

Designing Effective Evaluations for Direct Cash Programs

Considerations and Resources for States and Localities

Eve Mefferd and Heather Hahn

June 2025

Many states and localities across the United States have implemented some form of direct cash program in which money is provided directly to individuals or families as a public benefit.¹ Studying these programs can help improve them and can inform future programs across the country. To facilitate learning, direct cash programs can be designed in ways that set them up well for evaluation. In this summary, we summarize the reasons for evaluating direct cash programs, types of evaluations, and key considerations for designing evaluations.

WHY EVALUATE DIRECT CASH PROGRAMS?

States and localities should consider evaluating their direct cash programs for the following reasons:

- To understand if direct cash is having the desired impact so states and localities can spend money efficiently and effectively on programs that work
- To understand which components of the direct cash program design and implementation are key drivers of program impact
- To understand which parts of the program are working and which could be improved, including for specific populations²
- To improve future iterations of the program based on the evaluation's findings
- To generate buy-in from community members, policymakers, advocates, and funders, including through their involvement in evaluation and through sharing evaluation findings
- To add to the knowledge base that can help other states and localities interested in implementing similar programs, such as ideal amount and duration of direct cash, or how direct cash might complement other safety net programs³

In the following sections, we discuss multiple options for evaluating direct cash programs.

EXPERIMENTAL DESIGN FOR EVALUATING IMPACTS OF DIRECT CASH PROGRAMS

What Is Experimental Design?

Experimental evaluations involve randomly assigning study participants into one or more treatment groups, which receive an intervention (such as receiving the direct cash), or a control group, which does not receive the intervention.

How Does Experimental Design Work?

Experimental studies investigate whether an intervention, such as a direct cash payment, *causes* certain outcomes, such as improved mental health, reduced debt, or better housing or food security. A rigorous experimental design

can conclude that the intervention caused the outcomes because the randomly assigned treatment group and control group are basically the same, except for the difference in the intervention. Any differences researchers see are likely a result of the intervention. These analyses may also assess how much of the difference in outcomes can be explained by factors like age, race, gender, geography, or education (Hanson and Pergamit 2022).

When and How Should Experimental Design Be Used?

Many researchers and policymakers value experimental design because it enables them to understand the impact of a program or policy. For example, a study of a California direct cash pilot is currently underway to determine the impact of the program on health, financial well-being, employment, and perinatal and birth outcomes.⁴ Similarly, a study of the Long Beach Pledge direct cash initiative found that the program improved food, housing, and financial security, as well as physical and mental health of both parents and children.⁵

To conduct an experiment, researchers randomly put people in treatment and control groups, for instance through a lottery, so that if the groups are large enough, they will be equivalent. Researchers cannot use groups that are based on voluntary choices by participants, such as those who opted in to a program and those who opted out. This is because a fundamental difference could exist between the groups that led to their decision to opt in or out. For similar reasons, this method is not appropriate if a whole population receives the direct payment because you do not know what would have happened had you not delivered the cash.

The Urban Institute's evaluation of California's direct cash pilot is an example of how researchers and program administrators created treatment and control groups. First, they set eligibility criteria for the program, like having a low income, living in specific counties, and being pregnant or aging out of foster care. Individuals who met these criteria could apply for the program. Eligible applicants were randomly selected into groups. Some received the payments, while others did not. Researchers will analyze administrative data and survey both groups before the payments are disbursed, three and nine months into the program, at the end of the program, and six months after the program ends.

Certain experimental designs might not work for a state or locality because they are too costly or complex.⁶ The design above involves giving some individuals an intervention but not others within the same population, which may be complicated to implement or unfair to those who do not receive the intervention. However, this might happen naturally if funds are not sufficient to implement the intervention with the full population. In that case, a **randomized design** is the fairest way to deliver a program. Similarly, a **randomized rollout model** staggers the rollout of a program over time, allowing researchers to compare changes in outcomes before and after implementation at intervention and nonintervention sites while still giving everyone the treatment.

QUASI-EXPERIMENTAL DESIGNS FOR EVALUATING IMPACTS OF DIRECT CASH PROGRAMS

What Is Quasi-experimental Design?

Like experimental designs, **quasi-experimental evaluations** include treatment and control groups, but participants are not randomly selected into one group or the other. Rather, natural circumstances or other conditions make it so that only one group has access to the intervention. For example, people may live inside or outside a certain geographic area or may be born before or after a cutoff date that determines whether they receive the intervention.

How Does Quasi-experimental Design Work?

Quasi-experimental studies can point to correlations between an intervention and particular outcomes and make inferences about causation. For example, a quasi-experimental study could find that individuals who received direct

cash payments were less likely to experience anxiety and depression than those who did not receive payments. Researchers could infer that the money improved mental health outcomes with a certain level of confidence. However, quasi-experimental studies usually have less ability to assert that an intervention caused a particular outcome than randomized control trials.

Like in experimental designs, quasi-experimental analyses can control for factors other than the intervention to try to determine how much of a given outcome is explained by the intervention compared with how much is explained by other factors.

Quasi-experimental designs come in various forms: **Difference-in-difference analysis** compares differences in outcomes pre- and post-implementation for both a treatment and a control group. **Regression discontinuity design** uses an arbitrary numeric eligibility cut-off point, such as a certain amount of household income, and measures the differences in outcomes of individuals close to the cut-off on either side. A **synthetic control method** can be used when one whole population, like an entire state, receives an intervention and no other similar groups can be used for comparison. This method compares real outcomes to estimates of what would have happened if the program had not been implemented (McClelland and Mucciolo 2022).⁷

When and How Should Quasi-experimental Design Be Used?

Quasi-experimental methods can be more practical to conduct than randomized control trials because true random assignment with real-life policies is difficult to achieve. As an example, the Rx Kids research and evaluation team used a quasi-experimental approach to determine the impact of a direct cash program on individual and community-level outcomes, such as health care and breastfeeding service utilization, children's cognitive development and kindergarten readiness, educational attainment, physical and mental health, and economic security (Hanna and Shaefer 2024). As part of their research, they sampled individuals who gave birth in a specific hospital in Flint, Michigan. They compared families who received the direct cash to those who did not, either because they were not Flint residents or because they gave birth the year before the program began. They paired this analysis with additional qualitative, survey, and administrative data analyses of families in the program.

This is considered a quasi-experimental design because families were not randomly assigned to either study group—all eligible families received the funds. The team could still make robust inferences by using families in nearby areas and families who gave birth the year prior as comparison groups.

OTHER APPROACHES FOR EVALUATING DIRECT CASH PROGRAMS

Many other approaches to evaluating programs exist, including those that focus on important aspects beyond impact. These approaches are particularly useful for answering questions about implementation, such as, “When is direct cash a cost-effective solution?” or “How much cash might be needed and for how long for particular populations?” Implementation studies are often crucial for arriving at smarter impact study designs, especially in a field like direct cash where there are still many unknowns. Below, we highlight a few models states and localities might consider for implementation studies and evaluations.

Qualitative and Mixed Methods

Though experimental and quasi-experimental methods usually focus on quantitative analyses, evaluations may also include qualitative approaches, alone or in combination with quantitative methods. A **qualitative methodology** might include interviews and focus groups with people in the treatment and control groups, to hear about the experiences of individuals in both groups. Using mixed qualitative and quantitative methods can add depth and richness to a state or locality's understanding of a direct cash program.

Social Network Analysis

Social network analysis is a research method that maps and analyzes relationships between individuals, groups, organizations, and other entities that facilitate the exchange of information, resources, and influence (Bogle et al. 2018, 2021). Understanding these relationships and their structure can help investigators understand what drives processes and outcomes. Qualitative data on direct cash show that recipients often use it to “buy time” (e.g., drop a second job to spend more time with children after school); social network analysis is an ideal method for digging deeper into this output (Kappil et al. 2023). Urban Institute research has used social network analyses as part of an evaluation of a direct cash program in Austin, Texas (Bogle et al. 2024). The analysis demonstrated how direct cash strengthened relationships and led to a variety of positive outcomes.

Community-Engaged Methods and Participatory Research

Community-engaged methods and participatory research intentionally include impacted communities as partners and fellow experts in research, policymaking, or program planning. These methods can be applied to many types of evaluations. Using community-engaged methods can increase community buy-in, more effectively address local needs and challenges, and help research teams ask better questions and get more useful findings. The Urban Institute’s Community-Engaged and Participatory Methods Toolkits provide a wealth of information for conducting this work.⁸

Cost-Benefit and Cost-Effectiveness Evaluations

Cost-benefit analysis compares the total costs of implementing a program to the estimated monetary value of the program’s benefits. This approach has its limitations in effectively putting a price tag on some benefits of receiving direct cash, such as happiness, thriving, and mental and emotional well-being. **Cost-effectiveness analyses** are similar, but they compare a program’s costs to its goals and outcomes that are not necessarily monetary or related to cost-savings.⁹

Rapid Cycle Evaluation

The goal of **rapid cycle evaluation** is to quickly assess the effectiveness of different components of a program and provide actionable feedback that can be used to improve the program. Analysis is usually conducted during or throughout the implementation process using recent interim or real-time data. Rapid cycle evaluation designs are often flexible, staying nimble and pivoting as needed based on emerging findings and informational needs.¹⁰

Case Studies

Conducting a **case study** is another method of understanding the implementation and perceived impacts of cash transfer programs. By design, case studies do not attempt to be generalizable to different contexts. Instead, they take a deep dive into one particular program or area. Case studies can use qualitative or quantitative methods, and can serve as a record of the experiences, processes, and outcomes related to a program, which can be a useful tool both for the state or locality that implemented the program as well as others that are interested in the program and resulting lessons learned.¹¹

PREPARING FOR AN EVALUATION

Program planners should consider the following steps to prepare for a successful evaluation:

- **Plan out a timeline.** Build in plans for evaluation from the very beginning of program design, and incorporate evaluation throughout. It is more difficult to set up a rigorous evaluation if a program is already underway (Brooks Martin 2015).

- **Establish research partners.** Consider collaborating with researchers to help determine what type of evaluation is most appropriate for your program. Research organizations like the Urban Institute can help states and localities evaluate their programs.
- **Secure realistic funding.** The cost of an evaluation varies widely depending on many factors, such as the stage of the program's implementation, its size, and the feasibility of measuring impact. An evaluation partner, either external or internal to your organization, can help determine the resources necessary to conduct the evaluation properly. Dedicating funds to evaluation during the program design phase can help prevent this important task from being overlooked (Results for America 2024).
- **Think early on about possible metrics.** Determine what outcomes are most important for your program to achieve. Some evaluations of direct cash programs investigate things like how the funds impacted recipients' employment, or focus on largely economic, quantitative, or quantifiable outcomes. Qualitative outcomes can shed additional light on social, emotional, and family impacts of a program, and individuals' experiences of the program (Kappil et al. 2023).
- **Develop a theory of change.** The program should have a well-defined theory of change before implementation, and a logic model before embarking on an implementation or impact evaluation. This means that program leadership should be clear about what they believe is causing negative outcomes and what changes would improve them. They should also determine the specific inputs and activities that will logically lead to those improved outcomes (Hanson and Pergamit 2022).
- **Establish buy-in from involved staff and community members.** Evaluation requires cooperation and coordination with administrators, implementors, other program staff, and program beneficiaries, and may require these individuals to change how they operate or add to their workload, which can be difficult especially if they are already stretched thin. However, people on the ground and at all levels of program implementation are ultimately responsible both for the success or failure of a program, and that of an evaluation (Bryson et al. 2011).
- **Consider which time points to measure and report.** Think carefully about which points in time to test and report. Outcomes like mental health or perceived economic stability may decline at the end of a program if participants are worried about how they will get by without assistance, so measuring and reporting these measures halfway through their duration may deliver a clearer view of program impacts (Bogle et al. 2024).

ADDITIONAL RESOURCES

- Hatry, Harry P. 2020. *Evaluation Guide for Public Service Program Managers*. Washington, DC: Urban Institute. <https://www.urban.org/research/publication/evaluation-guide-public-service-program-managers>.
- Lery, Bridgette, Laura Packard Tucker, and Catherine Kuhns. 2022. *Evaluating Your Capacity for Impact Evaluation*. Washington, DC: Urban Institute. <https://www.urban.org/research/publication/evaluating-your-capacity-impact-evaluation>.
- Tatian, Peter A. 2016. "Performance Management to Evaluation." Washington, DC: Urban Institute. <https://www.urban.org/research/publication/performance-measurement-evaluation-0>.
- Edmonds, Leisha, Clair Minson, and Ananya Hariharan. 2021. "Centering Racial Equity in Measurement and Evaluation." Washington, DC: Urban Institute. <https://www.urban.org/research/publication/centering-racial-equity-measurement-and-evaluation>.
- Katz, Batia, and Shelley Metzenbaum. 2020. "Evaluation and Analytics to Improve, Not Just Evaluate, Program Performance." Washington, DC: Urban Institute. <https://www.urban.org/research/publication/evaluation-and-analytics-improve-not-just-evaluate-program-performance>.

NOTES

- ¹ "The Guaranteed Income Pilots Dashboard," Stanford Basic Income Lab, accessed May 8, 2025, <https://guaranteedincome.us/>.
- ² Example populations include elders on fixed incomes (City of Santa Monica. 2019. "Preserving Our Diversity (POD) Expansion to a Second Phase Pilot 2 Program." Santa Monica, CA: City of Santa Monica City Council Report), foster youth transitioning to adulthood (Harvey, Donovan, Shubhangi Kumari, Vivian Liu, and Mary Bogle. 2025. "Thrive Prince George's: Participant Characteristics at Enrollment." Washington, DC: Urban Institute), and pregnant individuals (Karasek, Deborah, Jazzmin C. Williams, Michaela A. Taylor, Monica M. De La Cruz, Stephanie Ateaga, Sabra Bell, Esperanza Castillo, Maile A. Chand, Anjeanette Coats, Erin M. Hubbard, Latrice Love-Goodlett, Breezy Powell, Solaire Spellen, Zea Malawa, and Any Manchikanti Gomez. 2025. "Designing the First Pregnancy Guaranteed Income Program in the United States: Qualitative Needs Assessment and Human-Centered Design to Develop the Abundant Birth Project." *JMIR Formative Research* 9. <https://doi.org/10.2196/60829>).
- ³ For example, the Housing Pledge demonstrates benefits of combining direct cash and housing vouchers. "Housing Pledge," Fund for Guaranteed Income, accessed May 29, 2025, <https://f4gi.org/pilot/housing-pledge/>.
- ⁴ Sarah Benatar, Bridgette Lery, and Jaclyn Chambers, "Evaluating California's Guaranteed Income Pilot Program," Urban Institute, <https://www.urban.org/projects/evaluating-californias-guaranteed-income-pilot-program>.
- ⁵ "Long Beach Pledge Dashboard," mySidewalk, City of Long Beach, February 21, 2024, <https://dashboards.mysidewalk.com/long-beach-pledge-data/long-beach-pledge-guaranteed-income-pilot-a0256deec04e>
- ⁶ Susan J. Popkin, "Getting off the Gold Standard," *Urban Wire* (blog), Urban Institute, October 23, 2012, <https://www.urban.org/urban-wire/getting-gold-standard>
- ⁷ For an example of a synthetic control methodology used to evaluate direct cash, see Jones, Damon, and Ioana Marinescu. 2022. "The Labor Market Impacts of Universal and Permanent Cash Transfers: Evidence from the Alaska Permanent Fund." *American Economic Journal: Economic Policy* 14(2). <https://doi.org/10.1257/pol.20190299>.
- ⁸ "Community-Engaged and Participatory Methods Toolkits," Urban Institute, accessed May 29, 2025, <https://www.urban.org/projects/cem-toolkit>.
- ⁹ For an example of a cost-effectiveness evaluation for a direct cash program, see McGuire, Joel. 2001. *The Cost-Effectiveness of Cash Transfers at Increasing Subjective Wellbeing and Affective Mental Health*. Oxford, UK: Happier Lives Institute.
- ¹⁰ For an overview of rapid cycle evaluation in the context of Temporary Assistance for Needy Families (TANF), see Scott Baumgartner and Katie Eddins, "Using Rapid-Cycle Evaluation to Inform Policy Decision Making," *Mathematica* (blog), March 17, 2020, <https://www.mathematica.org/blogs/using-rapid-cycle-evaluation-to-inform-policy-decision-making>
- ¹¹ For examples of case studies on direct cash programs, see "Our Key Findings," Stockton Demonstration (SEED), accessed May 29, 2025, <https://www.stocktondemonstration.org/#summary-of-key-findings> and Greenberg, Erica, Victoria Nelson, Justin B. Doromal, Heather Sandstrom, Soumita Bose, and Eleanor Lauderback. 2023. "Toward Pay Equity: A Case Study of Washington DC's Wage Boost for Early Childhood Educators." Washington, DC: Urban Institute.

REFERENCES

Bogle, Mary, Leiha Edmonds, and Ruth Gourevitch. 2018. "Connectedness Outcomes Produced by the Love Your Block Program." Washington, DC: Urban Institute.

Bogle, Mary, Lauren Fung, Rodrigo Garcia, Owen Noble, Fay Walker, Hala Ahmad, and Jack Landry. 2024. "Evaluation of the Austin Guaranteed Income Pilot." Washington, DC: Urban Institute.

Bogle, Mary, Matthew Gerken, Leiha Edmonds, and Christopher Davis. 2021. "Ascribing Outcomes of Complex Interventions to Social Networks." Washington, DC: Urban Institute.

Bogle, Mary, Owen Noble, and Lauren Fung. 2024. "Austin Guaranteed Income Pilot: Participant Outcomes at 12 Months." Washington, DC: Urban Institute.

Brooks Martin, Alison. 2015. "Plan for Program Evaluation from the Start." *NIJ Journal* 275: 24–28. <https://nij.ojp.gov/topics/articles/plan-program-evaluation-start>.

Bryson, John M., Michael Quinn Patton, and Ruth A. Bowman. 2011. "Working with Evaluation Stakeholders: A Rationale, Step-Wise Approach and Toolkit." *Evaluation and Program Planning* 34 (1): 1–12. <https://doi.org/10.1016/j.evalprogplan.2010.07.001>.

Hana, Mona and Luke Shaefer. 2024. "Results from the Rx Kids Participant Survey & Maternal Wellbeing Research Study." Flint, MI: RxKids.

Hanson, Devlin and Michael Pergamit. 2022. "Conducting a Randomized Controlled Trial (RCT) in Child Welfare." Washington, DC: Urban Institute.

Kappil, Tresa, Anna Jefferson, Swati Gayen, and AshLee Smith. 2023. *My Kids Deserve the World: How Children in the Southeast Benefit from Guaranteed Income*. Rockville, MD: Abt Global.

McClelland, Robert and Livia Mucciolo. 2022. "An Update on the Synthetic Control Method as a Tool to Understand State Policy." Washington, DC: Urban Institute.

Results for America. 2024. "Evaluation Policy Guide Chapter 8: Building a Culture of Evidence and Evaluation." Washington, DC: Results for America.

ABOUT THE AUTHORS

Eve Mefferd is a research analyst in the Family and Financial Well-Being Division. She researches child care and early education, economic security, and social safety nets.

Heather Hahn is the associate vice president of the Family and Financial Well-Being Division. Throughout her career, Hahn has conducted nonpartisan research on the wide range of issues related to the well-being of children and families, including cash assistance, nutrition assistance, and other supports for families with low incomes.

ACKNOWLEDGMENTS

This summary was funded by the Doris Duke Foundation as part of the Innovations in Cash Assistance for Children project. We are grateful to them and to all our funders, who make it possible for Urban to advance its mission. The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, or its funders. Funders do not determine research findings or the insights and recommendations of Urban experts. Further information on the Urban Institute's funding principles is available at urban.org/fundingprinciples. Copyright © June 2025. Urban Institute. Permission is granted for reproduction of this file, with attribution to the Urban Institute.