



214 Massachusetts Avenue, NE • Washington DC 20002 • (202) 546-4400 • heritage.org

CONGRESSIONAL TESTIMONY

Evidence-Based Policymaking: An Idea Whose Time Has Come

**Testimony before
the Committee on Ways and Means
Subcommittee on Human Resources
United States House of Representatives**

March 17, 2015

**David B. Muhlhausen, PhD
Research Fellow in Empirical Policy Analysis
The Heritage Foundation**

My name is David Muhlhausen. I am a Research Fellow in Empirical Policy Analysis in the Center for Data Analysis at The Heritage Foundation. I thank Chairman Charles Boustany, Ranking Member Lloyd Doggett, and the rest of the subcommittee for the opportunity to testify today on evidence-based policymaking. The views I express in this testimony are my own and should not be construed as representing any official position of The Heritage Foundation.

The federal government's total debt is over \$18.1 trillion.¹ Given the fiscal crises that the federal government is facing, holding federal social programs accountable for their performance is necessary to regain control over excessive spending. Operating with scarce resources, federal policymakers need to fund programs that work and defund programs that do not work. Americans, especially income tax payers, deserve better than Congress's current habit of continuing to spend taxpayer dollars on programs that do not produce their intended results.

The effectiveness of federal programs is often unknown. Many programs operate for decades without ever undergoing thorough scientific evaluations. In *Do Federal Social Programs Work?*, I reviewed 20 scientifically rigorous multisite experimental evaluations of federal social programs published since 1990.² Except for the welfare-to-work programs, federal social programs have been consistently found to be ineffective.

Evidence-Based Policymaking

To plug this information gap, the evidenced-based policy movement seeks to inform and influence policymakers through scientifically rigorous evaluations of the effectiveness of government programs.³ In other words, the movement provides tools to figure out what works and what does not work.

Evidence-based policymaking is based upon using scientifically rigorous impact evaluations to improve policy decisions. Rigorous impact evaluations that use random assignment provide policymakers improved capability to exercise oversight of government programs and be more effective stewards of the federal purse. There is little merit in continuing programs that fail to ameliorate their targeted social problems. Programs that are unknown to work or that do not work at all do not deserve continued funding.

Congress needs to take the lead in making sure that the social programs it funds are evaluated. First, when authorizing a new social program or reauthorizing an existing program, Congress should specifically mandate multisite experimental (random assignment) evaluation of the program.

Experimental evaluations are the only way to determine to a high degree of certainty the effectiveness of social programs. Thus, Congress should mandate that all recipients of federal funding, if selected for participation, must cooperate with evaluations in order to receive future funding.

Second, the experimental evaluations should be large-scale, nationally representative, multisite studies. When Congress creates social programs, the funded activities are intended to be spread out across the nation. For this reason, Congress should require nationally representative, multisite experimental evaluations of these programs. For multisite evaluations, the selection of the sites to be evaluated should be representative of the population of interest for the program. When program sites and sample participants are randomly selected, the resulting

evaluation findings will have high external validity.

The Problem of Replication and Scaling-Up

Many advocates of social programs have adopted the language of the “evidence-based” policy movement. Under the evidence-based policy movement, programs found to be effective using rigorous scientific methods are deemed “effective” or “evidence-based” and held up as “model” programs. The assumption is that the same successful impacts found at a particular setting can be replicated in other settings or on the national scale.

This faulty reasoning is based upon the “single-instance fallacy.”⁴ This fallacy occurs when a person believes that a small-scale social program that appears to work in one instance will yield the same results when replicated elsewhere. Compounding the effects of this fallacy, we often do not truly know why an apparently effective program worked in the first place. So how can we replicate it?

An excellent example of a federal government attempted replication of an effective local program is the Center for Employment Training (CET) Replication.⁵ Of 13 youth job-training programs evaluated, the JOBSTART Demonstration found only one program to have a positive impact on earnings—the CET in San Jose, California.⁶ Based on the results for the CET, the U.S. Department of Labor replicated and evaluated the impact of the CET in 12 other sites using random assignment.⁷ The CET model had little to no effect on short-term and long-term employment and earnings outcomes at these other locations. According to the evaluation’s authors, “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.”⁸

Just because an innovative program appears to have worked in one location does not mean that the program can be effectively implemented on a larger scale. Proponents of evidence-based policymaking should not automatically assume that allocating taxpayer dollars towards programs attempting to replicate previous successful findings will yield the same results.

Spending and Intentions Do Not Equal Success

Far too frequently, the amount of money spent to alleviate social problems and the good intentions of the social program advocates are considered measures of success. Instead, the actual degree to which social problems are reduced should be the measure of success. While continually spending taxpayer dollars on government programs may symbolize the compassion of program advocates, it does not mean that actual social problems are being alleviated.

Intentions are often confused with results. This wide-ranging problem is especially relevant to early childhood education programs. These programs are automatically assumed by advocates to level the playing field by helping disadvantaged children arrive at school without learning deficits. From time to time, an early childhood education program will appear to work. When a particular innovative early childhood education program seems to produce compelling evidence of success, policymakers and advocates of government social programs around the country appropriately take notice.

Such is the case with the High/Scope Perry Preschool and the Carolina Abecedarian Projects—two small-scale, highly intensive early education programs that served minority children. Based on the experiences of 58 preschoolers and 65 children not granted access to preschool, University of Chicago economist James Heckman and his team of researchers estimate that the Perry program produced \$7 to \$12 in long-term societal benefits for every dollar invested. The major benefit of the program is derived from reduced crime.⁹ The Abecedarian Project, and its study of 111 children, was found to have a long-term effect on increased educational attainment with the treatment and control groups averaging 13.46 years and 12.31 years of education, respectively.¹⁰ However, the Abecedarian Project had no measurable impact on income or criminal convictions.

Based on Heckman's research, President Barack Obama during his 2013 State of the Union Address made the broad generalization that "[e]very dollar we invest in high-quality early childhood education can save more than seven dollars later on—by boosting graduation rates, reducing teen pregnancy, even reducing violent crime."¹¹ In his fiscal year (FY) 2015 budget proposal, Obama states: "Research shows that one of the best investments we can make in a child's life is high-quality early education. This year, we will invest in new partnerships with States and communities across the country to expand access to high-quality early education, and I am again calling on the Congress to make high-quality preschool available to every four-year-old child."¹² This year President Obama released his FY 2016 budget recommendations that included a huge expansion in early childhood education programs.¹³

The President's proposal is well-meaning, but is based upon the single-instance fallacy.

There are good reasons to question the assumption that the federal government can replicate the beneficial outcomes purported to have been caused by the Perry and Abecedarian Projects. Ignoring the fact that these studies are not based upon well-implemented random-assignment studies, the evaluations of these small-scale programs are outdated. And despite all the hoopla, the results have never been replicated. In more than 50 years, not a single experimental evaluation of the Perry approach applied in another setting or on a larger scale has produced the same results. The same holds true for the Abecedarian program, which began in 1972.

Simply put, there is no evidence that these programs can produce the same results today. If we really knew how these programs actually produced success, would not these results have been replicated elsewhere?

In addition, the federal government has a poor track record of replicating successful programs on a national scale.¹⁴ This point is almost never raised by advocates of expanding the federal government's involvement in early childhood education programs. The Perry and Abecedarian programs are not realistic models to draw conclusions about the effectiveness of federal early childhood education programs.

And here is the problem. With no scientific certainty, advocates of expanding the federal role in early childhood education programs cannot answer the following question: Will increased federal spending on early childhood education programs improve children's futures? Instead, the decision to favor a federal expansion of preschool learning opportunities is most often based on the answer to a less scientifically rigorous question: Will proposing increased federal spending

on early childhood programs make advocates feel that they are making a difference in the lives of children?

The answer to the latter, simpler question is almost certainly yes. Unfortunately, this faulty decision-making process often results in federal boondoggles like Head Start and its sibling Early Head Start. Just consider what we really know about Head Start and Early Head Start.

Early Head Start. 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. The results of the multisite experimental evaluation of Early Head Start are particularly important because the program was inspired by the findings of the Abecedarian Project.¹⁵ By the time participants reached age three, Early Head Start had beneficial impacts on two out of six outcome measures for child cognitive and language development, while the program had beneficial effects on four out of nine measures of child-social-emotional development.¹⁶ While the short-term (age three) findings indicated modest positive impacts, almost all of the positive findings for all Early Head Start participants were driven by the positive findings for black children. 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.¹⁷ For the 11 child-social-emotional outcomes, none of the results were found to have statistically meaningful impacts. Further, Early Head Start failed to have statistically measurable effects on the 10 measures of child academic outcomes, including reading, vocabulary, and math skills.

Head Start. Created as part of the War on Poverty in 1965, Head Start is a preschool community-based program intended to help disadvantaged children catch up to children living in more fortunate circumstances. 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.¹⁸ According to *CQ News*, the 2012 study "revealed that children who attended Head Start had lost most of its benefits by the time they reached third grade."¹⁹ This assessment is entirely wrong. Almost all of the benefits of participating in Head Start disappeared by kindergarten.

Overall, the evaluation found that the program largely failed to improve the cognitive, socio-emotional, health, and parenting outcomes of children in kindergarten and first grade who participated compared with the outcomes of similar children who did not participate. By third grade, Head Start had little to no effect on cognitive, social-emotional, health, or parenting outcomes of participating children.²⁰

In addition to the failures of Early Head Start and Head Start, multisite experimental evaluations of the Enhanced Early Head Start with Employment Services, which provides early childhood care and employment training services to families, and the now-defunct Even Start Family Literacy Program, which was intended to meet the basic educational needs of parents and children, failed to produce beneficial impacts.²¹ The scientific rigor of these evaluations clearly demonstrates that the federal government has serious trouble operating early childhood education programs. These programs have done a poor job of improving the cognitive abilities and socio-

emotional development of children.

Causing Harm

The results of the multisite experimental evaluations of federal social programs generally find that these programs are ineffective.²² However, social program advocates too frequently concentrate on any beneficial, even if only modest, impacts that have been identified. Nevertheless, politicians and policy experts also need to recognize that federal social programs can produce harmful impacts too. These harmful effects are rarely mentioned in government press releases announcing the findings of evaluations. While all of the programs identified in this section as having produced some harmful impacts, all of these programs were created with the best of intentions. [Suggest: While all of the programs in this section are identified with producing some harmful impacts...]

For Early Head Start, white parents in the intervention group displayed higher dysfunctional parent-child interactions than their counterparts in the control group.²³ Further, participation in Early Head Start appears to have increased welfare dependency for Hispanics.

Enhanced Early Head Start with Employment Services is a demonstration program that involves regular Early Head Start services with the addition of employment and training services for parents. An experimental evaluation of the program based on two sites in Kansas and Missouri was performed. At the time of the 48-month follow-up, the longest job spells of mothers participating in the program were significantly shorter than the job spells of mothers in the control group.²⁴

For the three-year-old cohort of the Head Start Impact Study, kindergarten teachers reported that math abilities were worse than for similar children not given access to the program.²⁵ For the four-year-old cohort, teachers reported that Head Start children in the first grade were more likely to be shy or socially reticent than their peers. By the third grade, teachers reported that the four-year-old cohort with access to Head Start displayed a higher degree of unfavorable emotional symptoms than similar children without access to the program.²⁶ Further, children in the four-year-old cohort self-reported poorer peer relations with fellow children than their counterparts in the control group.²⁷

The role of the federal government in funding after-school programs increased substantially after passage of the Improving America's School Act of 1994, which created the 21st Century Community Learning Centers program. A multisite experimental impact evaluation of the 21st Century Community Learning Centers program found a whole host of harmful effects.²⁸ Overall, teachers found participating students to have disciplinary problems that were confirmed by student-reported data. According to their teachers, participating students were less likely to achieve at above average or high levels in class and were less likely to put effort into reading or English classes. These students were also more likely to have behavior problems in school than their counterparts. Teachers were more likely to have to call the parents of participating students about misbehavior. Participating students were also more likely to miss recess or be placed in the hall for disciplinary reasons, while also having parents come to school more often to address behavior problems. 21st Century students were also more likely to be suspended from school than similar students.

Upward Bound was created in 1965 and is an original War on Poverty social program. Through the provision of supplemental academic and support services and activities, Upward Bound is intended to help economically disadvantaged high school students successfully complete high school and attend college. Despite the program's lofty goal, Upward Bound participants with high expectations to earn a college degree were less likely than their counterparts to earn associate's degrees, while being no more or less likely to attain any other college degree.²⁹

The Department of Health and Human Services and Department of Labor funded the Employment Retention and Advancement (ERA) project, initiated in 1998, to assess the effectiveness of 12 different employment retention and advancement programs across the nation.³⁰ Participation in ERA programs targeting unemployed Temporary Assistance for Needy Families (TANF) recipients in Houston, Texas, and Salem, Oregon, was associated with increased dependence on the receipt of TANF benefits, while participation in the program in Fort Worth, Texas, was associated with increased dependence on food stamps. The Chicago ERA program targeting employed TANF recipients was associated with increased dependence on food stamps, while the Medford, Oregon, ERA program targeting employed individuals not on TANF was associated with decreased employment.

Conducted in five cities, the Moving to Opportunity (MTO) demonstration assessed the impact of offering families with children under 18 living in public housing developments or concentrated poverty areas the opportunity to move out of their neighborhoods. The evaluation consisted of two intervention groups, MTO voucher recipients and Section 8 voucher recipients, compared to a control group that did not receive MTO or Section 8 vouchers but was eligible to receive public housing assistance. For adults and children with access to MTO or Section 8 vouchers, several harmful impacts were produced.³¹ Access to a MTO voucher was associated with increased dependence on drugs and alcohol for adults. Also, MTO adults had higher participation rates in food stamps and received more food stamp benefits than their similar counterparts not given access to MTO or Section 8 vouchers. Youth from families given access to MTO vouchers were less likely to be employed and more likely to have smoked than their peers. These youth were also more likely to be arrested for property crimes. As for Section 8, adults offered access were more likely to be currently unemployed and less likely to have employment spells with the same job for at least a year. In addition, Section 8 adults were less likely to be currently working and not receiving TANF than their counterparts. Section 8 youth were more likely to have smoked than their peers in the control group.

Adult men participating in Job Training Partnership Act programs were more likely to be dependent on Aid to Families with Dependent Children (AFDC) benefits than similar men not given access to the training.³² Male youths with no criminal arrest record at the time of random assignment were more likely to be arrested after participating in federal job-training programs, while male youth with histories of arrest experienced long-term declines in income.

In an attempt to help Americans start businesses, the Department of Labor teamed with the Small Business Administration to create an employment program to assist people in creating or expanding their own business enterprises.³³ After receiving entrepreneurship training, Project GATE participants spent more time collecting Unemployment Insurance benefits than their counterparts who were not taught how to be entrepreneurs. While Project GATE had no effect on the self-employment income of participants, participants experienced initial periods of decreased wages and salaries earned from overall employment.

The Quantum Opportunity Program (QOP) demonstration, operated by the U.S. Department of Labor and the Ford Foundation from 1995 to 2001, offered intensive and comprehensive services with the intention of helping at-risk youth graduate from high school and enroll in postsecondary education or training. QOP provided services to participants year-round for five years. The findings from the QOP experimental evaluation, according to its authors, provide some insight about the effectiveness of WIA youth programs. For the initial post-intervention impacts, youth participating in QOP were less likely to find jobs that provided health insurance benefits.³⁴ At the six-year follow-up period, youth participating in QOP were more likely to be arrested.³⁵ Increasing criminality appears to be a common effect of federal job-training programs supposedly benefiting youth.

The previously discussed CET Replication job-training programs were associated with several harmful outcomes.³⁶ Men experienced periods of declines in employment, earnings, and number of months worked. Individual participants who possessed a high school diploma or GED at the time of random assignment experienced periods of declines in the number of months worked and earnings. In addition, participants in the high-fidelity sites were less likely to find jobs that provided health insurance. Also, those older than 18 and those with high school degrees or GEDs at the time of random assignment were less likely to have jobs that provided health insurance.

Job Corps is another federal training program that has negative effects. Created in 1964, Job Corps is a residential job-training program that serves disadvantaged youths ages 16 to 24 in 125 sites across the nation. A multisite experimental evaluation of Job Corps found, compared to non-participants, Job Corp participants were less likely to earn a high school diploma.³⁷ In addition, youth participating in the program worked fewer weeks and worked fewer hours per week than similar youth in the control group.³⁸

In sum, federal social programs that harm their participants are not uncommon. This fact is all too often ignored by advocates of these social programs.

Conclusion

With the enormous federal debt increasingly shaping policy debates in Washington, DC, Congress should subject all federal programs to rigorous evaluations to determine what works and what does not work. The adoption of evidence-based policymaking is an important step in helping Congress become wise stewards of the federal purse. To assist in accomplishing this goal, Representative Paul Ryan (R-WI) and Senator Patty Murray (D-WA) introduced the Evidence-Based Policymaking Commission Act of 2014 (H.R. 5754) in the 113th Congress. Shortly, Representative Ryan and Senator Murray are expected to introduce a revised version of the bill for the 114th Congress.

Changing the federal government's emphasis on measuring success by the amount of spending and intentions will not be easy. However, the creation of an Evidence-Based Policymaking Commission would be a step in the right direction for changing the culture in Washington towards funding programs that work and defunding those that do not work.

The Heritage Foundation is a public policy, research, and educational organization recognized as exempt under section 501(c)(3) of the Internal Revenue Code. It is privately supported and receives no funds from any government at any level, nor does it perform any government or other contract work.

The Heritage Foundation is the most broadly supported think tank in the United States. During 2013, it had nearly 600,000 individual, foundation, and corporate supporters representing every state in the U.S. Its 2013 income came from the following sources:

Individuals	80%
Foundations	17%
Corporations	3%

The top five corporate givers provided The Heritage Foundation with 2% of its 2013 income. The Heritage Foundation's books are audited annually by the national accounting firm of McGladrey, LLP.

Members of The Heritage Foundation staff testify as individuals discussing their own independent research. The views expressed are their own and do not reflect an institutional position for The Heritage Foundation or its board of trustees.

¹U.S. Department of the Treasury, "The Debt to the Penny and Who Holds It," <http://treasurydirect.gov/NP/debt/current> (accessed March 9, 2015).

²David B. Muhlhausen, *Do Federal Social Programs Work?* (Santa Barbara, CA: Praeger, 2013).

³See Karen Bogenschneider and Thomas J. Corbett, *Evidence-Based Policymaking: Insights from Policy-Minded Researchers and Research-Minded Policymakers* (New York: Routledge, 2010).

⁴Stuart M. Butler and David B. Muhlhausen, "Can Government Replicate Success?" *National Affairs*, No. 19 (Spring 2014), pp. 25–39, <http://www.nationalaffairs.com/publications/detail/can-government-replicate-success> (accessed March 9, 2015).

⁵Cynthia Miller, Johannes M. Ros, Kristen E. Porter, Fannie M. Tseng, and Yasuyo Abe, *The Challenge of Replicating Success in a Changing World: Final Report on the Center for Employment Training Replication Cites* (Manpower Demonstration Research Corporation, September 2005).

⁶George Cave, Hans Bos, Fred Doolittle, and Cyril Toussaint, *JOBSTART: Final Report on a Program for School Dropouts* (Manpower Demonstration Research Corporation, October 1993).

⁷Miller et al., *The Challenge of Replicating Success in a Changing World*.

⁸*Ibid.*, p. xi.

⁹James J. Heckman, Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz, "The Rate of Return to the HighScope Perry Preschool Program," *Journal of Public Economics*, Vol. 94 (2010), pp. 114–128.

¹⁰Frances A. Campbell, Elizabeth P. Pungello, Margaret Burchinal, Kirsten Kainz, Yi Pan, Barbara H. Wasik, Oscar A. Barbarin, Joseph J. Sparling, and Craig T. Ramey, "Adult Outcomes as a Function of an Early Childhood Educational Program: An Abecedarian Project Follow-Up," *Developmental Psychology*, Vol. 48, No. 4 (2012), pp. 1033–1043.

¹¹Barack Obama, "Remarks by the President in the State of the Union Address," The White House, Office of the Press Secretary, February 12, 2013, <http://www.whitehouse.gov/the-press-office/2013/02/12/remarks-president-state-union-address> (accessed June 18, 2013).

¹²Office of Management and Budget, *Fiscal Year 2015 Budget of the U.S. Government*, p. 3, <http://www.gpo.gov/fdsys/pkg/BUDGET-2015-BUD/pdf/BUDGET-2015-BUD.pdf> (accessed March 9, 2015).

-
- ¹³Office of Management and Budget, *The President's Budget for Fiscal Year 2016*, The White House, <http://www.whitehouse.gov/omb/budget> (accessed March 9, 2015).
- ¹⁴Butler and Muhlhausen, "Can Government Replicate Success?"
- ¹⁵Geoffrey D. Borman, "National Efforts to Bring Reform to Scale in High-Poverty Schools: Outcomes and Implications," in Barbara Scheider and Sarah-Kathryn McDonald, eds., *Scaled-Up in Education: Issues in Practice, Vol. II* (Lanham, MD: Rowman & Littlefield, Inc., 2007), pp. 41–67.
- ¹⁶John M. Love, Ellen Eliason Kisker, Christine M. Ross, Peter Z. Schochet, Jeanne Brooks-Gun, Diane Paulsell, Kimberly Boller, Jill Constantine, Cheri Vogel, Allison Sidle Fulongi, and Christi Brady-Smith, *Making a Difference in the Lives of Infants and Toddlers and Their Families: The Impacts of Early Head Start, Volume 1: Final Technical Report* (Princeton, NJ: Mathematica Policy Research, June 2002).
- ¹⁷Cheri A. Vogel, Yange Xue, Emily M. Moiduddin, Barbara Lepidus Carlson, and Ellen Eliason Kisker, *Early Head Start Children in Grade 5: Long-Term Follow-Up of the Early Head Start Research Evaluation Project Study Sample: Final Report*, OPRE Report No. 2011–8 (Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, December 2010).
- ¹⁸U.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; U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation, *Head Start Impact Study: Final Report*, January 2010; and Mike Puma, Stephen Bell, Ronna Cook, Camilla Heid, Pam Broene, Frank Jenkins, Andrew Mashburn, and Jason Downer, *Third Grade Follow-Up to the Head Start Impact Study: Final Report*, OPRE Report 2012–45 (Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, October 2012).
- ¹⁹*CQ News*, "Obama Details Plan for Expanding Preschool Education," February 14, 2013.
- ²⁰Puma et al., *Third Grade Follow-Up to the Head Start Impact Study: Final Report*.
- ²¹JoAnn 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*, OPRE Report No. 2012–05 (Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, February 2012); U.S. Department of Education, *Third National Even Start Evaluation: Program Impacts and Implications for Improvement*, 2003; and Anna E. Ricciuti, Robert G. St. Pierre, Wang Lee, Amanda Parsad, and Tracy Rimdzius, *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).
- ²²Muhlhausen, *Do Federal Social Programs Work?*
- ²³Love et al., *Making a Difference in the Lives of Infants and Toddlers and Their Families*, Table VII.11, pp. 381–385.
- ²⁴Hsueh and 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*, Table 3.2, pp. 36–37.
- ²⁵U.S. Department of Health and Human Services, *Head Start Impact Study: Final Report*, Exhibit 4.5, pp. 4–21–4–25.
- ²⁶Puma et al., *Third Grade Follow-up to the Head Start Impact Study Final Report*, p. 84.
- ²⁷*Ibid.*, Exhibit 4.3, pp. 81–82.
- ²⁸Susanne 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.
- ²⁹Neil S. Sefor, Arif Mamun, and Allen Schirm, *The Impacts of Regular Upward Bound on Postsecondary Outcomes 7–9 Years after Scheduled High School Graduation: Final Report* (Princeton, NJ: Mathematica Policy Research, January 2009), Table IV.2, p. 59.
- ³⁰Richard Hendra, Keri-Nicole Dillman, Gayle Hamilton, Erik Lundquist, Karin Martinson, Melissa Wavelet, Aaron Hill, and Sonya Williams, *How Effective Are Different Approaches Aiming to Increase Employment Retention and Advancement? Final Impacts for Twelve Models*, MDRC, April 2010.
- ³¹Lisa Sanbonmatsu, Jens Ludwig, Lawrence F. Katz, Lisa Gennetian, Greg J. Duncan, Ronald C. Kessler, Emma Adam, Thomas W. McDade, Stacy Tessler Lindau, Matthew Sciandra, Fanghua Yang, Ijun Lai, William Congdon, Joe Amick, Ryan Gillette, Michael A. Zabek, Jordon Marvakov, Sabrina Yusuf, and Nicholas A. Potter, *Moving to*

Opportunity for Fair Housing Demonstration Program: Final Impacts Evaluation (Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research, November 2011).

³²Larry L. Orr, Howard S. Bloom, Stephen H. Bell, Fred Doolittle, Winston Lin, and George Cave, *Does Training for the Disadvantaged Work?* (Washington, DC: Urban Institute Press, 1996).

³³Jeanne Bellotti, Sheena McConnell, and Jacob Benus, "Growing America through Entrepreneurship: Interim Report," Impaq International, August 2006, and Jacob Benus, Theodore Shen, Sisi Zhang, Marc Chan, and Benjamin Hansen, *Growing America through Entrepreneurship: Final Evaluation of Project GATE* (Columbia, MD: Impaq International, December 2009).

³⁴Allen Schirm and Nuria Rodriguez, *The Quantum Opportunity Program Demonstration: Initial Post Intervention Impacts* (Mathematica Policy Research, June 2004).

³⁵Allen Schirm, Elizabeth Stuart, and Allison McKie, *The Quantum Opportunity Program Demonstration: Final Impacts* (Mathematica Policy Research, July 2006).

³⁶Miller et al., *The Challenge of Replicating Success in a Changing World*.

³⁷Peter Z. Schochet, John Burghardt, and Sheena McConnell, "Does Job Corps Work? Impact Findings from the National Job Corps Study," *American Economic Review*, Vol. 98, Issue 5, pp. 1864–1886.

³⁸Peter Z. Schochet, John Burghardt, and Steven Glazerman, *National Job Corps Study: The Impacts of Job Corps on Participants' Employment and Related Outcomes* (Princeton, NJ: Mathematica Policy Research, Inc., June 2001).