

BACKGROUND

No. 2884 | MARCH 19, 2014

Do Federal Social Programs Work?

David B. Muhlhausen, PhD

Abstract

Federal social programs consistently fail to achieve their objectives, yet federal spending on social programs has grown faster than the U.S. economy and population over the past 50 years. Scientifically rigorous multisite experimental evaluations have found a consistent pattern of failure in federal social programs, including early childhood education programs such as Head Start. In the current budgetary environment of high deficits and ballooning federal debt, Congress has a moral obligation to cut funding for and even eliminate these ineffective social programs.

Do federal social programs work? This is a simple question. While the question may be straightforward, however, finding an answer is complicated. To answer in the affirmative, federal social programs must ameliorate the social problems they target. In essence, social programs seek to improve human behavior in ways that will make people better off. For example, the social programs of the Great Society sought to eradicate the fundamental causes of poverty by providing opportunity for the poor to join other Americans in prosperity.¹

As used in this paper and the book by the same name,² the term “social program” refers to efforts by the federal government that attempt to improve human behavior by increasing skills or awareness, chiefly through noncompensatory services. These programs engage in social engineering that attempts to enhance the well-being of citizens. Social programs are intended to fix social problems that individuals are assumed to be unable to solve themselves. Head Start is a classic example of a social program. Created as part

KEY POINTS

- Determining the effectiveness of federal social programs is particularly relevant to our nation’s current political debate over the federal government’s persistent deficits and debt.
- Using evidence from scientifically rigorous multisite experimental evaluations of national programs since 1990, the book *Do Federal Social Programs Work?* demonstrates that federal social programs—such as Early Head Start, Head Start, and numerous job-training programs—are ineffective.
- The American public should have nothing to fear from the elimination of ineffective programs. Now is the time for deep budget cuts in federal social programs.
- Multisite experimental evaluations are the best method for assessing the effectiveness of federal social programs. Yet to date, this method has been used on only a handful of federal social programs. Congress needs to reverse this trend.

This paper, in its entirety, can be found at <http://report.heritage.org/bg2884>

Produced by the Center for Data Analysis

The Heritage Foundation
214 Massachusetts Avenue, NE
Washington, DC 20002
(202) 546-4400 | heritage.org

Nothing written here is to be construed as necessarily reflecting the views of The Heritage Foundation or as an attempt to aid or hinder the passage of any bill before Congress.

of the War on Poverty in 1965, Head Start is a preschool community-based program that is intended to provide a boost to disadvantaged children before they enter elementary school.

In the federal budget, social programs are considered discretionary and grouped with “other mandatory” domestic programs. This classification includes numerous education, welfare, housing, and employment programs.

Evidence from scientifically rigorous multisite experimental evaluations of national programs published since 1990 demonstrates that federal social programs such as Head Start are ineffective. The American people have nothing to fear from the elimination of ineffective programs.

Determining the effectiveness of federal social programs is particularly relevant given the current political debate over the federal government’s persistent deficits and debt. Many of the budget plans in Congress reduce the rate of spending increases on federal social programs. Very few plans actually propose real spending reductions.

Opponents of spending reductions assert that spending any less on social programs will have disastrous effects on society. In April 2012, for example, President Barack Obama called the budget plan passed by the Republican-controlled U.S. House of Representatives a “Trojan Horse” and “thinly veiled social Darwinism.”³ In particular, the President said, “If this budget becomes law and the cuts were applied evenly, starting in 2014, over 200,000 chil-

dren would lose their chance to get an early education in the Head Start program.”⁴ The clear implication is that over 200,000 children will somehow be harmed by not attending Head Start. This would be true only if Head Start is an effective program that actually benefits the children it serves.

Using evidence from scientifically rigorous multisite experimental evaluations of national programs published since 1990, the book *Do Federal Social Programs Work?* demonstrates that federal social programs such as Head Start are ineffective. The American people have nothing to fear from the elimination of ineffective programs.

The findings of multisite experimental evaluations are reliable because they assess the performance and effectiveness of federal social programs in multiple locations. While individual programs operating in single locations may undergo experimental evaluations, these small-scale, single-site evaluations do not inform federal policymakers of the general effectiveness of national programs. The success of a single program that serves a particular jurisdiction or population does not necessarily mean that the same program will achieve similar success on a national scale. Thus, small-scale evaluations are poor substitutes for large-scale evaluations. Yet many advocates of social engineering make grandiose claims about the potential effectiveness of small-scale programs implemented on the national scale.⁵

Out-of-Control Spending

On December 31, 2011, the gross debt racked up by the federal government reached \$15.2 trillion—the legal limit authorized by Congress.⁶ In response, President Obama formally notified Congress on January 12, 2012, of his intent to raise the nation’s debt ceiling by \$1.2 trillion from \$15.2 trillion to \$16.4

1. Patrick McGuinn and Frederick Hess, “The Great Society and the Evolution of the Elementary and Secondary Education Act of 1965,” in Sidney M. Milkis and Jerome M. Mileur, eds., *The Great Society and the High Tide of Liberalism* (Boston: University of Massachusetts Press, 2005), pp. 311–312, and Jerome M. Mileur, “The Great Society and the Demise of New Deal Liberalism,” in Milkis and Mileur, *The Great Society and the High Tide of Liberalism*, p. 436.

2. This paper is a summary of David B. Muhlhausen, *Do Federal Social Programs Work?* (Santa Barbara, CA: Praeger, 2013).

3. Barack Obama, “Remarks by the President at the Associated Press Luncheon,” Washington, DC, April 3, 2012, <http://www.whitehouse.gov/the-press-office/2012/04/03/remarks-president-associated-press-luncheon> (accessed February 26, 2014).

4. Ibid.

5. For a detailed discussion, see Muhlhausen, *Do Federal Social Programs Work?*

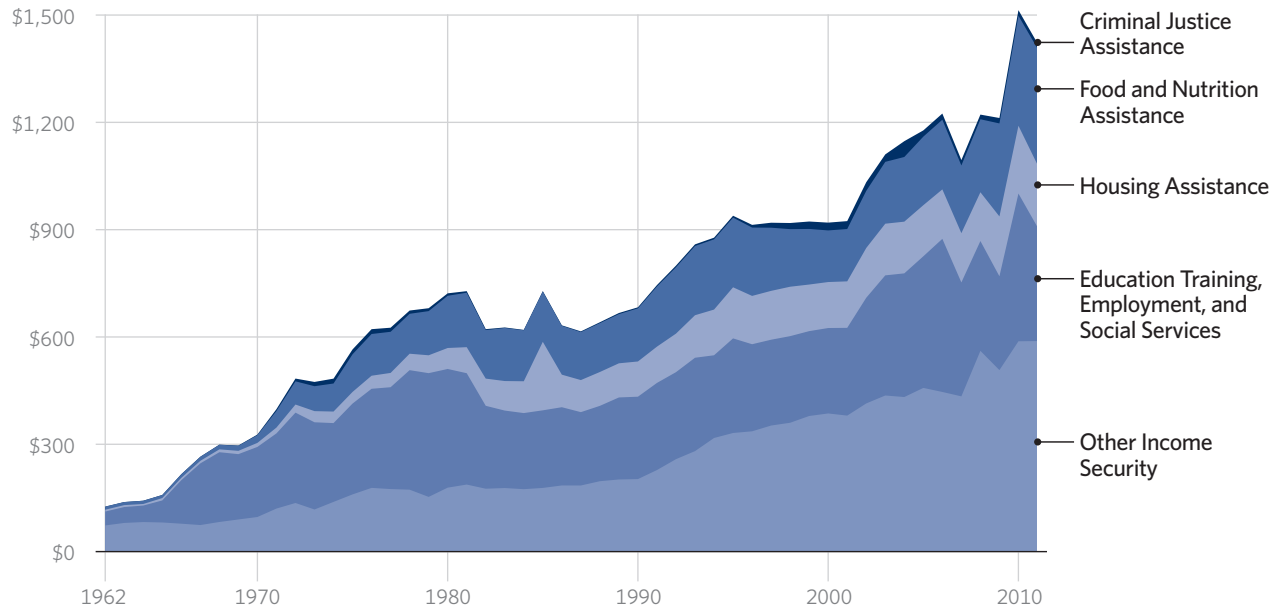
6. U.S. Department of the Treasury, Bureau of the Public Debt, “Monthly Statement of the Public Debt of the United States,” December 31, 2011, Table 1, <http://www.treasurydirect.gov/govt/reports/pd/mspd/2011/opds122011.pdf> (accessed February 26, 2014).

CHART 1

Per-Capita Spending on Social Programs Has Risen Significantly

From 1962 to 2011, per-capita spending on social programs increased by 1,031 percent.

PER-CAPITA SPENDING ON FEDERAL SOCIAL PROGRAMS, IN 2010 DOLLARS



Sources: U.S. Office of Management and Budget, *Budget of the United States, Fiscal Year 2013: Historical Tables* (Washington, DC: U.S. Government Printing Office, 2012), Tables 3.2, 10.1; and U.S. Census Bureau, Population Estimates Program, Population Division, "Historical National Population Estimates."

B2884 heritage.org

trillion.⁷ At the end of fiscal year (FY) 2013 on September 30, 2013, the federal government's gross debt was expected to reach \$17.5 trillion or 107.4 percent of gross domestic product (GDP).⁸ As of February 14, 2014, the gross debt was \$17.3 trillion.⁹ This is a staggering sum that is difficult for Americans to grasp. If we did, we would be truly frightened at the prospect of paying it off.

While entitlement spending is the primary driver of the federal debt, spending on social programs is not trivial. Obtaining exact data on how much the federal government spends on social

programs is difficult. The Office of Management and Budget (OMB) started classifying spending by subfunction categories in 1962. The Education, Training, Employment, and Social Services function and the Housing Assistance, Food and Nutrition Assistance, Other Income Security, and Criminal Justice Assistance subfunctions are used to estimate the amount of taxpayer dollars spent on federal social programs.¹⁰ Within these categories are a range of social programs intended to change human behavior for the better. While this measure is imperfect, it is a practical estimate of social

7. David Nakamura, "Obama Asks Congress for Debt Limit Hike," *The Washington Post*, January 12, 2012, http://www.washingtonpost.com/blogs/44/post/after-delay-obama-asks-congress-for-debt-limit-hike/2012/01/12/gIQA3ADuP_blog.html (accessed February 26, 2014).

8. U.S. Office of Management and Budget, *Budget of the United States, Fiscal Year 2013: Historical Tables* (Washington, DC: U.S. Government Printing Office, 2012), pp. 139-140, Table 7.1, <http://www.whitehouse.gov/sites/default/files/omb/budget/fy2013/assets/hist.pdf> (accessed February 26, 2014).

9. United States Department of the Treasury, "The Debt to the Penny and Who Holds It," February 14, 2014, <http://www.treasurydirect.gov/NP/debt/current> (accessed February 26, 2014).

10. *Ibid.*, Table 3.2.

program spending. Social programs funded under these categories include Temporary Assistance for Needy Families (TANF), job-training and education programs, juvenile delinquency prevention programs, and many other programs. The figures presented include only federal outlays and ignore spending by state and local governments.

A growing population means more spending on social programs. To account for population growth, Chart 1 presents federal social program spending on a per capita basis in 2010 dollars. In 1962, social program spending was \$125.67 per capita. In 2011, the figure reached \$1,421.02 per capita—a 1,031 percent increase. Total U.S. population grew from 186,537,737 people in 1962¹¹ to 311,591,917 people in 2011—an annual growth rate of 1.1 percent.¹² An annual growth rate of 1.1 percent means that the total population will double in size every 91 years. In contrast, the annual growth rate of 5.1 percent in social program spending per capita means that spending in this category doubles every 19 years.

Why does the federal government overspend so much? To arrive at an answer, we need to understand how greatly the federal government's scope, power, and responsibilities have expanded. Starting with the Progressive Era, running through the New Deal, and ending with the Great Society, the original understanding of the government's role in protecting our freedoms that was established during the American Founding was redefined with the call for a much more activist federal government. Each of these political waves sought to transform America into something very different from what the Founding Fathers envisioned.¹³

Based on the Founders' conception, once government secures our natural rights or formal freedom as expressed in the Declaration of Indepen-

dence, the responsibility for obtaining our hopes, desires, and economic security—while not guaranteed—is up to us. The Progressives replaced the Founders' notion of natural rights with a new positive or effective freedom that required government to assist individuals in achieving their full potential as human beings. Securing formal freedoms was not enough to allow individuals to be truly free.¹⁴ Individuals must be given the capacity and resources to achieve “effective” freedom.¹⁵ Instead of protecting our formal freedoms as understood by the Founders, Woodrow Wilson asserted that “the individual must be assured the best means, the best and fullest opportunities, for complete self-development.”¹⁶ Federal social programs must be created to help individuals acquire the necessary resources to be effectively free.

This fundamental transformation of the notion of freedom allowed Franklin D. Roosevelt to assert that “[e]very man has a right to life; and this means that he has also a right to make a comfortable living.”¹⁷ Instead of requiring that Americans not harm each other, we are, according to Roosevelt, obligated to ensure that every American has a comfortable living.

Unlike the New Dealers, the advocates of the Great Society sought not to provide relief and social insurance programs, but rather to social engineer a better society. In May 1964, President Lyndon B. Johnson envisioned the Great Society as

a place where every child can find knowledge to enrich his mind and to enlarge his talents. It is a place where leisure is a welcome chance to build and reflect, not a feared cause of boredom and restlessness. It is a place where the city of man serves not only the needs of the body and the

11. U.S. Census Bureau, Population Division, “Historical National Population Estimates: July 1, 1900 to July 1, 1999,” June 28, 2000, <http://www.census.gov/popest/data/national/totals/pre-1980/tables/popclockest.txt> (accessed February 26, 2014).
12. U.S. Census Bureau, Population Division, “Annual Estimates of the Population for the United States, Regions, States, and Puerto Rico: April 1, 2010 to July 1, 2011,” December 2011, <http://www.census.gov/popest/data/state/totals/2011/tables/NST-EST2011-01.xls> (accessed February 26, 2014).
13. Charles R. Kesler, *I Am the Change: Barack Obama and the Crisis of Liberalism* (New York: Broadside Books, 2012).
14. John Dewey, *Liberalism and Social Action* (Amherst, NY: Prometheus Books, 2000), and John Dewey, *Individualism Old and New* (Amherst, NY: Prometheus Books, 1999).
15. John Dewey and James H. Tufts, *Ethics* (New York: Henry Holt and Company, 1910).
16. Woodrow Wilson, *The State: Elements of Historical and Practical Politics* (Boston: D.C. Heath & Co., Publishers, 1910), p. 633.
17. Franklin D. Roosevelt, “Campaign Address on Progressive Government at the Commonwealth Club in San Francisco, California,” September 23, 1932, <http://www.presidency.ucsb.edu/ws/index.php?pid=88391> (accessed February 26, 2014).

demands of commerce but the desire for beauty and the hunger for community.¹⁸

Johnson expressed freedom in terms of reaching the highest ideals. From the Progressive Era through the New Deal to the Great Society, the role of the federal government was greatly expanded. Instead of securing natural rights based on the principles of the Founding, the mission of the federal government now revolves around the Great Society's objective of providing effective freedom that alleviates material hardship and promotes the realization of individual fulfillment.

Operating with increasingly scarce resources, federal policymakers need to start denying funds to ineffective programs, even if calls for funding these programs seem morally compelling.

Democrats and Republicans have embraced the progressive cause of providing social programs to assist individuals in achieving their full potential as human beings. Further, support for federal social engineering programs crosses ideological lines. Despite the transformation of freedom coming from the Left, many on the Right have embraced federal social engineering.¹⁹

Holding Social Programs Accountable

Given the fiscal crises that the federal government is facing, holding federal social programs accountable for their performance is necessary to regain control of excessive spending. Why should Congress routinely spend taxpayer dollars on failed social programs that do not work?

Operating with increasingly scarce resources, federal policymakers need to start denying funds to ineffective programs, even if calls for funding these programs *seem* morally compelling. Calling for more spending on social programs may seem morally compelling, but continuing to spend taxpayer dol-

lars on programs that do not produce their intended results is morally indefensible.

Social programs should be carefully evaluated to determine whether they do in fact work. Determining whether these programs work requires reliably sorting out the effect of a social program from confounding factors, which is a difficult task. While large-scale experimental evaluations are the best method for assessing cause-and-effect relationships, there are additional methods to judge the merits of social programs. Most notable is the legal matter of whether or not the U.S. Constitution provides Congress the authority to create social programs in the first place. In addition, the worthiness of spending hundreds of billions of dollars on social programs should be judged in light of the fact that the federal government's gross debt is \$17.3 trillion.

Science Versus Anecdotal Observations.

There are numerous methods available for making sense of the world around us. We frequently make personal observations of events around us to bring order to our lives. We often assign cause-and-effect relationships to events we personally experience. For instance, learning that touching a hot stove will burn one's hand is an easy cause-and-effect association that does not need to be tested more than once. We can easily correlate the act of touching the stove with the pain felt. Firsthand experience is often instrumental to developing knowledge. Every day, we make personal observations that guide us in our activities. We often seek the advice of others based on their personal experiences.

However, the usefulness of personal observations or experiences can be undermined when assessing complex social interactions that can have multiple causes. This problem is particularly acute when assessing the effectiveness of social programs where multiple factors can cause the outcomes of interest. The best way to determine whether federal social programs do in fact work is to conduct large-scale, multisite experimental evaluations that attempt to isolate the direct effects of social programs from other factors that affect the outcomes of interest.

Impact evaluations often assess impacts by comparing treatment or intervention groups to control

18. Lyndon B. Johnson, "Remarks at the University of Michigan," May 22, 1964, <http://www.lbjlib.utexas.edu/johnson/archives.hom/speeches.hom/640522.asp> (accessed February 26, 2014).

19. For a detailed discussion on how the Left and Right have embraced federal social engineering, see Muhlhausen, *Do Federal Social Programs Work?*

or comparison groups. Determining the impact of social programs requires comparing the conditions of those who received assistance with the conditions of an equivalent group that did not experience the intervention. However, evaluations differ by the quality of methodology used to separate the net impact of programs from other factors that may explain differences in outcomes between comparison and intervention groups.²⁰

Experimental evaluations are the “gold standard” of evaluation designs. Randomized experiments attempt to demonstrate causality by (1) holding all other possible causes of the outcome constant, (2) deliberately altering only the possible cause of interest, and (3) observing whether the outcome differs between the intervention and control groups. In reality, all evaluation methods, including experimental designs, can never establish 100 percent certainty that all of the potential causes (confounding factors) were held constant.

When conducting an impact evaluation of a social program, identifying and controlling for all of the possible factors that influence the outcomes of interest is impossible. We simply do not have enough knowledge to accomplish this task. Even if we could identify all possible causal factors, collecting complete and reliable data on all of these factors would likely still be beyond our abilities. For example, it is impossible to isolate a person participating in a social program from his family in order to “remove” the influences of family.

This is where the benefits of random assignment become clear. Because we do not know enough about all possible causal factors to identify and hold them constant, randomly assigning test subjects to intervention and control groups allows us to have a high degree of confidence that these unidentified factors will not confound our estimate of the intervention’s impact. Random assignments should evenly distribute these unidentified factors between the intervention and control groups of an experimental evaluation.

Standards for Assessing the Effectiveness of Federal Social Programs

Congress can take several steps to ensure that federal social programs are properly assessed using experimental evaluations. These experimental evaluations should be large in scale and based on multiple sites to avoid the problems of simplistic generalizations. Given the multitude of confounding factors that may influence the performance of social programs, the larger the size of the evaluation, the more likely that the federal social program will be assessed under all of the conditions under which it operates.

When Congress creates social programs, especially state and local grant programs, the funded activities are implemented in multiple cities or towns. While individual social programs operating in a single location and funded by the federal government may undergo experimental evaluations, these small-scale, single-site evaluations do not inform policymakers of the general effectiveness of national social programs. Small-scale evaluations assess only the impact on a small fraction of the people served by federal social programs. The success of a single program that serves a particular jurisdiction or population does not necessarily mean that the same program will achieve similar success in other jurisdictions or among different populations. Put simply, small-scale evaluations are poor substitutes for large-scale evaluations.

Thus, federal social programs should be evaluated in multiple sites so that social programs can be tested in the various conditions in which they operate and in the numerous types of populations that they serve. In addition, a multisite experimental evaluation that examines the performance of a particular program in numerous and diverse settings can potentially produce results that are more persuasive to policymakers than results from a single locality.²¹

The case of police departments performing mandatory arrests in domestic violence incidents is a poignant example of why policymakers should exercise caution when generalizing findings from a single

20. For a detailed discussion of evaluation methodology, see *ibid.*

21. Erica B. Baum, “When the Witch Doctors Agree: The Family Support Act and Social Science Research,” *Journal of Policy Analysis and Management*, Vol. 10, No. 4 (Autumn 1991), pp. 603–615, and Judith M. Gueron, “The Politics of Random Assignment: Implementing Studies and Affecting Policy,” in Frederick Mosteller and Robert Boruch, eds., *Evidence Matters: Randomized Trials in Education Research* (Washington, DC: Brookings Institution, 2002), pp. 15–49.

evaluation.²² During the 1980s, criminologists Lawrence W. Sherman and Richard A. Berk analyzed the impact of mandatory arrests for domestic violence incidents on future domestic violence incidents in Minneapolis, Minnesota.²³ Compared to less severe police responses, the Minneapolis experiment found that mandatory arrests lead to significantly lower rates of domestic violence. Sherman and Berk urged caution, but police departments across the nation adopted the mandatory arrest policy based on the results of one evaluation conducted in one city.

Federal social programs should be evaluated in multiple sites so that social programs can be tested in the various conditions in which they operate and in the numerous types of populations that they serve.

However, what worked in Minneapolis did not always work in other locations. Experiments conducted by Sherman and others in Omaha, Nebraska; Milwaukee, Wisconsin; Charlotte, North Carolina; Colorado Springs, Colorado; and Dade County, Florida, found mixed results.²⁴ Experiments in Omaha, Milwaukee, and Charlotte found that mandatory arrests lead to long-term *increases* in domestic violence. Apparently, knowing that they would automatically be arrested prompted repeat offenders to become more abusive. It seems the offender followed the sick logic that if he were going to be automatically arrested and spend the night in jail, he might as well beat his wife even more. In a subsequent analysis of the disparate findings, Sherman postulated that arrested individuals who lacked a stake in conformity within their communities were significantly more likely to engage in domestic violence after arrest,

while arrested individuals who were married and employed were significantly less likely to commit further domestic violence infractions.²⁵

Large-Scale Multisite Experimental Evaluations

Despite the trillions of dollars that Congress has spent on federal social programs, only a few have undergone large-scale experimental impact evaluations. We have done our best to include all of the relevant multisite experimental evaluations of federal social programs that have been published since 1990. These 20 evaluations assessed the impact of 21 federal social programs:

- Early Head Start;
- Enhanced Early Head Start with Employment Services;
- Head Start;
- Even Start Family Literacy Program;
- 21st Century Community Learning Centers;
- Abstinence Education;
- Upward Bound;
- Food Stamp (renamed Supplemental Nutrition Assistance Program, or SNAP) Employment and Training Program;
- Welfare-to-Work;
- Employment Retention and Advancement (ERA) Project;
- Building Strong Families (BSF);

22. For a detailed discussion on the problem of overgeneralizing the results of small-scale evaluations and the difficulty of replicating the results of small-scale programs, see Muhlhausen, *Do Federal Social Programs Work?*

23. Lawrence W. Sherman and Richard A. Berk, "The Specific Deterrent Effects of Arrest for Domestic Assault," *American Sociological Review*, Vol. 49, No. 2 (April 1984), pp. 261-272.

24. Lawrence W. Sherman, *Domestic Violence: Experiments and Dilemmas* (New York: Free Press, 1992); Lawrence W. Sherman et al., "Crime, Punishment, and Stake in Conformity: Legal and Informal Control of Domestic Violence," *American Sociological Review*, Vol. 57, No. 5 (October 1992), pp. 680-690; and Lawrence W. Sherman et al., "The Variable Effects of Arrest on Criminal Careers: The Milwaukee Domestic Violence Experiment," *Journal of Criminal Law & Criminology*, Vol. 83, No. 1 (Spring 1992), pp. 137-169.

25. Sherman, *Domestic Violence*.

- Supporting Healthy Marriage;
- Moving to Opportunity;
- Section 8 Housing Vouchers;
- Job Training Partnership Act (JTPA) programs;
- Unemployment Insurance Self-Employment Demonstrations;
- Project GATE (Growing America Through Entrepreneurship);
- Job Corps;
- JOBSTART;
- Center for Employment Training (CET) Replication; and
- Quantum Opportunity Program Demonstration.

Federal Social Programs for Children and Families

This section covers federal social programs that are intended to benefit a wide range of clients—from infants and toddlers to adult heads of household. Federal programs target a host of social problems, including low academic skills, poverty, personal relations, hard-to-employ workers, and low wages.

Young Children. Commenting on the lack of demonstrated effectiveness of federal social programs over the past 20 years, Isabel V. Sawhill of the Brookings Institution and Jon Baron of the Coali-

tion for Evidence-Based Policy wrote, “Only one program—Early Head Start (a sister program to Head Start, for younger children)—was found to produce meaningful, though modest, positive effects.”²⁶ While the short-term results of the evaluation were publicly available at the time that Sawhill and Baron made this statement, a subsequent long-term follow-up study released in 2012 found that the short-term benefits of Early Head Start quickly faded away—not an uncommon finding for social programs.²⁷

Early Head Start, created during the 1990s, is a federally funded community-based program that serves low-income families with pregnant women, infants, and toddlers up to age three. While the short-term findings indicated modest positive impacts, almost all of the positive findings for all Early Head Start Participants are driven by the positive findings for blacks.²⁸ The program had little to no effect on white and Hispanic participants, who are the majority of program participants.

For the long-term findings, the overall initial effects of Early Head Start at age three clearly faded away by the fifth grade. For the 11 child-social-emotional outcomes, none of the results were found to have statistically meaningful impacts.²⁹ Further, Early Head Start failed to have statistically measurable effects on the 10 measures of child academic outcomes, including reading, vocabulary, and math skills.

Despite Head Start’s long life, the program never underwent a thorough, scientifically rigorous evaluation of its effectiveness until Congress mandated an evaluation in 1998. The Head Start Impact Study began in 2002, and the immediate-term, short-term, and long-term results released in 2005, 2010, and 2012, respectively, are disappointing.³⁰ Overall, the

26. Isabel V. Sawhill and Jon Baron, “Federal Programs for Youth: More of the Same Won’t Work,” *Youth Today*, May 1, 2010, <http://coalition4evidence.org/wordpress/wp-content/uploads/Viewpoint-Essay-Sawhill-Baron-Youth-Today-May-2010.pdf> (accessed February 26, 2014).

27. John M. Love et al., *Making a Difference in the Lives of Infants and Toddlers and Their Families: The Impacts of Early Head Start, Volume 1: Final Technical Report*, Mathematica Policy Research, June 2002, http://www.mathematica-mpr.com/publications/redirect_PubsDB.asp?strSite=PDFs/ehsfinalvol1.pdf (accessed February 26, 2014), and Cheri A. Vogel et al., *Early Head Start Children in Grade 5: Long-Term Follow-Up of the Early Head Start Research Evaluation Project Study Sample: Final Report*, U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation, *OPRE Report No. 2011-8*, December 2010, <http://www.acf.hhs.gov/sites/default/files/opre/grade5.pdf> (accessed February 26, 2014).

28. Love et al., *Making a Difference in the Lives of Infants and Toddlers and Their Families*, pp. 381–385, Table VII.11, and pp. 386–388, Table VII.12.

29. Vogel et al., *Early Head Start Children in Grade 5*, Table III.2, pp. 24–25.

30. U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation, *Head Start Impact Study: First Year Findings*, June 2005, http://www.acf.hhs.gov/sites/default/files/opre/first_yr_finds.pdf (accessed February 26, 2014), and *Head Start Impact Study: Final Report*, January 2010, http://www.acf.hhs.gov/sites/default/files/opre/hs_impact_study_final.pdf (accessed February 26, 2014).

evaluation found that the program largely failed to improve the cognitive, social-emotional, health, and parenting outcomes of children who participated compared with the outcomes of similar children who did not participate. According to the report, “[T]he benefits of access to Head Start at age four are largely absent by 1st grade for the program population as a whole.”³¹ Alarmingly, Head Start actually had a harmful effect on three-year-old participants once they entered kindergarten. Teachers reported that non-participating children were more prepared in math skills than the children who participated in Head Start.

The third-grade follow-up to the Head Start Impact Study followed students’ performance through the end of third grade.³² The results shed further light on the ineffectiveness of Head Start. By third grade, Head Start had little to no effect on cognitive, social-emotional, health, or parenting outcomes of participating children.

In addition to these failures, evaluations of the Enhanced Early Head Start with Employment Services, which provides early childhood care and employment training services to families, and the Even Start Family Literacy Program, which is intended to meet the basic educational needs of parents and children, failed to produce statistically meaningful impacts.³³

School-Age Children. Federal social programs targeting school-age children do not fare any better. An evaluation of the 21st Century Community Learning Centers, an after-school program intended to improve the academic performance of students, found that the program actually decreased the academic performance and increased the behavioral problems of participating students.³⁴

By third grade, Head Start had little to no effect on cognitive, social-emotional, health, or parenting outcomes of participating children.

Created in 1965, Upward Bound is an original War on Poverty social program intended to help economically disadvantaged students successfully complete high school and attend college.³⁵ However, an evaluation found that in general, the program has almost entirely no effect on these goals.³⁶ Other school-age social programs have been found to be ineffective too.³⁷

Families. Of all the social programs reviewed, welfare-to-work strategies have had the most positive results. In 1989, the federal government funded

-
31. U.S. Department of Health and Human Services, *Head Start Impact Study: Final Report*, p. xxxviii.
 32. Michael Puma et al., *Third Grade Follow-Up to the Head Start Impact Study: Final Report*, U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation, *OPRE Report 2012-45*, October 2012, http://www.acf.hhs.gov/sites/default/files/opre/head_start_report.pdf (accessed February 26, 2014).
 33. JoAnn Hsueh and Mary E. Farrell, *Enhanced Early Head Start with Employment Services: 42-Month Impacts from the Kansas and Missouri Sites of the Enhanced Services for the Hard-to-Employ Demonstration and Evaluation Project*, U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation, *OPRE Report No. 2012-05*, February 2012, http://www.acf.hhs.gov/sites/default/files/opre/kansas_missouri.pdf (accessed February 26, 2014); U.S. Department of Education, *Third National Even Start Evaluation: Program Impacts and Implications for Improvement*, 2003, <http://www2.ed.gov/rschstat/eval/disadv/evenstartthird/toc.html> (accessed February 26, 2014); and Anna E. Ricciuti et al., *Third National Even Start Evaluation: Follow-Up Findings from the Experimental Design Study*, U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, December 2004, <http://ies.ed.gov/ncee/pdf/20053002.pdf> (accessed February 26, 2014).
 34. Susanne James-Burdumy, Mark Dynarski, and John Deke, “When Elementary Schools Stay Open Late: Results from the National Evaluation of the 21st Century Community Learning Centers Program,” *Educational Evaluation and Policy Analysis*, Vol. 29, No. 4 (December 2007), pp. 296–318.
 35. David Myers and Allen Schirm, *The Short-Term Impacts of Upward Bound: An Interim Report*, Mathematica Policy Research, May 1997, p. 1, http://www.mathematica-mpr.com/publications/PDFs/education/upwardbound_interim.pdf (accessed February 26, 2014).
 36. U.S. Department of Education, Office of the Under Secretary, Policy and Program Studies Service, *The Impacts of Regular Upward Bound: Results from the Third Follow-Up Data Collection*, 2004, and Neil S. Seftor, Arif Mamun, and Allen Schirm, *The Impacts of Regular Upward Bound on Postsecondary Outcomes 7–9 Years After Scheduled High School Graduation: Final Report*, Mathematica Policy Research, January 2009, <http://mathematica-mpr.com/publications/pdfs/upwardboundoutcomes.pdf> (accessed February 26, 2014).
 37. Christopher Trenholm et al., *Impacts of Four Title V, Section 510 Abstinence Education Programs: Final Report*, Mathematica Policy Research, April 2007, <http://www.mathematica-mpr.com/publications/pdfs/impactabstinence.pdf> (accessed February 26, 2014).

the National Evaluation of Welfare-to-Work Strategies (NEWWS), which assessed the long-term effects of 11 mandatory welfare-to-work programs operating in seven sites during the late 1980s and 1990s.³⁸ According to Lawrence M. Mead, a professor of politics at New York University,

NEWWS aimed mainly to determine whether programs that emphasized training or work first were more effective and whether work tests for mothers would harm children. A clear verdict emerged that work first was best and that children were little affected. The congressional drafters of PRWORA [Personal Responsibility and Work Opportunity Reconciliation Act] in 1996 had already assumed that a radical, work-first reform was best, but the NEWWS results ratified that judgment.³⁹

Overall, NEWWS found that welfare-to-work strategies were more effective at increasing the earning of participants. In particular, strategies that focused on quick entry into the labor force moved welfare recipients into jobs more quickly and were more effective than strategies that focused on job training.⁴⁰

While the 1989 welfare-to-work evaluation showed consistent positive impacts on earnings, welfare reform experts argue that the spread of work-focused strategies encouraged many individuals to find employment in the first place rather than seeking welfare assistance. This effect should

surely be considered a positive benefit of welfare reform and was not part of the NEWWS evaluation. The NEWWS evaluation, with similar randomized experiments, used a control group that participates in some form of welfare assistance. This fact means that these experimental evaluations cannot assess the potential benefit of encouraging individuals to obtain employment instead of seeking welfare assistance and are an inadequate method for estimating changes in welfare caseloads.

Despite the success of welfare-to-work strategies, other social programs focusing on families that have undergone multisite experimental evaluations have been found to be ineffective.⁴¹ For example:

- The Food Stamp (renamed Supplemental Nutrition Assistance Program) Employment and Training Program failed to affect earnings and employment outcomes;⁴²
- The Employment Retention and Advancement Project, a program designed to provide additional employment and training services to Temporary Assistance to Needy Families participants and others with employment difficulties, was largely ineffective in improving earnings and employment outcomes;⁴³ and
- The evaluation of Moving to Opportunity/Section 8 housing vouchers found that housing subsidies intended to improve the lives of parents and

38. Gayle Hamilton et al., *National Evaluation of Welfare-to-Work Strategies: How Effective Are Different Welfare-to-Work Approaches? Five-Year Adult and Child Impacts for Eleven Programs*, U.S. Department of Health and Human Services and U.S. Department of Education, December 2001, <http://aspe.hhs.gov/hsp/NEWWS/5yr-11prog01/> (accessed February 26, 2014).

39. Lawrence M. Mead, *Expanding Work Programs for Poor Men* (Washington, DC: AEI Press, 2011), pp. 114-115.

40. For a discussion of the policy consequences of the NEWWS finding that strategies that emphasized work were more effective than strategies that focused on education and training, see Muhlhausen, *Do Federal Social Programs Work?*, p. 156, and Robert Rector and Brian M. Riedl, "Why Successful Welfare Reform Must Strengthen Work Requirements," Heritage Foundation *Backgrounders* No. 1568, July 12, 2002, <http://www.heritage.org/research/reports/2002/07/why-successful-welfare-reform-must-strengthen-work-requirements>.

41. Robert G. Wood et al., "Strengthening Unmarried Parents' Relationships: The Early Impacts of Building Strong Families," Mathematica Policy Research, May 2010, http://www.acf.hhs.gov/sites/default/files/opre/15_impact_main_rpt.pdf (accessed February 26, 2014); Robert G. Wood et al., *The Long-Term Effects of Building Strong Families: A Relationship Skills Education Program for Unmarried Parents: Final Report*, U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation, *OPRE Report No. 2012-28A*, November 2012, http://www.acf.hhs.gov/sites/default/files/opre/bsf_36_mo_impact_report.pdf (accessed February 26, 2014); and JoAnn Hsueh et al., *The Supporting Healthy Marriage Evaluation: Early Impacts on Low-Income Families*, U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation, *OPRE Report No. 2012-11*, February 2012, http://www.acf.hhs.gov/sites/default/files/opre/early_impacts_low.pdf (accessed February 26, 2014).

42. Michael J. Puma and Nancy R. Burstein, "The National Evaluation of the Food Stamp Employment and Training Program," *Journal of Policy Analysis and Management*, Vol. 13, No. 2 (Spring 1994), pp. 311-330.

43. Richard Hendra et al., *How Effective Are Different Approaches Aiming to Increase Employment Retention and Advancement? Final Impacts for Twelve Models*, MDRC, April 2010, http://www.mdrc.org/sites/default/files/full_390.pdf (accessed February 26, 2014).

children consistently failed to produce statistically meaningful results.⁴⁴

Federal Social Programs for Workers

This section covers federal social programs that are intended to boost the job skills and employability of workers. The federal government has spent decades trying to improve the earnings of low-income individuals through various employment and training programs, but the Government Accountability Office has concluded that there is little evidence to show that these programs are effective.⁴⁵

Adults and Youth. Conducted in 16 sites across the nation during the late 1980s and early 1990s, the Job Training Partnership Act evaluation tracked program effects for more than 20,000 adult men, adult women, and out-of-school youths over the course of 30 months.⁴⁶ Overall, the performance of JTPA programs is widely considered to be a failure. While adult females had several positive outcomes, the results were generally not large enough to be considered very meaningful. Further, JTPA programs were largely ineffective in raising the incomes of adult males or male and female youths.

Adults. To assess the effectiveness of self-employment training programs for those receiving Unemployment Insurance benefits, the federal government sponsored two experimental impact evaluations of programs in Washington State and Massachusetts.⁴⁷ Both programs had some very small and fleeting effects on increasing self-employ-

ment, but they were largely ineffective at raising the earnings of participants.

To help Americans start new businesses, the federal government established an employment program to assist people in creating or expanding their own business enterprises.⁴⁸ Begun in 2003, Project GATE operated in Pennsylvania, Minnesota, and Maine. Overall, Project GATE appears to have an initial impact on business ownership and self-employment, but it quickly faded away. Most important, Project GATE failed to increase the self-employment earnings of participants while temporarily reducing their total earnings.⁴⁹

Youth. Three multisite experimental evaluations of employment and training programs that specifically target youth strongly suggest that these programs are poorly serving this population.

Created in 1964, the Job Corps is a residential job-training program that serves disadvantaged youth in 125 sites across the nation. An experimental evaluation found the program to be ineffective.⁵⁰ Over the course of the 48-month study, Job Corps participants actually worked less than the control group. The study also revealed that the Job Corps had little impact on increasing the number of hours worked per week. If the Job Corps is effective in improving the skills of its participants, then it should have substantially raised the hourly wages they received. Initially, Job Corps participants earned an average of \$0.24 more per hour than non-participants, but this difference decreased to \$0.22 per hour after one year.

-
44. Larry Orr et al., *Moving to Opportunity Interim Impacts Evaluation: Final Report*, U.S. Department of Housing and Urban Development, Office of Policy Development and Research, June 2003, http://www.abtassociates.com/reports/2003302754569_71451.pdf (accessed February 26, 2014), and Lisa Sanbonmatsu et al., *Moving to Opportunity for Fair Housing Demonstration Program: Final Impacts Evaluation*, U.S. Department of Housing and Urban Development, Office of Policy Development and Research, November 2011, <http://www.huduser.org/portal/publications/pubasst/MTOFHD.html> (accessed February 26, 2014).
 45. U.S. Government Accountability Office, *Multiple Employment and Training Programs: Providing Information on Collocating Services and Consolidating Administrative Structures Could Promote Efficiencies*, GAO-11-92, January 2011, <http://www.gao.gov/new.items/d1192.pdf> (accessed February 26, 2014).
 46. Larry L. Orr et al., *Does Training for the Disadvantaged Work?* (Washington, DC: Urban Institute Press, 1996).
 47. Jacob M. Benus et al., "Self-Employment Programs: A New Reemployment Strategy: Final Impact Analysis of the Washington and Massachusetts Self-Employment Demonstrations," U.S. Department of Labor, Employment and Training Administration *Unemployment Insurance Occasional Paper* No. 95-4, December 1995, http://workforsecurity.doleta.gov/dmstree/op/op95/op_04-95.pdf (accessed February 26, 2014).
 48. Jeanne Bellotti, Sheena McConnell, and Jacob Benus, *Growing America Through Entrepreneurship: Interim Report*, Impaq International, August 2006, p. I, http://wdr.doleta.gov/research/keyword.cfm?fuseaction=dsp_resultDetails&pub_id=2337&mp=y (accessed February 26, 2014).
 49. Jacob Benus et al., *Growing America Through Entrepreneurship: Final Evaluation of Project GATE*, Impaq International, December 2009, <http://www.impaqint.com/sites/default/files/project-reports/Project%20GATE%20Final%20Report%20revised%20April%202010.pdf> (accessed February 26, 2014).
 50. Peter Z. Schochet, John Burghardt, and Steven Glazerman, *National Job Corps Study: The Impacts of Job Corps on Participants' Employment and Related Outcomes*, Mathematica Policy Research, June 2001, <http://wdr.doleta.gov/opr/fulltext/01-jcimpacts.pdf> (accessed February 26, 2014).

The JOBSTART Demonstration evaluated the impact of 13 job-training programs that were offered by community-based organizations, schools, and the Job Corps.⁵¹ Overall, the programs failed to increase the earnings of participants. Of the 13 sites, 12 were found to be ineffective at raising the incomes of participants.⁵² However, one site—the Center for Employment Training (CET) in San Jose, California—had a positive impact on earnings. Thus, at a single site, the CET appears to have been effective at raising the incomes of participants. For policymakers, the important question is whether these results can be replicated at different sites and for different populations.

Despite the success of welfare-to-work strategies, other social programs focusing on families that have undergone multisite experimental evaluations have been found to be ineffective.

Based on the JOBSTART evaluation results for the CET, the U.S. Department of Labor sought to replicate the program at 16 other sites.⁵³ However, only 12 of these sites were evaluated. The CET model had little to no effect on short-term and long-term employment and earnings outcomes at these other locations. The multisite experimental evaluation of the CET, according to its authors, “shows, that even in sites that best implemented the model, CET had no overall employment and earnings effects for youth in the program, even though it increased participants’ hours of training and receipt of credentials.”⁵⁴

Similarly, the Quantum Opportunity Program demonstration, which offered intensive and compre-

hensive services with the intention of helping at-risk youth graduate from high school and enroll in post-secondary education or training, failed to increase the employability or earnings of youth.⁵⁵

Conclusion

Do federal social programs work? Based on the scientifically rigorous multisite experimental evaluations published since 1990, the answer certainly cannot be in the affirmative. Despite the best social engineering efforts, overwhelming evidence points to the conclusion that federal social programs are ineffective.

Ameliorating such problems as low academic achievement, poor cognitive ability, poverty, joblessness, low wages, and personal relations appears to be out of reach for federal social programs. The most notable exception is welfare-to-work programs, which increased earnings, but participants still received some government assistance.

The evidence clearly shows that federal social programs are ineffective. It cannot be just a coincidence that the many multisite evaluations published since 1990 overwhelmingly find that this is true. Our nation faces a severe debt crisis that threatens our very future. Americans should not fear eliminating social programs. Now is the time for deep budget cuts in federal social programs.

The social programs that Congress continues to fund need to undergo large-scale experimental evaluations. Multisite experimental evaluations are the best method for assessing the effectiveness of federal social programs. Yet to date, this method has been used to evaluate only a handful of federal social programs. Congress needs to reverse this trend.

—*David B. Muhlhausen, PhD, is Research Fellow in Empirical Policy Analysis in the Center for Data Analysis at The Heritage Foundation.*

51. George Cave et al., *JOBSTART: Final Report on a Program for School Dropouts*, Manpower Demonstration Research Corporation, October 1993, p. xvii, http://www.mdrc.org/sites/default/files/full_416.pdf (accessed February 26, 2014).

52. *Ibid.*, p. 158, Table 5.13.

53. Cynthia Miller et al., *The Challenge of Replicating Success in a Changing World: Final Report on the Center for Employment Training Replication Sites*, Manpower Demonstration Research Corporation, September 2005, http://www.mdrc.org/sites/default/files/full_530.pdf (accessed February 26, 2014).

54. *Ibid.*, p. xi.

55. Allen Schirm, Elizabeth Stuart, and Allison McKie, *The Quantum Opportunity Program Demonstration: Final Impacts*, Mathematica Policy Research, July 2006, <http://www.mathematica-mpr.org/publications/PDFs/QOPfinalimpacts.pdf> (accessed February 26, 2014).