

Board of Advisors

Robert Boruch
University of Pennsylvania

Jonathan Crane
Coalition for Evidence-Based
Policy

David Ellwood
Harvard University

Judith Gueron
MDRC

Ron Haskins
Brookings Institution

Blair Hull
Matlock Capital

Robert Hoyt
Jennison Associates

David Kessler
Former FDA Commissioner

Jerry Lee
Jerry Lee Foundation

Dan Levy
Harvard University

Diane Ravitch
New York University

Howard Rolston
Abt Associates

Isabel Sawhill
Brookings Institution

Martin Seligman
University of Pennsylvania

Robert Solow
Massachusetts Institute of
Technology

Nicholas Zill
Westat, Inc.

President

Jon Baron
jbaron@coalition4evidence.org
202-683-8049

1725 I Street, NW, Suite 300
Washington, DC 20006
202-349-1130
www.coalition4evidence.org

CBO/OMB Budget Scoring Guidance Could Greatly Accelerate Development of Rigorous Evidence About “What Works” To Achieve Major Budget Savings

Overview: *The solution to the nation’s long-term deficit problem is sometimes portrayed as a choice among sharp budget cuts, major tax increases, or a combination of the two. Given the magnitude of the problem, some level of sacrifice is probably unavoidable. But largely overlooked in the discussion are clear examples, from welfare and health care policy in the 1980s and 1990s, in which rigorous randomized experiments identified program reforms that produced important budgetary savings without adverse effects and, in some cases, with improvements in people’s lives. Similar cost-saving opportunities already exist in a few areas, and many more could likely be identified through rigorous research.*

Importantly, we believe that a modest clarification in CBO and OMB budget scoring guidance, described below, could provide a strong new impetus for the policy and research community to conduct such research, and thereby create a smarter, less painful path to major budget savings.

1. Rigorous studies have identified a few programs/reforms that produce sizable savings in federal entitlement spending, illustrating how such studies can build the evidence needed for significant and smart spending reductions.

- **Example: Welfare reform.** In the 1980s and 90s, federal officials sponsored many large experimental (i.e., randomized) evaluations of state and local welfare reforms. These studies showed convincingly that certain reform models that emphasized moving participants quickly into the workforce through short-term job search assistance and training – as opposed to providing remedial education – produced large gains in employment and earnings, reductions in welfare, and net entitlement savings (in AFDC and Food Stamps) of \$1700-\$6000 per participant.¹ Such findings helped shape the 1996 federal welfare reform act and the major work-focused reforms in state and local welfare programs that followed.
- **Example: Medicare home health agencies.** In 1995, federal officials launched a rigorous experimental evaluation to test prospective payment of Medicare home health agencies – i.e., paying such agencies an up-front lump sum per patient – against the usual cost-reimbursement approach. The evaluation found that prospective payment reduced costs to Medicare by 20% over three years, compared to cost reimbursement, with no adverse effects on patient health.² This finding helped shape Medicare’s nationwide implementation of prospective payment for home health agencies in 2000, producing large cost savings in this \$15 billion program.³
- **Example: Transitional Care Model.** Older adults are discharged from U.S. hospitals 13 million times each year; more than one-third are rehospitalized within 90 days, generating major costs to Medicare.⁴ The Transitional Care Model is a nurse-led hospital discharge and home follow-up program for chronically-ill older adults, designed to prevent health problems and rehospitalizations. It has been identified as one of the few social programs meeting the highest (Congressional “Top Tier”) standard of evidence, based on two well-conducted randomized experiments carried out in real-world hospital settings.⁵ In these studies, the program was found to produce a 30-50% reduction in rehospitalizations, and net savings in health care expenditures of about \$4,000 per patient, within 5-12 months after patient discharge – without any adverse effects on patient health or quality of life.

2. To identify enough of these cost-saving strategies to produce sizable long-term deficit reduction, many more rigorous experiments testing a wide range of strategies are needed.

In part, this is because experience suggests that many of the strategies tested, including those backed by promising preliminary evidence, will be found not to work. For example, the welfare experiments – in addition to identifying a few reform strategies with large effects – showed that many strategies thought to be effective based on expert opinion or preliminary studies actually had little impact on welfare and employment.

Similarly, the federal government’s ongoing Medicare Coordinated Care Demonstration (MCCD) is a large experimental evaluation of 15 cost-saving strategies for Medicare patients with chronic conditions, designed to coordinate care among their many physicians. A systematic review of earlier, more preliminary studies (quasi-experiments and small randomized trials) suggested that such programs reduced hospitalizations and health care costs, often by 25% or more.⁶ By contrast, the more definitive MCCD study has found – in follow-up one to three years after random assignment – that none of the 15 strategies is producing net savings in Medicare costs. On average, they have actually increased such costs by 11%.⁷ This is a typical pattern across diverse policy areas: promising preliminary evidence is often not confirmed in more definitive experimental studies.

3. Recommendation: That CBO and OMB make clear in their budget scoring guidance that when they consider available evidence to estimate a program’s effect on entitlement spending, they will give greater weight to findings from rigorous experimental research, particularly when conducted in real-world community settings with replication across multiple sites or studies.

- **Given budget scoring’s central role in the legislative process, such guidance would create a strong new incentive for federal agencies and others to build reliable evidence about how to reduce entitlement spending.** Specifically, it would incentivize the policy and research community – including the federal agencies that sponsor many of the large studies – to carry out rigorous experimental studies that measure program impact on entitlement spending as one of the key outcomes. Doing so would have a clear payoff: programs able to establish entitlement savings through such research would be scored as costing less than comparable programs without such evidence, and thus would have an important advantage in obtaining Congressional funding. Currently, there is little awareness in the policy and research community that such advantage is possible; and experimental studies, when they are undertaken, rarely measure entitlement savings as an outcome.
- **Such a statement would be consistent with evidence standards articulated by the National Academies,⁸ Institute of Education Sciences,⁹ U.S. Preventive Services Task Force,¹⁰ FDA,¹¹ and other respected scientific bodies** – all of which recognize findings from well-conducted randomized experiments, carried out in real-world environments, as the strongest method for measuring an intervention’s impact.
- **OMB has already taken a key step forward, recently advising the federal agencies that it will score entitlement savings demonstrated through rigorous experimental research.** However, this guidance is not yet available to the larger policy community. Furthermore, CBO – whose budget scoring decisions govern Congressional legislation – would need to provide similar guidance in order to send the community a clear signal about the potential policy impact of such research.

Conclusion: Experimental research has shown that it can produce important evidence about “what works” to reduce entitlement spending. CBO and OMB guidance could create a strong new impetus for such research, generating reliable evidence to help point the way to major budget savings.

References

¹ These are 2011 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 Valerie Cheh, *The Final Evaluation Report on the National Home Health Prospective Payment Demonstration: Agencies Reduce Visits While Preserving Quality.* Report submitted by Mathematica Policy Research, Inc. to the Health Care Financing Administration, April 30, 2001.

³ U.S. General Accounting Office, *Medicare Home Health Care: Prospective Payment System Will Need Refinement as Data Become Available*, GAO/HEHS-00-9, April 2000.

⁴ Stephen Jencks, Mark Williams, and Eric Coleman. "Rehospitalization Among Patients in Medicare Fee-For-Service Program," *New England Journal of Medicine*, 2009, vol. 360, no. 14, pp. 1418–1428. K. Levitt, L. Weir, E. Strangest, K. Ryan, and A. Eli Hauser, *HCUP Facts and Figures: Statistics on Hospital-based Care in the United States*, 2007. Rockville, MD: Agency for Healthcare Research and Quality, 2009.

⁵ Mary D. Naylor, Dorothy A. Bronte, Roberta L. Campbell, Greg Maslin, Kathleen M. McCauley, and J. Sanford Schwartz. "Transitional Care of Older Adults Hospitalized with Heart Failure: A Randomized, Controlled Trial." *Journal of American Geriatric Society*, 2004, vol. 52, no. 7, pp. 675-684. Mary D. Naylor, Dorothy Bronte, Roberta Campbell, Barbara S. Jacobsen, Math D. Mosey, Mark V. Pauli, and J. Sanford Schwartz. "Comprehensive Discharge Planning and Home Follow-up of Hospitalized Elders: A Randomized Clinical Trial." *JAMA*, 1999, vol. 281, no. 7, pp. 613-620. Top Tier Evidence initiative, Coalition for Evidence-Based Policy, http://toptierevidence.org/wordpress/?page_id=582.

⁶ Arnold Chen, Randall Brown, Nancy Archibald, Sherry Aliotta, and Peter Fox, *Best Practices in Coordinated Care*, prepared for the Health Care Financing Administration by Mathematica Policy Research, Inc., March 22, 2000.

⁷ Deborah Peikes, Arnold Chen, Jennifer Schore, and Randall Brown, "Effects of Care Coordination on Hospitalization, Quality of Care, and Health Care Expenditures Among Medicare Beneficiaries: 15 Randomized Trials," *JAMA*, vol. 301, no. 6, February 11, 2009, pp. 603-618.

⁸ National Research Council and Institute of Medicine, *Preventing Mental, Emotional, and Behavioral Disorders Among Young People: Progress and Possibilities*, Mary Ellen O'Connell, Thomas Boat, and Kenneth E. Warner, Editors (Washington DC: National Academies Press, 2009), recommendation 12-4, p. 371, [linked here](#).

⁹ U.S. Department of Education, "Scientifically-Based Evaluation Methods: Notice of Final Priority," *Federal Register*, vol. 70, no. 15, January 25, 2005, pp. 3586-3589, [linked here](#).

¹⁰ U.S. Preventive Services Task Force, "Current Methods of the U.S. Preventive Services Task Force: A Review of the Process," *American Journal of Preventive Medicine*, vol. 20, no. 3 (supplement), April 2001, pp. 21-35.

¹¹ The Food and Drug Administration's standard for assessing the effectiveness of pharmaceutical drugs and medical devices, at 21 C.F.R. §314.126, [linked here](#).