

Pay-for-Outcomes: Transforming Federal Social Programs to Expand Individual Well-Being

Leslie Ford and Robert Rector

KEY TAKEAWAYS

Federal social programs spend billions but fail to show that they meaningfully help at-risk Americans. It is time to change how these programs pay for services.

Instead of using ineffective payment models, federal social programs should reward providers for achieving verifiable outcomes that change lives for the better.

Paying directly for outcomes that show clear improvements for vulnerable Americans respects their dignity in society while using taxpayer dollars more effectively.

Every year, the federal government spends billions of dollars on social programs that are intended to help at-risk individuals, who are vulnerable to poverty and dependence on government safety-net programs, achieve self-sufficiency. Leading examples include employment and training programs (\$18.9 billion in federal spending in fiscal year 2019); substance abuse treatment programs (\$10.5 billion in FY 2017); and recidivism reduction programs (\$68 million in FY 2015).

These and other social programs are intended to increase lasting well-being through the acquisition of education, skills, and other positive changes in behavior. They promise concrete change: for the unemployed, obtaining a job; for former inmates, lower rates of reincarceration; and for those with addictions, treatment that will enable them to live free from substance

This paper, in its entirety, can be found at <http://report.heritage.org/bg3550>

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.

dependence. In contrast to other federal programs that offer transfer payments (e.g., cash, food, or housing assistance), these social programs are premised on helping individuals to build and maintain a meaningful and free life.

Unfortunately, most of these social programs have little to show despite steep expenditures. Many—particularly the “open grant model,” which is usually one of providing federal block grants to states for social programs—do not track whether federal spending produces key outcomes such as greater employment, reduced inmate recidivism, successful substance abuse treatment, and improved self-sufficiency. Even the data available from legislative attempts to manage program performance and distribute funds to “evidenced-based” services over the past two decades do not demonstrate that those efforts produce overall meaningful outcomes.

Meanwhile, “pay-for-success” efforts during the past decade, while showing some promise, remain largely outside the federal government’s main expenditures on social programs. Moreover, these efforts can be judged primarily by reductions in federal government expenditures, not direct impacts in building recipients’ capacity for self-sufficiency.

Overall, these federal social programs are not demonstrating that they achieve the outcomes for which they were designed. This both disservices recipients and shortchanges taxpayers. It is time for a transformation of federal investments in social programs, and this requires a new funding model. Federal policymakers should pay directly for outcomes in social programs in a way that truly respects the dignity of vulnerable individuals by expanding their capacity and adding to their well-being.

Summary of Current Federal Social Program Payment Models

Three principal methods are currently used by federal policymakers to distribute funds to social programs that seek to increase capacities among vulnerable recipients:

- The open grant model,
- The performance management model, and
- The social impact bond or pay-for-success model.

Each of these models has shortcomings. To achieve the goals of capacity-building among recipients, a new funding model is required.

This paper proposes a new payment model to bring about transformation in these social programs: the pay-for-outcome model. Before describing this proposed model, however, a review of the current models is provided.

The Open Grant Model

The open grant model characterizes many federal programs initiated as part of the 1960s Great Society as well as others that have proliferated since then. Federal executive agencies allocate these funds to states or directly to service providers according to the intended goals set by congressional legislation. One of the aspirations of the open grant model is to allow states or direct service providers to innovate in social programs based on distinctive local circumstances or the unique needs of the vulnerable individual.

Unfortunately, the vast majority of these initiatives do not produce convincing evidence that the programs work for their intended beneficiaries. There is often no serious data tracking or evidence that the outcomes for beneficiaries are achieved. As summarized by Yale University Law Professor Peter Schuck in 2014, “less than 1 percent of government spending is backed by even the most basic evidence of cost-effectiveness.”¹ This use of federal funds is fundamentally flawed, representing both a misuse of taxpayer funds and a failure to keep the promises made to the vulnerable individuals whom federal policymakers seek to aid.

Three large federal block grants intended to reduce poverty exemplify the problems inherent in the open grant model. These include the Social Services Block Grant (SSBG); the Community Services Block Grant (CSBG); and the Community Development Block Grant (CDBG).

- The SSBG is a \$1.7 billion-per-year federal block grant to states with five broad statutory goals intended to move individuals to self-sufficiency and eliminate dependence.² States must identify their intended uses of the funds (over a quarter of which is spent on child protective services and foster care),³ but there is no consistent collection of

1. Peter H. Schuck, *Why Government Fails So Often: And How It Can Do Better* (Princeton, NJ: Princeton University Press, 2014), p. 391.

2. U.S. Department of Health and Human Services, Administration for Children and Families, Office of Community Services, “SSBG Legislative Authority Summary,” last reviewed June 10, 2019, <https://www.acf.hhs.gov/ocs/resource/legislative-authority> (accessed October 2, 2020).

3. Karen E. Lynch, “Social Services Block Grant,” Congressional Research Service *In Focus* No. 10115, updated November 1, 2018, <https://fas.org/sgp/crs/misc/IF10115.pdf> (accessed September 15, 2020).

outcome data that shows whether states achieve the SSBG program's intended goals.

- The CSBG distributes nearly \$700 million in federal funds annually to more than 1,000 local antipoverty programs through a state formula. The programs are duplicative, poorly targeted, and not funded on the basis of their achievement of measured performance outcomes. In 2019, the U.S. Government Accountability Office performed its most recent survey and audit of the data that the U.S. Department of Health and Human Services (HHS) collects on the efficacy of the program and concluded that there is little measurable evidence that this significant funding source achieves the goals that Congress intended it to achieve.⁴
- Operated by Department of Housing and Urban Development (HUD), the federal CDBG program spends \$3 billion per year, \$900 million of which HUD allocates to states. The CDBG was intended to provide housing assistance for low-income families, but its funds have often been funneled to high-income communities and to wasteful pork-barrel projects. For example, \$500,000 from the program was spent to expand a brewery in New York,⁵ and \$1 million was used to restore a Michigan hotel.⁶

When not distributed directly to states, open grant model funds are distributed through federal agencies. The agencies outline the type of recipients they seek to assist and the general aims or activities that the recipients promise to undertake. In many cases, the federal agency overseeing the open grant does not track whether the services were supplied. Even when administrative agencies collect data from grant recipients, the data collection is focused on tracking provision of services, not outcomes for the beneficiaries. Grant recipients do not have to specify the intended number of services they expect to provide or the number of beneficiaries that will be reached per federal dollar expended.

4. U.S. Government Accountability Office, "Community Services Block Grant: Better Alignment of Outcome Measures with Program Goals Could Help Assess National Effectiveness," GAO-20-25, November 2019, *passim*, <https://www.gao.gov/assets/710/702722.pdf> (accessed October 2, 2020).

5. "State Grant to Support Wilson Brewery's NT Expansion," *Lockport Union-Sun Journal*, August 24, 2018, https://www.lockportjournal.com/news/local_news/state-grant-to-support-wilson-brewerys-nt-expansion-plan/article_0e38eea0-6130-54b6-a360-781f6e449b1a.html (accessed October 2, 2019).

6. "Dilworth Hotel Project Gets \$1 Million Grant," *News-Review*, December 23, 2015, https://www.petoskeynews.com/featured-pnr/dilworth-hotel-project-gets-million-grant/article_9141283b-01f6-58d4-95ae-c2ea4ed1185e.html?mode=jqm (accessed October 2, 2020).

Because open grants are largely distributed based on stated goals without outcome data collection and analysis, this model is unable to produce evidence that it achieves intended outcomes in the lives of vulnerable individuals. The overall persistence of the problems these programs seek to address suggests that they do not achieve those ends.

The Performance Management Model

Recognizing that the open grant model has significant flaws, over the past decade and a half a significant reform movement tried to improve the effectiveness of federally funded social programs.⁷ Federal legislators used program reauthorizations to implement the performance management model, which can take the form of three key reforms: collecting data on service provision and program performance, mandating that services deployed be supported by evidence, and requiring third-party outcome evaluations.

One of the most significant examples of congressional reauthorization that aimed to collect data and obtain performance analysis is the Workforce Innovation and Opportunity Act (WIOA),⁸ enacted in 2014 to replace the Workforce Investment Act of 1998 (WIA).⁹ In this reform, Congress attempted to move federal employment and training (E&T) programs from the open grant model to an evidence-collection and performance management model. The WIOA created a common performance accountability system for six core E&T programs. Despite that, in a 2019 report on the overall effectiveness of these government training programs, the White House Council of Economic Advisers concluded that “government job training programs appear to be largely ineffective and fail to produce sufficient benefits for workers to justify the costs.”¹⁰

Another example of service data collection and outcome feedback based on available performance metrics is the Substance Abuse Prevention and Treatment (SAPT) Block Grant, which is intended to combat substance abuse.¹¹ The main source of data to track these funds is the Treatment

7. David Leonhardt, “The Quiet Movement to Make Government Fail Less Often,” *The New York Times*, July 15, 2014, <https://www.nytimes.com/2014/07/15/upshot/the-quiet-movement-to-make-government-fail-less-often.html?abt=0002&abg=0> (accessed September 15, 2020).

8. H.R. 803, Workforce Innovation and Opportunity Act, Public Law 113-128, 128 Stat. 1425, 113th Cong., July 22, 2014, <https://www.govinfo.gov/content/pkg/PLAW-113publ128/pdf/PLAW-113publ128.pdf> (accessed October 6, 2020).

9. H.R. 1385, Workforce Investment Act of 1998, Public Law 105-220, 112 Stat. 936, 105th Cong., August 7, 1998, <https://www.govinfo.gov/content/pkg/PLAW-105publ220/pdf/PLAW-105publ220.pdf> (accessed October 6, 2020).

10. Executive Office of the President of the United States, Council of Economic Advisers, *Government Employment and Training Programs: Assessing the Evidence on Their Performance*, June 2019, p. 2, <https://www.whitehouse.gov/wp-content/uploads/2019/06/Government-Employment-and-Training-Programs.pdf> (accessed October 2, 2020).

11. See 42 U.S.C. §300, <https://www.law.cornell.edu/uscode/text/42/300> (accessed October 6, 2020).

Episode Data Set (TEDS), a national data system of publicly funded substance abuse treatment facilities' admissions and discharges.¹² Again, this system attempts to track provision of services, not outcomes for individuals at risk of continuing addiction or, in the worst case, death through overdose. The Substance Abuse and Mental Health Services Administration (SAMHSA) within HHS also used the Performance Management Reporting Tool (PMRT) to review the effects of these funds, but "there is no empirical evidence that such feedback leads to an improvement in performance."¹³

A second key aspect of the performance management model is the requirement that direct service providers deliver "evidenced-based" services, meaning that the program should be able to provide only services that have been evaluated and shown to have made a positive, statistically significant difference for the beneficiary. Unfortunately, many "evidence-based" programs fail to produce outcomes as intended. Generally, "evidence-based" programs are approved after investigators test the effect of services at an original demonstration site.

The evidence-based paradigm involves a presumed linked chain leading from random control trials to a flourishing social program sector. In reality, however, each link in this chain is weak. Overall, the paradigm rests on the following sequence of questionable assumptions:

- In any given field, a small number of programs can be evaluated by random control trials (RCTs); among the small set of trials, a number of clearly successful programs will be found.
- These successful programs can be studied and then serve as "model" programs that can be replicated across the relevant social program sector.
- The "replicated" versions of the programs will perform as effectively as the original RCT-tested versions.
- Replicated models will readily displace other older programs.

12. U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, "Treatment Episode Data Set (TEDS)," <https://www.samhsa.gov/data/data-we-collect/teds-treatment-episode-data-set> (accessed October 6, 2020).

13. Lawrence A. Palinkas, Suzanne E. Spear, Sapna J. Mendon, Juan Villamar, Thomas Valente, Chi-Ping Chou, John Landsverk, Shepperd G. Kellam, and C. Hendricks Brown, "Measuring Sustainment of Prevention Programs and Initiatives: A Study Protocol," *Implementation Science*, Vol. 11, Issue 1 (December 2015), <https://link.springer.com/article/10.1186/s13012-016-0467-6> (accessed October 6, 2020).

- A desirable level of effective service delivery will thereby be established and maintained throughout the sector even though the operation of most individual providers will not be meaningfully measured or tested.

The evidence-based paradigm is inadequate for at least six reasons.

First, at any given time, there are thousands of individual social programs in operation in a wide variety of fields such as drug treatment, anti-recidivism, and employment services. How policymakers determine which services are evidenced-based is normally based on a single demonstration of results, which can be subject to “single-instance fallacy.”¹⁴ RCT experiments will probably examine no more than a dozen of these thousands of operations each year. Only a small fraction of those evaluated will be found to have significant effects, and those effects will often be modest. It is therefore unlikely that a handful of RCT evaluations can uncover many “hidden gems” within this myriad of programs.

Second, the performance management model also runs afoul of replication fallacy: a common assumption that if a program can be evaluated and proven to have results in the initial location, scaling up or replicating the original method in another location is both possible and advisable. Most RCT evaluations are like “black boxes.” Although some programs appear successful, it is often unclear why they succeeded, and the most important factors behind their success (such as a charismatic leader) are not measurable. Even positive studies do not guarantee that cloning those programs and replicating them throughout the country will prove to be successful.

Third, even when an evaluation documents a highly successful model program, history suggests that it is nearly impossible to replicate that model effectively in other locations. Where there is evidence of original single-instance effectiveness and subsequent studies of wider implementation, the broader program results often do not live up to the promise of

14. A single study is not normally considered scientifically significant evidence. According to Food and Nutrition Administration guidance, for example, “A single clinical experimental finding of efficacy, unsupported by other independent evidence, has not usually been considered adequate scientific support for a conclusion of effectiveness.” U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER), “Guidance for Industry: Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products,” May 1998, p. 4, <https://www.fda.gov/files/drugs/published/Providing-Clinical-Evidence-of-Effectiveness-for-Human-Drug-and-Biological-Products..pdf> (accessed October 6, 2020). This is because one positive outcome could be as good as random chance (one in 20). Unfortunately, in determining which services are “evidenced-based,” the policymaker looks at the evidence of one instance of services, which leaves the “evidence-based” services open to being statistical flukes. John P. A. Ioannidis, “Why Most Published Research Findings Are False,” *PLoS Medicine*, Vol. 2, No. 8 (August 2005), p. e124, <https://journals.plos.org/plosmedicine/article/file?id=10.1371/journal.pmed.0020124&type=printable> (accessed October 6, 2020).

the original demonstration.¹⁵ As Straight Talk on Evidence puts it, “even a well-conducted RCT with blockbuster findings...generally does not provide a sufficient basis for widespread program implementation; there is a good chance that the program would not produce the hoped-for effects if implemented in new jurisdictions.”¹⁶

Fourth, even in the unlikely event that a highly successful, replicable model is found, it is unlikely that a social program sector will readily “retool” to follow the model. Change is difficult, and replicating the success of one model suggests that many existing organizations, skills, and workers are unsuccessful and should be replaced or reformed. Bureaucratic inertia, historic ties, institutional ethos, and the instinct for organizational self-preservation all bolster the status quo. “Replicating a new model” often means little more than inserting a new set of buzzwords into the next grant application.

Fifth, even if the effort at replication is sincere, the “evidence-based” paradigm at best establishes broad programmatic guidance within a service sector. But the paradigm does not assess the thousands of individual service providers and therefore cannot separate “gems from lemons” at the specific provider level.

The Hawaii Opportunity Probation with Enforcement (HOPE) program, for example, had initial positive findings that did not carry over to the subsequent expansion. HOPE was a probation program that was found to have a significant effect on recidivism: Participants had a 55 percent reduction in rearrests in the 12 months following release.¹⁷ Based on this evidence, the program was quickly replicated and expanded to 28 locations in four other states (Arkansas, Massachusetts, Oregon, and Texas). However, the National Institute of Justice later commissioned another randomized controlled trial of the expansion programs and found far less effect.¹⁸

Another example is the Maternal, Infant, and Early Childhood Home Visiting (MIECHV) program for vulnerable mothers that has been widely acclaimed as evidence-based. Congress authorized the program in 2010,

15. Ioannidis, “Why Most Published Research Findings Are False,” p. e124.

16. Larry Orr, “If at First You Succeed, Try Again!” Straight Talk on Evidence, August 16, 2017, <https://www.straighttalkonevidence.org/2017/08/16/if-at-first-you-succeed-try-again/> (accessed October 6, 2020).

17. Angela Hawken and Mark Kleiman, “Managing Drug Involved Probationers with Swift and Certain Sanctions: Evaluating Hawaii’s HOPE,” submitted to the U.S. Department of Justice, December 2009, <https://www.ncjrs.gov/pdffiles1/nij/grants/229023.pdf> (accessed October 6, 2020), and Coalition for Evidence-Based Policy, “Hawaii’s Opportunity Probation with Enforcement (HOPE) Program,” *Newsletter Summary*, February 2011, <https://www.straighttalkonevidence.org/content/uploads/2018/03/HOPE-Probation-Evidence-Summary.pdf> (accessed October 22, 2020).

18. Pamela K. Lattimore, Doris Layton MacKenzie, Gary Zajac, Debbie Dawes, Elaine Arsenault, and Stephen Tueller, “Outcome Findings from the HOPE Demonstration Field Experiment: Is Swift, Certain, and Fair an Effective Supervision Strategy?” *Criminology & Public Policy*, Vol. 15, Issue 4 (November 2016), pp. 1103–1141.

despite the finding of a 2004 meta-analysis of home visitation programs that overall outcomes were not achieved.¹⁹ As of June 2017, HHS determined that 18 individual services were “evidenced-based.”²⁰ The program currently expends \$400 million a year to fund those services.²¹ However, in 2019, HHS found in its first early evaluation of program effectiveness that there was little difference in outcomes between the treatment and control groups: “No outcome area stands out as one where home visiting programs had large effects.”²²

Sixth, the third key attribute of the performance management model is the requirement for randomized controlled trials²³ and third-party evaluations.²⁴ However, the best application of the scientific gold standard analysis of federal programs’ data still indicates that the social programs are not achieving measurable impacts for participants.²⁵ As summarized in a 2018 review of the 13 known large RCTs of federal programs, “Eleven of the 13 RCTs found that the programs produced either no significant positive effects on the key targeted outcomes or small positive effects that dissipated shortly after participants completed the program.”²⁶

-
19. Monica A. Sweet and Mark Applebaum, “Is Home Visiting an Effective Strategy? A Meta-Analytic Review of Home Visiting Programs for Families with Young Children” *Child Development*, Vol 75, No. 5 (September 2004), pp. 1435–1456.
 20. Adrienne L. Fernandes-Alcantara, “Maternal, Infant and Early Childhood Home Visiting (MIECHV) Program: Background and Funding,” Congressional Research Service *Report for Members and Committees of Congress*, updated November 21, 2018, p. 20, <https://fas.org/sgp/crs/misc/R43930.pdf> (accessed October 6, 2020).
 21. U.S. Department of Health and Human Services, Administration for Children and Families, Health Resources and Services Administration, Maternal and Child Health Bureau. “The Maternal, Infant and Early Childhood Home Visiting Program: Partnering with Parents to Help Children Succeed,” April 2020, p. [1], <https://mchb.hrsa.gov/sites/default/files/mchb/MaternalChildHealthInitiatives/HomeVisiting/pdf/programbrief.pdf> (accessed October 6, 2020).
 22. Charles Michalopoulos, Kristen Faucetta, Carolyn J. Hill, Ximena A. Portilla, Lori Burrell, Helen Lee, Anne Duggan, and Virginia Knox, *Impacts on Family Outcomes of Evidence-Based Early Childhood Home Visiting: Results from the Mother and Infant Home Visiting Program Evaluation*, U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research, and Evaluation *Report* No. 2019-07, January 18, 2019, p. iii, <https://www.acf.hhs.gov/opre/resource/impacts-family-outcomes-evidence-based-early-childhood-home-visiting-results-mother-infant-home-visiting-program-evaluation> (accessed October 6, 2020).
 23. Tom Kalil, “Funding What Works: The Importance of Low-Cost Randomized Controlled Trials,” White House Blog, July 9, 2014, <https://obamawhitehouse.archives.gov/blog/2014/07/09/funding-what-works-importance-low-cost-randomized-controlled-trials> (accessed October 6, 2020).
 24. Coalition for Evidence-Based Policy, “Practical Evaluation Strategies for Building a Body of Proven-Effective Social Programs: Suggestions for Research and Program Funders,” October 2013, <http://coalition4evidence.org/wp-content/uploads/2014/05/Practical-Evaluation-Strategies-2013.pdf> (accessed October 6, 2020).
 25. David B. Muhlhausen, “Do Federal Social Programs Work?” Heritage Foundation *Backgrounder* No. 2884, March 19, 2014, http://thf_media.s3.amazonaws.com/2014/pdf/BG2884.pdf, and David Muhlhausen, “Evidence-Based Policymaking: An Idea Whose Time Has Come,” testimony before the Subcommittee on Human Resources, Committee on Ways and Means, U.S. House of Representatives, March 17, 2015, <https://www.heritage.org/article/testimony-evidence-based-policymaking-idea-whose-time-has-come>.
 26. These gold-standard program evaluations include education-focused programs such as Head Start, Even Start, Community Learning Centers, Abstinence Education, Teacher Incentive Fund, Student Mentoring, and Upward Bound. The evaluations also include job training programs like Job Corps and National Guard Youth Challenge. See Straight Talk on Evidence, “When Congressionally-Authorized Federal Programs Are Evaluated in Randomized Controlled Trials, Most Fall Short. Reform Is Needed,” June 13, 2018, <https://www.straighttalkonevidence.org/2018/06/13/when-congressionally-authorized-federal-programs-are-evaluated-in-randomized-controlled-trials-most-fall-short-reform-is-needed/> (accessed October 6, 2020).

One social program with a large RCT analysis is the Teen Pregnancy Prevention (TPP) program. TPP is a federal grant program that, pursuant to reauthorization and reform in 2009, should be distributing 75 percent of its funds to “evidence-based” approaches. Analysis of TPP’s grants from 2010–2014, however, reveals that only four of 24 funded models reduced teen pregnancy rates.²⁷

Like HOPE, MIECHV, and TPP, the Head Start program is another illustration of initial promise unsupported by subsequent RCT studies. Head Start is one of the longest-running programs designed to help underprivileged children. The more than \$10 billion-per-year federal preschool program was founded as part of Lyndon B. Johnson’s 1964 War on Poverty and modeled on the promise of the small-scale Perry Preschool and Carolinian Abecedarian programs.²⁸ A large-scale, rigorous study was required by Congress in 1998, and in October 2012, HHS released its comprehensive final report, which demonstrated that any advantages gained through Head Start are short-term and undetectable by the time a participant reaches the third grade.²⁹

It is a step in the right direction that policymakers and executive agencies began to collect and track data on outcomes for beneficiaries and that Congress began to require randomized controlled trials of major programs’ outcome measures. However, the available evidence shows that the key reforms of the performance measurement model generally have not produced results for vulnerable individuals.

The Social Impact Bond/Pay-for-Success Model

Over the past decade, there has been another evolution in social programs as policymakers sought provable and sustainable results in social programs. Social impact bonds (SIBs), as they are known in the United Kingdom, or pay-for-success programs, as they are known in the U.S., are models in which private investors fund up-front costs for social programs to innovate in interventions to assist vulnerable individuals.

27. Straight Talk on Evidence, “Evidence-Based Policy ‘Lite’ Won’t Solve U.S. Social Problems: The Case of HHS’s Teen Pregnancy Prevention Program,” June 18, 2019, <https://www.straighttalkonevidence.org/2019/06/18/evidence-based-policy-lite-wont-solve-u-s-social-problems-the-case-of-hhss-teen-pregnancy-prevention-program/> (accessed October 15, 2020).

28. U.S. Department of Health and Human Services, Administration for Children and Families, Head Start, Early Childhood Learning and Knowledge Center, “Head Start History,” <https://eclkc.ohs.acf.hhs.gov/about-us/article/head-start-history> (accessed September 15, 2020), and Lawrence J. Schweinhart, “Long-term Follow-up of a Preschool Experiment,” *Journal of Experimental Criminology*, Vol. 9, Issue 4 (December 2013), p. 389–409, <https://link.springer.com/content/pdf/10.1007/s11292-013-9190-3.pdf> (accessed October 15, 2020).

29. Michael 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*, U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation, OPRE Report No. 2012-45, October 2012, https://www.acf.hhs.gov/sites/default/files/opre/head_start_report.pdf (accessed October 15, 2020).

In a major difference from both the open grant and performance management models, in this model, the government repays the private investor only after a predefined goal is achieved, sometimes with an additional financial return. This moves social programs from paying for services to paying only for predefined and measurable outcomes. Most pay-for-success models have these primary players: government (national, regional, or local); the private-sector up-front investor; a direct service provider; and a third-party evaluator. The private investors and service provider are financially entirely responsible for achieving the results for the intended beneficiaries.

Generally, pay-for-success programs to date have invested in criminal justice, homelessness, child welfare, early childhood education, and youth development programs. The “outcome” or success metric will vary by each of these interventions. In many of the youth-focused programs, the success of the social impact bond will be judged by whether they avoided common problems such as, for instance, achieving a reduction in recidivism—i.e., returning to incarceration—or realizing a reduction in the percentage of children who need special education in kindergarten.

As of July 2019, 132 known SIBs had been launched in 25 countries.³⁰ The movement began in the United Kingdom just as the country began to search for alternative funding models to resource and transform its public welfare programs after the Great Recession and subsequent austerity measures.

One of the first SIBs ever funded was focused on recidivism within the criminal justice system. The Peterborough prison in eastern England created a SIB called One Service in 2011. The intervention was focused on prisoners who served terms of less than 12 months. The goal was to reduce recidivism by a target 10 percent and at least 7.5 percent among the recently released prisoners compared to a control group.³¹ The RAND Corporation evaluation showed an 8.4 percent reduction (but not the target 10 percent) in recidivism in the first cohort.³²

In the United States, the pay-for-success projects began as state initiatives, with the first launching in 2013. There are now 26 ongoing projects in 13 states. It takes years for many of these projects to be implemented and to demonstrate results, and most last a minimum of four years. Therefore, because many currently funded projects were launched in 2015 or

30. Social Finance, “Impact Bond Global Database,” <https://sibdatabase.socialfinance.org.uk/> (accessed October 15, 2020).

31. Emma Disley, Chris Giacomantonio, Kristy Kruithof, and Megan Sim, *The Payment by Results Social Impact Bond Pilot at HMP Peterborough: Final Process Evaluation Report*, United Kingdom Ministry of Justice *Analytical Series*, 2015, pp. 3 and 14, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/486512/social-impact-bond-pilot-peterborough-report.pdf (accessed October 15, 2020).

32. *Ibid.*

later, most do not yet have outcome data. For that reason, the U.S. review is limited to projects with available data, and most cases still have only preliminary results.

Because the United States has a high recidivism rate (within three years, nearly half of all prisoners released in 2005 were imprisoned again),³³ many U.S. pay-for-success projects are focused on lowering the recidivism rate. This includes the first U.S. project, the \$9.6 million New York City Rikers Adolescent Behavioral Learning Experience (ABLE) program was begun in 2013 to reduce recidivism among 4,458 formerly incarcerated youth who were 16–18 years old. The project’s goal was a 10 percent reduction in the recidivism rate from the historical Rikers rate of 50 percent. A study published by the Vera Institute of Justice concluded that the project’s recidivism reduction was not statistically significant, and the initiative was ended in August 2015.³⁴

Other recidivism projects have been initiated, including a State of New York project funded with a \$12 million U.S. Department of Labor grant to increase employment by five points and reduce recidivism by eight points among 2,000 formerly incarcerated individuals in Rochester and New York City, with services to be provided by the Center for Economic Opportunities.³⁵ However, third-party evaluations for these projects are not yet publicly available.

Other early U.S. pay-for-success projects have concentrated on early education. Utah focused on low-income children with a \$7 million project that funded preschool slots for three-year-old and four-year-old low-income children who were at risk of needing special education.³⁶ The project began in 2013 for a seven-year period. In 2015, investors and the state reported that 99 percent of the first cohort of “at-risk” children in the Utah program had avoided special education and that \$267,000 had been

-
33. Matthew R. Durose, Alexia D. Cooper, and Howard N. Snyder, “Recidivism of Prisoners Released in 30 States in 2005: Patterns from 2005 to 2010,” U.S. Department of Justice, Office of Justice Programs, Bureau of Labor Statistics Special Report No. NCJ 244205, April 2014, p. 1, <https://www.bjs.gov/content/pub/pdf/rprts05p0510.pdf> (accessed October 15, 2020).
 34. Jim Parsons, Chris Weiss, and Qing Wei, “Impact Evaluation of the Adolescent Behavioral Learning Experience (ABLE) Program,” Vera Institute of Justice, September 2016, <https://www.vera.org/downloads/publications/rikers-adolescent-behavioral-learning-experience-evaluation.pdf> (accessed October 15, 2020).
 35. News release, “Governor Cuomo Announces New York the First State in the Nation to Launch Pay for Success Project in Initiative to Reduce Recidivism,” New York State, Office of Governor Andrew M. Cuomo, December 20, 2013, <https://www.governor.ny.gov/news/governor-cuomo-announces-new-york-first-state-nation-launch-pay-success-project-initiative> (accessed October 15, 2020).
 36. Jeff Edmondson, Bill Crim, and Allen Grossman, “Pay-for-Success Is Working in Utah,” *Stanford Social Innovation Review*, October 27, 2015, https://ssir.org/articles/entry/pay_for_success_is_working_in_utah (accessed October 15, 2020), and Benjamin Wood, “Preschool Paying off for Goldman Sachs and Utah Kids, According to United Way,” *The Salt Lake Tribune*, October 7, 2015, <https://archive.satrib.com/article.php?id=3032598&itype=CMSID> (accessed October 15, 2020).

paid out to investors.³⁷ However, some researchers cited concerns with the initial estimation of the number of children potentially needing special education.³⁸

Chicago funded another preschool pay-for-success program. The city appropriated \$17 million for 2,600 slots across nine preschools for four-year-old low-income children.³⁹ The project initially saw 4.38 percent of children needing special education programming in kindergarten, 1.73 percentage points lower than the cohort's comparison group.⁴⁰ Investors were paid \$9,100 plus interest for every child in the first cohort who did not need special education.⁴¹

The first federal pay-for-success initiative, the Social Impact Partnerships to Pay for Results Act (SIPPRRA), was included in the Bipartisan Budget Act of 2018.⁴² The act gave \$100 million to the Department of the Treasury to pay for feasibility analysis and pay-for-success projects across the U.S. Of that total, \$66 million will directly reimburse outcome payments. With half of the funds set aside for projects that directly benefit children.⁴³ The 2019 Notice of Funding Availability outlines 21 options for "social benefit," and participants must demonstrate both the social benefit and federal, state, or local government savings.⁴⁴ The goals identified in the SIPPRRA program currently have alternate federal funding. These projects are a fraction of overall expenditures when compared to similar federal governmental programs.

The key positive aspects of this model are that the goals and measurement (usually a singular measure) of the alternative funding approach are defined at the outset. Further, the goals are evaluated by

-
37. "Pay for Success: Utah Program Reports Positive Results," New Profit, Amplify Blog, October 8, 2015, <http://blog.newprofit.org/amplify/pay-for-success-utah-program-reports-positive-results> (accessed October 15, 2020), and Nathaniel Popper, "For Goldman, Success in Social Impact Bond That Aids Schoolchildren," *The New York Times*, October 7, 2015, https://www.nytimes.com/2015/10/08/business/for-goldman-success-in-social-impact-bond-that-aids-schoolchildren.html?smid=fb-share&referrer=http://m.facebook.com&_r=1 (accessed October 15, 2020).
 38. See, for example, Libby Doggett and David Wilkinson, "Op-ed: Utah's Pay for Success Pre-K Model Is a Worthy Innovation," *The Salt Lake Tribune*, January 26, 2016, <https://archive.sltrib.com/article.php?id=3462523&itype=CMSID> (accessed October 15, 2020).
 39. News release, "Mayor Emanuel Closes the Gap on Pre-Kindergarten Education for 4-Year Old Children in Low Income Families," City of Chicago, Office of the Mayor, October 7, 2014, https://www.chicago.gov/city/en/depts/mayor/press_room/press_releases/2014/oct/mayor-emanuel-closes-the-gap-on-pre-kindergarten-education-for-4.html (accessed October 15, 2020).
 40. Illinois Facilities Fund, "Fact Sheet: Chicago Pay for Success/Social Impact Bond Program Evaluation Year Three," April 26, 2018, p. [1], <https://iff.org/wp-content/uploads/2018/04/Fact-Sheet-Chicago-SIB-Year-3-FINAL-4.26.2018.pdf> (accessed October 15, 2020).
 41. Melissa Sanchez, "Investors Earn Max Initial Payment from Chicago's 'Social Impact Bond,'" *The Chicago Reporter*, May 16, 2016, <https://www.chicagoreporter.com/investors-earn-max-initial-payment-from-chicagos-social-impact-bond/> (accessed October 15, 2020).
 42. Title VIII—Supporting Social Impact Partnerships to Pay for Results, in H.R. 1892, Bipartisan Budget Act of 2018, Public Law 115-123, 115th Cong., February 9, 2018, <https://www.congress.gov/bill/115th-congress/house-bill/1892/text> (accessed October 15, 2020).
 43. U.S. Department of the Treasury, Office of Economic Policy, "Social Impact Partnerships to Pay for Results Act Demonstration Projects," Notice of Funding Availability, Funding Opportunity Number UST-SIPPRRA-2019-001, <https://home.treasury.gov/system/files/226/SIPPRRA-NOFA-FINAL-FY2019.pdf> (accessed October 15, 2020).
 44. *Ibid.*

a third-party evaluator. In addition, the private-sector investor and direct service provider are compensated only if they achieve the pre-defined goals.

However, pay-for-success models use new funds instead of drawing from the federal funds already appropriated to existing social programs that serve the same purpose. This means that the model is not transforming the underlying programs and the bulk of federal funding directed toward capacity-building. The small scale of pay-for-success projects is also necessarily limited to single experimental project sites and therefore subject to the “single-instance fallacy.”⁴⁵ Because of the limited number of pay-for-success projects, the active projects are testing one service method out of hundreds of service options with similar goals. They are not able to show whether the success is a statistical fluke or due to the unique circumstances of the project site.

A New Paradigm: The Pay-for-Outcomes Model

The current evidence-based paradigm within social programs is inadequate. Few programs are found to work, those that do work usually have modest impacts, and the rare successful programs typically cannot be replicated. Running social programs with few if any expected results is hard on the taxpayer and likely to generate despair and alienation among recipients who deserve better.

What can be done? The answer is a pay-for-outcomes system built around clear principles.

Policymakers should transform federal social programs by rewarding providers directly for a clear and verifiable capacity-building outcome in the lives of vulnerable individuals. After nearly two decades of performance management and pay-for-success models, policymakers should combine the best features of these reforms by implementing the pay-for-outcomes model. The pay-for-outcomes model would be applied to entire federal social programs so that contracts are awarded to direct service providers only on the basis of their ability to achieve positive and clear-cut behavioral changes in the lives of vulnerable individuals, based on third-party verification of concrete outcomes as specified at the beginning of the project.

Because direct service providers would be awarded annual contracts only for predetermined measurable effects, paying for outcomes for individuals

45. Ioannidis, “Why Most Published Research Findings Are False,” p. e124.

would ensure that direct service providers are either rewarded or held responsible for their effectiveness in producing positive changes in the lives of vulnerable individuals. With this responsibility, programs would be empowered to focus on responding to the distinctive needs of the individuals they serve. This payment model would weed out providers who are not achieving results for vulnerable individuals. Service providers that did not achieve the outcome predetermined by policymakers would not be awarded subsequent contracts.

Unlike the pay-for-success model, which involves small-scale experiments outside of and in addition to the regular program funding stream, the pay-for-outcomes model would redirect the billions in currently appropriated federal funds for social programs. Direct service organizations that already participate in these social programs would be able to continue their work, but in subsequent years, contracts would be renewed based only on the ability of these service providers to produce individual outcomes in a cost-effective manner.

In the pay-for-outcomes model, policymakers would finally have solid evidence that programs either are or are not working. Without evidence of positive outcomes, service providers would not be paid, and this would give them a strong incentive to prove that outcomes are not only measured, but actually achieved.

Finally, the pay-for-outcomes model gives service providers on the ground both freedom and responsibility. Those closest to human problems are in the best position to address the complexities of those problems. Instead of imposing a handful of allegedly effective models, the pay-for-outcomes model encourages innovation and allows people to apply their intelligence and intuition to meet the on-the-ground situation. The pay-for-outcomes principle also requires that service providers be financially responsible, evaluated based on clear outcomes, and rewarded only for success.

New York City Contracting. A successful example of the pay-for-outcomes approach is found in the welfare reforms initiated by Mayor Rudy Giuliani in New York City.⁴⁶ In 1999, the Giuliani administration replaced the standard fee-for-service contracts for job placement/training organizations with pay-for-outcomes contracts. The contractors responded to the financial incentives embodied in the new contracts quickly and effectively.

46. The reform featured a “full engagement” policy requiring welfare recipients to participate in job search, training, or community service five days per week. The payment-for-outcome system complemented full engagement.

For example, the new contracts prioritized increasing the number of job placements (i.e., new hires) among recipients; within a year, the number of job placements generated by the contractors doubled.⁴⁷

A few years later, revised pay-for-outcomes contracts were issued under the administration of Mayor Michael Bloomberg. These revised contracts provided larger financial rewards for job retention relative to job placements. The contractors again responded to these incentives, and the average share of clients who retained jobs for more than 90 days was as much as 198 percent higher under the new contracts compared to the old.⁴⁸

Under the system in both administrations, potential providers were required to bid for contracts. The bidding process generated financial pressure that weeded out the less efficient providers. The original set of 15 prime contractors in 1999 was pruned to seven by 2010.⁴⁹

Five Policy Principles for Implementation of Pay-for-Outcomes

Five key policy principles are necessary for successful implementation of a pay-for-outcomes model.

1. **Set clear quantitative outcome measures for service providers.**

The pay-for-outcomes model requires a clear, concrete, and quantifiable outcome measure in the lives of those the program is intended to help. Each individual service provider should be measured according to the concrete outcomes obtained for its clientele. The desired outcomes would be defined, quantified, and objectively measured using reliable sources independent from the service provider. All providers would be under contract and paid for clear results. These payments could take place on a quarterly basis, based either on the total number of outcomes or on outcomes above an assumed nonintervention baseline.

47. Swati Desai, Lisa Garabedian, and Karl Snyder, "Performance-Based Contracts in New York City: Lessons Learned from Welfare-to-Work," Rockefeller Institute *Brief*, June 2012, p. 13, <https://rockinst.org/wp-content/uploads/2018/02/Performance-Based-Contracts-in-New-York-City.pdf> (accessed October 15, 2020).

48. Chart, "Average Quarterly 90-Day Retention," in Swati Desai, Lisa Garabedian, and Karl Snyder, "History of Welfare-to-Work Performance-based Contracts in NYC: Lessons Learned," paper presented at the International Conference of the Association for Public Policy Analysis and Management (APPAM), Moscow, Russia, June 2011, p. 9, http://umdcipe.org/conferences/Moscow/papers/Desai_History%20of%20Welfare-to-work%20Performance-based%20Contracts%20in%20NYC_Lessons%20Learned.docx (accessed October 15, 2020).

49. *Ibid.*, p. 13.

Complex measurements like benefit/cost analysis, while useful for some purposes, should be avoided as primary outcome measures because they involve complex calculations with many hidden assumptions that can easily be manipulated to achieve desired results. Even accurately reported data and well-conducted randomized control trials tend to track a wide array of data outcomes, and this can result in selective reporting of outcome findings. In some studies, researchers identify encouraging outcomes by focusing on data points that demonstrate positive effects while disregarding negative data points.⁵⁰

In order to avoid this pitfall, policymakers must identify the particular outcome they are seeking among recipients at the outset. This model should be implemented in social programs with either a control group or a known benchmark for comparison. For instance, a prison recidivism project would be awarded a contract based on the number of beneficiaries who avoided prison; a drug treatment program would be rewarded for the number of beneficiaries with clean drug tests for 12 months; an employment and training program would be judged by the number of beneficiaries who obtained a job and the average compensation of such employment; and an education program could be measured objectively by increases in high school completion or improved math and reading scores.

- 2. Provide payment for outcomes, not services provided or savings achieved.** Contractors would provide services but would be paid, wholly or in part, not for the services delivered but for the objective and measured outcomes achieved.⁵¹ For example, providers could be paid for the number persons actively employed and average wages received or based on the percent of reentering inmates who were not arrested or incarcerated in the three years after release. Multiple measures could be used within a single program.

50. Straight Talk on Evidence, "Newly-Published Study of Federal Job Corps Program Inaccurately Claims to Demonstrate Long-term Positive Effects," July 21, 2020, <https://www.straighttalkonevidence.org/2020/07/21/newly-published-study-of-federal-job-corps-program-inaccurately-claims-to-demonstrate-long-term-positive-effects/> (accessed October 15, 2020).

51. However, a very simple pay-for-outcomes model does not automatically ensure that the "treatment" provided actually produced the desired outcome. For example, random assignment evaluations of job training and job placement services routinely show that most of the individuals in the non-treatment control group obtain jobs on their own. The difference between the experimental group receiving the service and the non-treated control group is often slight. This means that a simplistic version of the pay-for-outcomes model would reward contractors for the many positive "outcomes" that would have occurred without the contractor's intervention. There are multiple ways to deal with this issue. One is to pay the contractor only for outcomes that are in excess of the norm expected without the intervention. For example, if 50 percent of the individuals entering a training program are expected to find jobs without the program, the contractor would be paid only for placements above the 50 percent rate.

Unlike the pay-for-success models operating in the states—and even at the federal level in the SIPBRA program—the measurement cannot be whether the program saved government money. The pay-for-outcomes model seeks concrete results in terms of building positive social outcomes and the capacity of individuals for self-support. Taxpayer savings may still follow when, for example, individuals become more likely to work or avoid reincarceration, but such savings would be unconnected to the flow of funds between government and program operators.

To achieve better outcomes, providers must have a financial stake in the success of their operations. They must be, at least in part, fiscally accountable for their own success or failure. Government financial rewards should be contingent on the outcomes achieved, and the provider should bear at least part of the cost of ineffective and failed services.

Similarly, pay-for-outcomes contracts can also be established between different levels of government. The government organization with funding authority could distribute those funds through pay-for-outcomes contracts with other government entities rather than through open-ended grants. For example, the federal government could channel its current social funds through pay-for-outcomes contracts with state and local governments. The state and local agencies receiving federal funds should be rewarded for achieving outcomes but should bear a greater share of the financial burden when programs are ineffective. Increasing the fiscal accountability of the agencies delivering the programs should incentivize and invigorate program improvement.

- 3. Place multiple service providers in competition.** Under a pay-for-outcomes system, service providers would be held financially accountable for their own success or failure. In addition, funding agencies should establish pay-for-outcomes contracts with multiple contractors who would be placed in competition with each other.⁵² The contractors could be either nonprofit or for-profit organizations.

52. Although providing competition in a small city or county where only one contractor can operate efficiently may seem difficult, a degree of competition can be maintained as long as the service contractor is operating on a payment-for-outcome basis using data that are similar to those of other firms in nearby locales. Each firm should be required to rebid its contract periodically against competitive bidders with the contract going to the low-cost bidder. This would ensure that over time, inefficient providers would be replaced by more efficient ones.

Contractors would not be allowed to select their own clients, as this would lead to choosing clients who are likely to succeed while ignoring more difficult clients. Some recent outcome studies have demonstrated this risk by highlighting subgroups with positive outcomes as opposed to the full sample of participants, misrepresenting program results.⁵³ Instead, a central administrator would allocate a fixed share of clients to each service provider as fairly as possible. The providers would then compete to achieve better outcomes for their roster of clients.

Service providers would be free to innovate and devise their own approach to achieving the desired outcomes. Creative approaches and diligent management are likely to flourish because the prosperity and survival of the organization depend on achieving successful outcomes.

4. **Shift funding automatically from low to high performers.** After the early years' implementation, successful programs will begin to compete against each other comparing best individual outcome results and lowest cost per result. The most successful will be awarded the highest portion of the subsequent years' funding automatically. This competition can also lead to clear and continued improvement over time. Funding would automatically be shifted from low performers to high performers. Effective providers who demonstrated better actual outcomes would not only receive more funds for their achievements. Over time, they would also increase their "market share" while that of inefficient providers would automatically shrink.

Providers who produced few measured outcomes would lose money and quickly drop out of the program. When this occurred, funds would not be cut; instead, they would automatically be shifted to the more efficient service providers.

All contracts would be renegotiated periodically. Contractors would bid for the price they sought to be paid for each unit or numeric goal achieved (number of persons employed, etc.). Inefficient providers would not gain contracts, and contractors that did not achieve targets

53. Straight Talk on Evidence, "Newly-Published Study of Federal Job Corps Program Inaccurately Claims to Demonstrate Long-term Positive Effects."

would not be renewed. Lower bidders would receive contracts and a greater share of future clients, especially if they meet program goals. This process would reduce the unit cost of outcomes, and the savings achieved would be redirected into expanded or enhanced services.

Future clients would benefit as funds are automatically shifted from ineffective high-cost service providers to more effective lower-cost providers and services in the community are expanded. Pay-for-outcomes is an exercise in improving the lives of vulnerable Americans. Budget savings for taxpayers are a byproduct of prioritizing these human outcomes.

5. **Continue RCT evidence-based evaluations.** To a considerable degree, the process described above is self-rectifying. Efficiency is automatically increased as funds are shifted from ineffective high-cost providers to more effective low-cost providers. Nonetheless, periodic reassessment of the overall social impact of the spending program would be critical. This could be accomplished by performing recurring random assignment evaluations on various providers within the program. RCT evaluations would be particularly important for ascertaining baseline outcomes that would occur in the absence of the overall program intervention.

Current Federal Programs Where Pay-for-Outcomes Should Be Implemented

Policymakers should begin to implement this model in social programs that have the best potential for the measurement of clear results. The following are some of the strongest candidates for implementation of the pay-for-outcomes model.

- **Employment and Training (E&T).** Employment and training programs are particularly ripe for reform because of their continually documented lack of success despite long-standing congressional efforts to reform them. There are 43 employment and training programs across nine federal agencies, primarily at the Departments of Labor,

54. For a full list of these 43 programs and their respective agencies, see Figure 2, "Federally Funded Employment and Training Programs, by Agency, Fiscal Year 2017," in U.S. Government Accountability Office, *Employment and Training Programs: Department of Labor Should Assess Efforts to Coordinate Services Across Programs*, GAO-19-200, March 2019, p. 11, <https://www.gao.gov/assets/700/698080.pdf> (accessed October 15, 2020).

Education, and Health and Human Services.⁵⁴ They cover all federal expenditures to help beneficiaries build skills and identify and obtain employment. According to the Council of Economic Advisers, “[a]ggregate spending on these programs totaled \$18.9 billion in 2019 alone.”⁵⁵

Although programs were reformed most recently in 2014 with the intention of funds being distributed according to the evidenced-based model, that reform has not demonstrated successful outcomes for recipients.⁵⁶ Congress should take a further step and transform E&T programs to measure clear outcomes for beneficiaries, principally employment attained and earnings achieved.

- **Substance Abuse Treatment.** The U.S. struggle against drug addiction spans decades. In 2018, more than 67,000 Americans died from drug overdoses—with synthetic opioids being the main driver.⁵⁷ The federal government expended more than \$10.5 billion on substance abuse treatment in FY 2017.⁵⁸

The Substance Abuse and Mental Health Services Administration (SAMHSA) has authority over the main block grant, called the Substance Abuse Block Grants. In addition to this block grant, the previously discussed Substance Abuse Prevention and Treatment (SAPT) Block Grant, and other SAMSHA substance abuse grants, the 21st Century Cures Act passed by Congress in 2016 allocates nearly \$500 million per year through the Opioid State Targeted Response (STR) program, which is distributed to all states and territories based on the degree of their opioid addiction problem.⁵⁹

55. Executive Office of the President of the United States, Council of Economic Advisers, *Government Employment and Training Programs*, June 2019, p. 1.

56. See, for example, *ibid.*, pp. 9 et seq.

57. Holly Hedegaard, Arialdi M. Miniño, and Margaret Warner, “Drug Overdose Deaths in the United States, 1999–2018,” U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics *Data Brief* No. 356, January 2020, <https://www.cdc.gov/nchs/data/databriefs/db356-h.pdf> (accessed October 15, 2020).

58. Table 1, “Federal Drug Control Spending by Function,” in Executive Office of the President, Office of National Drug Control Policy, *National Drug Control Budget: FY 2018 Funding Highlights*, May 2017, p. 16, https://www.whitehouse.gov/sites/whitehouse.gov/files/ondcp/Fact_Sheets/FY2018-Budget-Highlights.pdf (accessed October 2, 2020).

59. Elinor McCance-Katz, Assistant Secretary for Mental Health and Substance Use, Substance Abuse and Mental Health Services Administration, U.S. Department of Health and Human Services; Deborah Houry, Director, National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, HHS; Francis Collins, Director, National Institutes of Health, HHS; and Scott Gottlieb, Commissioner, Food and Drug Administration, HHS, “The Federal Response to the Opioid Crisis,” written testimony to the Committee on Health, Education, Labor and Pensions, U.S. Senate, October 5, 2017, <https://www.drugabuse.gov/about-nida/legislative-activities/testimony-to-congress/2017/the-federal-response-to-the-opioid-crisis> (accessed October 2, 2020).

While these programs may increase access to treatment, there is little evidence of their effect on other outcomes. Most of these programs collect evidence on four performance metrics, including sobriety rates immediately after release from treatment, completion of the program, client interviews, and internal studies. None of these data sets will tell the public whether medium-term or long-term sobriety was achieved. The true measurement that Congress should set for these programs is whether sobriety (established, for instance, by clean drug tests) is maintained for a set period (12 months) after treatment.

- **Recidivism Reduction.** With the U.S. historical recidivism rate of nearly 50 percent and the possibility for a clear outcome metric, social programs that focus on recidivism should be a priority for reform according to the pay-for-outcomes model.⁶⁰ In early 2020, almost 2.3 million Americans, including approximately 52,000 minors, were incarcerated in federal, state, and local prisons and jails.⁶¹ In FY 2015, the Department of Justice distributed \$68 million in recidivism reduction grants.⁶²

Reducing recidivism is clearly a goal of many federal programs outside of those funded through the Department of Justice, and there is some evidence that mental health or substance abuse programs may have promise, but there is no updated, systematic evidence that these programs work at large.⁶³ “Across the country,” according to a 2019 study, “government agencies and non-profit organizations operate a wide variety of programs designed to improve prisoner reentry outcomes. Though these programs are well-intentioned, few have undergone rigorous evaluation of their effects.”⁶⁴

60. Durose, Cooper, and Snyder, “Recidivism of Prisoners Released in 30 States in 2005: Patterns from 2005 to 2010.”

61. Wendy Sawyer and Peter Wagner, “Mass Incarceration: The Whole Pie 2020,” Prison Policy Initiative, March 24, 2020, <https://www.prisonpolicy.org/reports/pie2020.html> (accessed October 2, 2020).

62. Table B-2, “Authorized and Appropriated Funding for the Second Chance Act Grant Programs Administered by the Department of Justice, FY2009–FY2015,” in Nathan James, “Offender Reentry: Correctional Statistics, Reintegration into the Community, and Recidivism,” Congressional Research Service Report for Members and Committees of Congress, January 12, 2105, pp. 32–33, <https://fas.org/sgp/crs/misc/RL34287.pdf> (accessed October 2, 2020).

63. Executive Office of the President of the United States, Council of Economic Advisers, *Returns on Investment in Recidivism-Reducing Programs*, May 2018, <https://www.whitehouse.gov/wp-content/uploads/2018/05/Returns-on-Investments-in-Recidivism-Reducing-Programs.pdf> (accessed October 2, 2020).

64. Jennifer L. Doleac, “Wrap-Around Services Don’t Improve Prisoner Reentry Outcomes,” *Journal of Policy Analysis and Management*, Vol. 38, Issue 2 (Spring 2019), p. 508.

While Congress and administrators of these programs have relied heavily on evidence-based services that in some cases may deter reincarceration, a reformed recidivism program measurement would include payment based on whether the individual avoids reentry for from one to three years.

- **Federal Grants to Improve Self-Sufficiency.** The social program block grants that are currently distributed according to an open grant model—the Social Services Block Grant (SSBG), Community Services Block Grant (CSBG), and Community Development Block Grant (CDBG)—are clear examples of the need for action by Congress. As noted, these programs may have commendable goals, but they do not produce quantifiable results for their intended beneficiaries.

Congress should separate and define the goals of these programs. In the SSBG, for instance, Congress would have to clarify the statutory goals. If the new goal was to provide for foster care, then that goal and the corresponding measurement should be defined. A new measurable goal could be a shorter time frame during which children remain in foster care without permanency.⁶⁵ Likewise, in the CSBG, Congress would have to define and measure goals—in this case whether an individual achieves self-sufficiency and no longer depends on key federal transfer programs. Finally, in the CDBG, Congress should limit funding to service providers who find permanent (at least 12 months) housing (ideally nongovernmental housing) for low-income families.

Conclusion

Social programs aim to open transformational pathways for vulnerable individuals. They seek to help the unemployed obtain meaningful work, the formerly incarcerated to be rehabilitated, and those trapped in addiction to be freed from dependence. They aspire to help low-income and vulnerable Americans build their capacity and truly improve their lives and well-being.

The federal government has made bold promises with good intentions, spending billions on social programs for individuals who are vulnerable to

65. Lynch, "Social Services Block Grant."

poverty and dependence on government safety-net programs to achieve self-sufficiency. But today's social programs are not succeeding, partly because the design of current programs does not hold service providers accountable for delivering the results that policymakers, the public, and—most importantly—program beneficiaries desire. The recipients of these services and their families deserve better, and simply expanding funding for current programs that lack the necessary accountability to deliver those desired results is wholly inadequate to the task.

Instead, federal policymakers should transform social programs by paying directly for positive outcomes. A pay-for-outcomes system built around clear principles would transform federal social programs by rewarding providers directly for clear and verifiable outcomes in the lives of vulnerable individuals. This model would deliver systemic change in social programs and finally deliver the results that policymakers seek and vulnerable individuals deserve.

Leslie Ford is a Visiting Fellow in Domestic Policy Studies, of the Institute for Family, Community, and Opportunity, at The Heritage Foundation. **Robert Rector** is a Senior Research Fellow in Domestic Policy Studies.