each year, rather than a select few (as is currently done).** As noted above, the government already funds a vast array of program models/strategies each year, but only a small fraction are rigorously evaluated. With impetus from policy officials, low-cost, rigorous evaluations could likely be embedded across a broad spectrum of such activities (recognizing that they may not be feasible in every area¹⁸).
3. **As illustrative examples –**
 - **The Subsidized Guardianship randomized trial, described above, cost a total of about \$100,000 over nine years –** because outcomes for the 2,400 children in the sample were measured with administrative data on foster care outcomes (such as placement in a permanent home) that the state of Illinois already collected for other purposes.
 - **The Reemployment and Eligibility Assessment trial, described above, cost about \$320,000 through the 12-18 month follow-up, based on the researchers’ rough estimate.** Even though the study had a very large sample (135,000 UI claimants), it was conducted at modest cost by measuring all outcomes using administrative data on UI receipt and earnings that the participating states collect already for other purposes.
 - **We recently developed a [brief](#) with five additional examples of trials costing between \$50,000 and \$300,000 – a fraction of the usual multimillion-dollar cost for such studies.** These studies all produced valid evidence that is of policy and practical importance – and, in some cases, identified actionable strategies that generate budget savings.¹⁹
4. **We recognize that the Subcommittee has played a leadership role on this issue with its recent bipartisan approval of H.R. 1896 – increasing researcher access, with appropriate privacy protections, to HHS’s National Directory of New Hires (NDNH).** We believe this legislation, if enacted, will greatly lower the cost and burden of conducting rigorous evaluations of employment programs, by enabling such studies to measure employment and earnings outcomes through NDNH data rather than engaging in costly new data collection (e.g., individual interviews).
5. **OMB is also seeking to advance greater use of low-cost rigorous evaluation methods, and could be a valuable partner to the Subcommittee in advancing such studies.** For example, OMB prominently featured the concept of low-cost RCTs in its May 2012 memo to the heads of the federal agencies on [Use of Evidence and Evaluation](#), and cited the brief we developed on such studies. The President’s FY 2014 budget does the same ([link](#), page 94).
6. **Given the great potential, we encourage the Subcommittee to explore, through oversight hearings or other means, what more can be done by Congress and/or the agencies to build, integrate, and facilitate researcher access to administrative data,** so as to enable low-cost, rigorous evaluations across a broad range of program areas.
- C. **Once an intervention has been proven through rigorous evaluations to improve participant outcomes and/or reduce cost, authorize agency programs to use waivers and other administrative actions to facilitate its widespread adoption with program funds** (while ensuring close adherence to the proven approach).

The reason we suggest this is that federal social programs generally do not have the statutory authority to use evidence of effectiveness as a key criterion for allocating program funds. (An important, but still relatively small, exception is the set of “tiered evidence” initiative that Congress has enacted in recent years.²⁰) As a result, research-proven, cost-saving interventions such as Reemployment and Eligibility Assessments and Subsidized Guardianship, described above, may never be widely implemented unless Congress steps in to change the authorizing legislation for the UI program (in the case of Reemployment and Eligibility Assessments²¹) or the Title IV-E Foster Care program (in the case of Subsidized Guardianship). Congress did act in the case of Subsidized Guardianship, as noted earlier, but this can be an unusual occurrence and, when it does happen, it often takes years to achieve.

If Congress were to authorize programs such as UI and Foster Care to take *administrative* action to foster wide implementation of proven, cost-saving interventions, it would create a much more efficient mechanism for translating credible research findings into practice, so as to improve the lives of program participants and/or produce savings to the taxpayer. Doing so would inject a dynamic for evidence-driven improvements into a social spending process where evidence currently has little role.

Conclusion: Evidence-based policy offers a demonstrated path to more effective, less expensive government. As discussed in my testimony, we believe it could provide the basis for a bipartisan effort to reinvent U.S. social spending, so as to greatly increase its effectiveness in improving people’s lives.

References

¹ These are 2012 dollars. Examples include: (i) the Riverside Greater Avenues for Independence (GAIN) Program (Stephen Freedman, Daniel Friedlander, Winston Lin, and Amanda Schweder, *The GAIN Evaluation: Five-Year Impacts on Employment, Earnings, and AFDC Receipt*, Working Paper 96.1, MDRC, July 1996; James Riccio, Daniel Friedlander, and Stephen Freedman, *GAIN: Benefits, Costs, and Three-Year Impacts of a Welfare-to-Work Program*, MDRC, September 1994); (ii) Los Angeles Jobs-First GAIN (Stephen Freedman, Jean Tansey Knab, Lisa A. Gennetian, and David Navarro, *The Los Angeles Jobs-First GAIN Evaluation: Final Report on a Work First Program in a Major Urban Center*, MDRC, June 2000); and (iii) Portland Job Opportunities and Basic Skills (JOBS) Training Program (Susan Scrivener, Gayle Hamilton, Mary Farrell, Stephen Freedman, Daniel Friedlander, Marisa Mitchell, Jodi Nudelman, Christine Schwartz, *National Evaluation of Welfare-to-Work Strategies: Implementation, Participation Patterns, Costs, and Two-Year Impacts of the Portland (Oregon) Welfare-to-Work Program*, MDRC, May 1998; Gayle Hamilton, Stephen Freedman, Lisa Gennetian, Charles Michalopoulos, Johanna Walter, Diana Adams-Ciardullo, Anna Gassman-Pines, Sharon McGroder, Martha Zaslow, Jennifer Brooks, Surjeet Ahluwalia, Electra Small, and Bryan Ricchetti, *National Evaluation of Welfare-to-Work Strategies: How Effective Are Different Welfare-to-Work Approaches? Five-Year Adult and Child Impacts for Eleven Programs*, MDRC and Child Trends, December 2001).

² Ron Haskins, "What Works Is Work: Welfare Reform and Poverty Reduction," *Northwestern Journal of Law and Social Policy*, vol. 4, no. 1, 2009, pp. 29-60. Ron Haskins, in *Rigorous Evidence: The Key To Progress Against Crime and Substance Abuse? Lessons From Welfare, Medicine, and Other Fields*, Proceedings of a National Policy Forum Sponsored by the U.S. Department of Justice and Coalition for Evidence-Based Policy, June 14, 2004, pp. 30-36. Judith M. Gueron, "Building Evidence: What It Takes and What It Yields," *Research on Social Work Practice*, vol. 17, no. 1, January 2007, pp. 134-142.

³ Mark F. Testa, *Subsidized Guardianship: Testing the Effectiveness of an Idea Whose Time Has Finally Come*, Children and Family Research Center, School of Social Work, University of Illinois at Urbana-Champaign, May 2008.

⁴ Mark F. Testa and Jennifer Miller, "Evolution of Legal Guardianship as a Child Welfare Resource," in G. Mallon and P. McCartt Hess (editors), *Child Welfare for the 21st Century: A Handbook of Practices, Policies, and Programs*, second edition, Columbia University Press, in press.

⁵ Eileen Poe Yamagata, Jacob Benus, Nicholas Bill, Hugh Carrington, Marios Michaelides, and Ted Shen, *Impact of the Reemployment and Eligibility Assessment Initiative*, Impaq International, June 2011. Marios Michaelides, Eileen Poe-Yamagata, Jacob Benus, and Dharmendra Tirumalasetti, *Impact of the Reemployment Eligibility Initiative In Nevada*, Impaq International, January 2012.

⁶ The study measured the program's effect on job earnings in the Florida site, and found that it produced a statistically-significant \$476 increase in earnings per claimant over a 12-month follow-up period, roughly offsetting claimants' loss in UI benefits. However, this was a regression-adjusted effect on earnings; the unadjusted effect was near zero and not statistically significant. Because the effect differed under these two different estimation approaches, we believe the positive findings for earnings in Florida are best viewed as tentative, and need corroboration in future studies before being accepted as valid.

⁷ The \$1.5 billion in net savings is calculated by multiplying the savings per claimant (\$180) by the number of claimants potentially eligible for REA nationwide. We estimate that there are 7.8 million such claimants in the United States, based on (i) Department of Labor data showing a total of 19.4 million UI claims filed in 2012, and (ii) the study's finding that, on average, about 40% of UI claimants met the REA eligibility requirements in the states participating in the study. A per-claimant savings of \$180 multiplied by 7.8 million claimants totals roughly \$1.5 billion.

⁸ Because of the initial positive findings in Nevada, the researchers conducted a longer-term follow-up, which found that the program produced \$672 in per-person net savings during the 20-26 months after random assignment. We

estimated national savings of \$5 billion by multiplying \$672 (in per-person net savings) by 7.8 million (the number of claimants nationwide that we estimate are eligible for REA, as described in the previous endnote).

⁹ A summary of the evidence on NFP, including citations to the original study reports, is [linked here](#).

¹⁰ Coalition for Evidence-Based Policy, *Randomized Controlled Trials Commissioned by the Institute of Education Sciences Since 2002: How Many Found Positive Versus Weak or No Effects*, July 2013, [linked here](#).

¹¹ This is based on a count of results from the Department of Labor RCTs that have reported results since 1992, as identified through the Department's research database ([link](#)). We are preparing a short summary of these findings, to be released in the next few weeks.

¹² Ioannidis 2005, Zia et. al., 2005, Chan et. al., 2008 (see footnote 6 for full citations).

¹³ Jim Manzi, *Uncontrolled: The Surprising Payoff of Trial-and-Error for Business, Politics, and Society*, Perseus Books Group, New York, 2012, pp. 128 and 142. Jim Manzi, *Science, Knowledge, and Freedom*, presentation at Harvard University's Program on Constitutional Government, December 2012, [linked here](#).

¹⁴ Carmen DeNavas-Walt, Bernadette D. Proctor, and Jessica C. Smith, U.S. Census Bureau, Current Population Reports, P60-243, *Income, Poverty, and Health Insurance Coverage in the United States: 2011*, U.S. Government Printing Office, Washington, DC, 2012. U.S. Census Bureau, *Official and National Academy of Sciences (NAS) Based Poverty Rates: 1999 to 2010*, 2011. Kathleen Short, U.S. Census Bureau, HHES Division, *Estimating Resources for Poverty Measurement, 1993 – 2003*, 2005. Panel on Poverty and Family Assistance, National Academy of Sciences, *Measuring Poverty: A New Approach*, 1995, pp. 31-36.

¹⁵ Rampey, B.D., G.S. Dion, and P.L. Donahue, P.L. *NAEP 2008 Trends in Academic Progress*, NCES 2009-479, National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education, Washington, D.C., 2009.

¹⁶ Cornman, S.Q., and A.M. Noel, *Revenues and Expenditures for Public Elementary and Secondary School Districts: School Year 2008-09 (Fiscal Year 2009)* (NCES 2012-313). U.S. Department of Education. Washington, DC: National Center for Education Statistics, 2011. Richard H. Barr, *Revenues and Expenditures for Public Elementary and Secondary Education, 1973-74* (NCES-76-140). U.S. Department of Health, Education & Welfare, National Institute of Education. Washington, DC: National Center for Education Statistics, 1976.

¹⁷ U.S. Census Bureau, Current Population Reports, 2012, op. cit., no. 1. This refers to inflation-adjusted income. It includes income from the economy (such as earnings) but not government transfers (such as Food Stamps). However, there is evidence to suggest that the overall story of income stagnation for the bottom 40% of households changes little even when one adjusts income for government transfers and other items that affect household living standards. Specifically, the Census Bureau's alternative, National Academy of Sciences-based poverty measures make adjustments for government transfers, as well as factors such as state and local taxes, work expenses such as child care, out-of-pocket medical expenses, and geographic differences in housing costs. These adjustments change the poverty rate in any given year, as well as the composition of those in poverty, but do not change the overall trend in the poverty rate over time – i.e., no overall progress since the 1970s. (The relevant citations are in endnote 11.) Although the National Academy-based poverty measures only apply to a subset of the bottom 40% of U.S. households, their corroboration of no meaningful improvement for that key subset suggest that similar findings would be obtained for the larger group.

¹⁸ To do a low-cost randomized evaluation, certain conditions must exist, such as (i) low-cost access to administrative data of reasonable quality to measure key outcomes; (ii) a sizable number of individuals (or other units) available to participate in the study without special recruitment efforts; and (iii) the approval of key policy or program officials to do a randomized study. When carried out, such studies may not be able to examine all questions studied in a traditional, more comprehensive evaluation, such as the program's effect on outcomes other than those measurable with administrative data; the nature of any services received by the control group and how they differ from those delivered by the program; and possible reasons why a program has (or does not have) an effect. But,

increasingly, they can be used to answer the fundamental policy question: does the program actually improve people's lives?

¹⁹ Coalition for Evidence-Based Policy, *Rigorous Program Evaluations on a Budget: How Low-Cost Randomized Controlled Trials Are Possible in Many Areas of Social Policy*, March 2012, [link](#)

²⁰ The tiered evidence initiatives are described in Ron Haskins and Jon Baron, *Building the Connection between Policy and Evidence: The Obama Evidence-Based Initiatives*, commissioned by the U.K. National Endowment for Science, Technology and the Arts (NESTA), September 2011, [link](#).

²¹ The REA program is currently implemented in 42 states but, due to limited program funding, only serves a small percentage of UI claimants who are eligible for REA in those states. For example, in the four states that participated in the REA randomized trial, the program served an average of about 10% of REA-eligible claimants. See Yamagata et. al., 2011 (endnote 5), page 16.