



# The Role of Pay for Success Evaluators

## Lessons Learned

*Sarah Gillespie*

*December 2016*

**Evidence is at the core of the pay for success (PFS) movement, and evaluation considerations are key to negotiating, implementing, and assessing PFS projects. Evaluators can play an important role during each stage of PFS development by grounding decisions in evidence and navigating stakeholders through evaluation design questions. Laying early groundwork for considerations such as evaluation enrollment and data collection also supports successful PFS implementation (box 1).**

Drawing on the experiences of Urban Institute (Urban) researchers engaged in PFS evaluations, this brief focuses on the unique characteristics and primary responsibilities of the evaluator in designing and implementing PFS evaluations, as well as ways in which the evaluator can leverage its expertise to go beyond these primary roles to work with each PFS stakeholder to strengthen the project and evaluation.

---

### BOX 1

#### **What Is PFS?**

Pay for success, or PFS, is an innovative financing mechanism that shifts risk for a new or scaling evidence-based social program from a traditional funder (usually a government) to a third-party investor (usually a private organization or nonprofit). At the heart of all PFS projects is a test of whether a social program can improve outcomes for a specific group of people. If the program works (as measured by a rigorous evaluation), the project is a success. Investors get their money back (with a potential positive return), the government realizes potential future cost savings, families and society benefit from better outcomes, and social service providers strengthen the case for funding their model. For more information on PFS, visit <https://pfssupport.urban.org/>.

---

Urban’s experience has shown that the early and continuous engagement of an evaluator is important as PFS projects proceed through development. Especially early on, evaluators can help build knowledge of the existing evidence base, assess the capacity for using evidence among PFS stakeholders, and apply lessons learned in the feasibility stage toward project structuring and implementation. The evaluator is the only PFS partner without a stake in the financial outcome—its agenda should be focused on collecting and rigorously evaluating the data, bringing an important objective perspective to a PFS project.

This brief is written for those interested in the role of evaluators in a PFS project. Published separately is a more technical companion report, designed primarily for evaluators, that describes some of the unique practical considerations when conducting a PFS evaluation (box 2). By sharing our experiences, we hope to shorten the learning curve for future PFS evaluators and partners.

---

## BOX 2

### Additional Resources

Urban has also published a more technical companion report on practical considerations for PFS evaluations and a report on the evaluation and research design of Denver’s PFS project, which Urban is evaluating. Further, the Pay for Success Initiative at the Urban Institute has authored additional resources that may be helpful for PFS project design and evaluation methods. These resources include “Foundational Concepts and Terms of Pay for Success,” “From Evidence to Outcomes: Using Evidence to Inform Pay for Success Project Design,” “Measuring Success in Pay for Success: Randomized Controlled Trials as the Starting Point,” and “An Introduction to Evaluation Designs in Pay for Success Projects,” available at [pfs.urban.org](https://pfs.urban.org).

---

## The Evaluator’s Role during PFS Project Development and Implementation

Stewarding a PFS project from development through implementation can be a long, resource-intensive, and complex process (Azemati et al. 2013). Advancing a project to launch may take at least a year, with most requiring several years. This process has many moving parts and has been discussed in different stages (Liebman and Sellman 2013). In this brief, we will discuss the evaluator’s perspective as it relates to four key project stages: feasibility, structuring, implementation, and evaluation.<sup>1</sup>

This paper also considers the five primary partners within a PFS project: the government, investors, intermediaries, service providers, and evaluators. Government partners function as the project initiator, identifying what problem to address through a PFS project and paying if established performance metrics are met. Private and philanthropic investors provide up-front capital to launch the program, assuming the financial risk in exchange for potential financial returns if the project meets established performance metrics. The intermediary provides financial and project management expertise and is a party to the legal contract with the government, managing other relationships with investors, service providers, and sometimes evaluators. The service providers implement the chosen program model for the target population to achieve the agreed-upon outcomes. Finally, the evaluator assesses whether the intervention has met the performance metrics specified in the PFS contract.

Although the evaluator will work with all PFS partners, the intensity of work with certain partners varies in each phase. In the following sections, we discuss the key tasks of the evaluator in each phase of the PFS project and the partners with which the evaluator will work most closely. We also include several roles that an evaluator can play and that we believe can improve the strength of the overall project, but we recognize that not every project will seek to fully engage the evaluator. The accompanying figures highlight the relationship of the evaluator to the other PFS partners in each phase with whom the evaluator's work is most intensive.

## Feasibility

During the feasibility stage, evaluators can support PFS partners, particularly the government and intermediary exploring a project, in understanding the likelihood of a proposed program achieving the outcomes of interest and how an evaluation could effectively measure program effects. This grounds early project decisions in the best available evidence.

**Identify a program model and interpret the evidence base.** Evidence should be at the core of PFS for all stakeholders; it mitigates risk by improving the likelihood of achieving outcomes and decreasing the risk of program failure. Evidence is only as good as the evaluation that produces it. The most evidence-based programs have undergone rigorous evaluations several times. Few programs meet this very high standard of evidence (Milner and Eldridge 2016). Because the selected program's evidence base informs the next phases of PFS development, the evaluator plays an important role in ensuring that other partners understand the strengths and weaknesses of an existing program's evidence base and ultimately use that to select a program model.

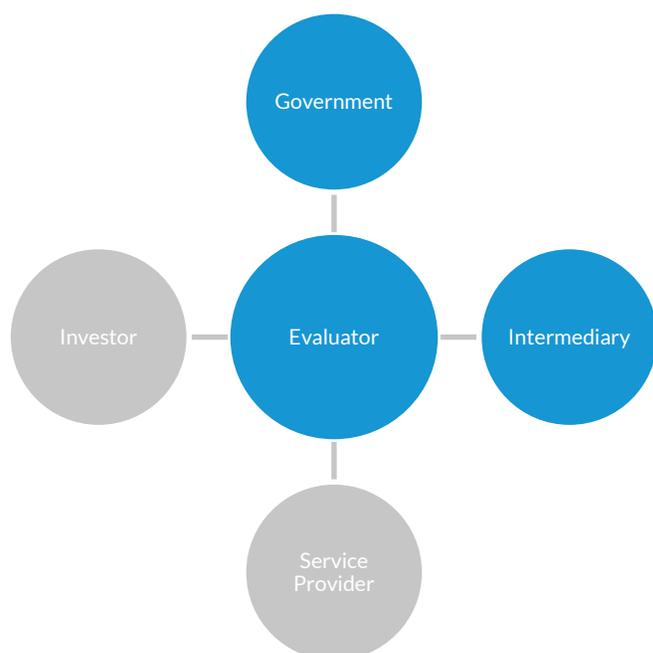
**Define a target population as well as the referral and take-up rates needed to achieve sufficient analytic power.** Once a program is selected, the feasibility tasks shift to data analysis. Access to administrative datasets is confirmed and negotiated during the feasibility stage. Administrative datasets are often "matched," or connected to one another so that many pieces of information are brought together to identify a specific target population. This type of data match is often done to identify people who use several public systems, such as people who access homeless shelters and spend time in jail or frequent emergency health departments. Long-standing relationships or existing databases that maintain matched data are important to efficiently identifying the target population and the potential cost implications of serving that target population.

Analysis of the target population helps determine the total size of the eligible population and whether the program can be implemented on a large enough scale to reliably detect any effects of the program. When program effects, such as a decrease in jail bed-days, are expected to be large, a relatively smaller number of participants may be sufficient for a rigorous evaluation. If program effects are more modest, a larger number of participants may be required for the evaluation. The evaluator is critical for the work of identifying and analyzing a target population because the information uncovered will inform important evaluation design considerations.

**Identify outcomes of interest.** After defining a target population, administrative data are used to understand historical baselines for the outcomes of interest. Evidence of the program model can be mined to understand expected program effects, including as they relate to potential cost savings that may inform how payments will be made to investors. Some of this work may move into the structuring phase, but the evaluator will need to ensure that the administrative work pertaining to data sharing and analysis, such as consent and data sharing agreements, are feasible and in development as early as possible.

FIGURE 1

### The Role of PFS Evaluators in the Feasibility Stage



**Start to develop an evaluation design.** Using what is learned from data analysis on the target population and outcomes of interest, the evaluator can begin to draft options for an evaluation design. The evaluation design is part of the legally binding PFS contract and dictates in detail how program success will be measured for payment. The design should outline several components important to the feasibility of a strong evaluation, including assumptions about anticipated effects, the program’s theory of change, the expected pace of referrals and enrollment in or take-up of the program, the data collection process, analysis of impacts, and reporting of results. The evaluation design will be finalized in the structuring phase through close collaboration with other key PFS partners.<sup>2</sup>

**Design a pilot stage.** Even though PFS practice urges a program design based on evidence, the evidence is typically incomplete and does not account for the implementation risk that can occur when a program that has been evaluated in one place is adopted in another context or for a different target population. During the feasibility phase, the evaluator may help the PFS project plan for a pilot stage in which implementation issues can surface and be addressed, establishing a feedback process between

the evaluator and other partners that will be important during implementation. A pilot phase can illuminate unanticipated challenges to program implementation and will inform and shape program assumptions made in the evaluation design, such as the feasibility of intended referral practices, the pace at which individuals enter the program, and the rate at which individuals choose to enroll in the offered program. The pilot is designed during the feasibility phase but runs during implementation (see the Implementation section later in this brief).

## Structuring

During the structuring stage, evaluators support all PFS partners in the final negotiations that produce the PFS contract, outlining project implementation, governance, and the parameters by which success payments will be made.

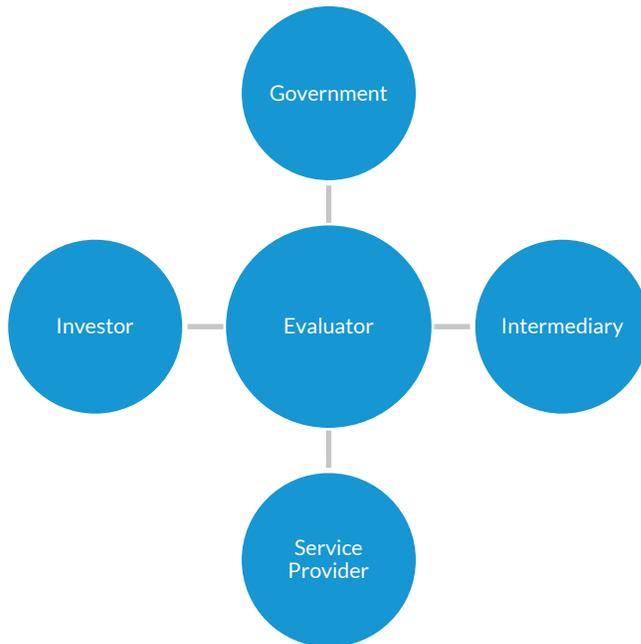
**Inform development of success payments based on evidence.** Evaluators play an important role during the structuring phase, grounding decisions in evidence and raising any feasibility or research design questions that may affect success payments. Considerations include, for example, whether success payments are based on interim program outcomes (for those directly served by the program), final program impacts (based on a comparison to outcomes for a control or comparison group), or both.<sup>3</sup>

**Build understanding, consensus, and buy-in on a rigorous evaluation design.** PFS investors and government partners should seek the most rigorous method—often an experimental design with random assignment—for determining whether the program achieved the established performance metrics.<sup>4</sup> Evaluators must work closely with these stakeholders and those who will be directly involved in PFS implementation to determine the most rigorous design possible. All partners should feel engaged in discussions around this decision; consensus-building should be a goal so that all partners are committed to the chosen design. The evaluator can outline evaluation design options and actively involve stakeholders, especially service providers, to identify and address concerns or potential challenges to each design option.

**Finalize data-sharing agreements and address any ethical considerations.** If not completed, evaluators should during the structuring phase finalize data-sharing agreements both with service providers to obtain program monitoring data and with those agencies (usually government agencies) that will provide the administrative data needed to measure outcomes. Data-sharing agreements will likely raise data security considerations, such as how consent procedures will be navigated, how data will be kept secure and confidential, and who will own the data. Many of these decisions also factor into the review and approval of the research design by an institutional review board (another task for the evaluator) before implementation begins. Evaluators may need to educate one or more of the partners on the need to have an institutional review board review the research procedures. An institutional review board review is vital to ensuring the appropriate protections for people participating in a program evaluation.

FIGURE 2

The Role of PFS Evaluators in the Structuring Stage



**Work with investors to understand financial risks related to the program model and research design.** As Gordon Berlin (2016) discussed in his reflections on the Riker’s Island project in New York City, “due diligence for a social impact bond bears little resemblance to the due-diligence process with which most social investors are familiar ... private investors are unlikely to possess the specialist content expertise needed to master social program design and context; the understanding of evaluation research needed to assess the potential effectiveness of an intervention” (13). Evaluators may be called upon by PFS investors to help explain the nuances of the program model and research design as they conduct due diligence and make the case for investment. Depending on investors’ motivations, considerations may include, for example, the effect of different program take-up rates on statistical power, the differences in statistical estimation methods and implications of choosing one over another, or the role of statistical significance in determining the credibility of a measured impact.

## Implementation

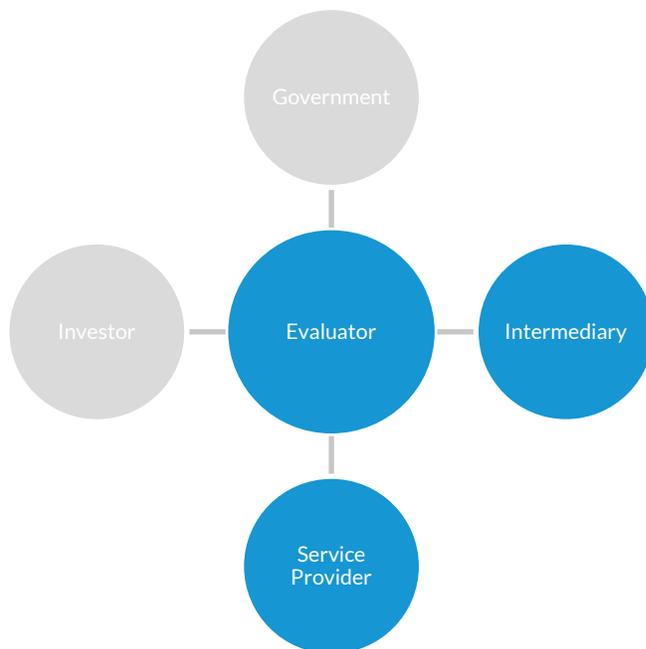
During the implementation stage, evaluators support PFS partners, particularly the intermediary and service providers involved in day-to-day operations, in monitoring program implementation and recommending course changes as necessary to maximize the success of the program and evaluation.

**Implement a pilot phase and recommend necessary course changes.** A pilot phase is planned in the feasibility stage and conducted during early implementation. It allows evaluators and other key stakeholders to test the assumptions made in the evaluation design, such as the way the referral and enrollment process works and the expected flow of individuals into the program. A pilot allows time for

necessary course correction to improve the program and refine the evaluation plan before performance metrics are tied to success payments. For example, a pilot phase may reveal that program referrals are slower than anticipated or that the enrolled population has different characteristics than expected. As a result, eligibility criteria should be adapted to generate the number and type of referrals necessary for the success of the program and evaluation.

FIGURE 3

### The Role of PFS Evaluators in the Implementation Stage



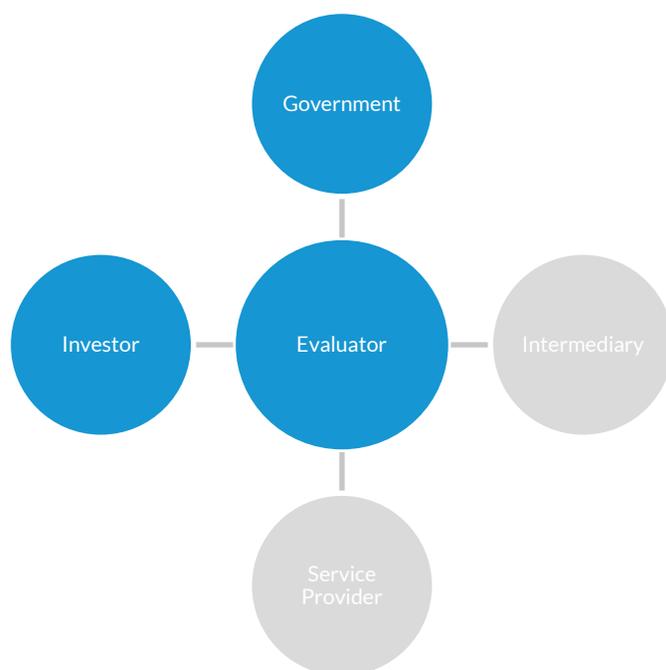
**Develop a program monitoring dashboard.** Throughout implementation, evaluators will work to closely monitor program implementation indicators and help raise and navigate any operating concerns (e.g., rates of referral, take-up, and attrition) that should be addressed before they negatively affect the evaluation. Program monitoring data at the individual level are collected from PFS service providers, often through a tool and with a process created by the evaluators. These data are then reported at the aggregate level in a dashboard regularly shared with all PFS partners. The evaluator should help ensure these data are understood by an operating committee composed of those who will be involved in day-to-day program implementation and by a governance body that includes leadership from each PFS stakeholder.

## Evaluation

During the evaluation stage, evaluators help PFS partners, particularly the government and investors seeking to translate outcomes to success payments, understand the results of the project and ensure the project adds to the knowledge base of what works.

**Document outcomes for success payments.** PFS evaluations are similar to other evaluations in many respects, but integrating established evaluation methods into the PFS model requires some unique practical considerations. These considerations are discussed in detail in a companion paper (Gillespie et al. 2016), and they include strategies for protecting evaluator objectivity and independence by, for example, reporting outcomes publicly and by building evaluation flexibility into the PFS contract to allow the analysis plan to change as data become available.

**FIGURE 4**  
**The Role of PFS Evaluators in the Evaluation Stage**



**Work with an independent validator.** To date, a few PFS projects have engaged an independent validator. Engaging such an entity allows an external party to confirm that evaluation activities are conducted accurately and rigorously, subsequently ensuring that payments are triggered only when appropriate. The evaluator will work with the independent validator to ensure they have both the necessary data to conduct data verification and a way to achieve consensus if a disagreement arises about analysis plans or statistical methods.

**Expand the evaluation.** PFS contracts focus heavily on outcomes tied to success payments, specifying in detail how they are measured. Often, success payments are tied to the outcomes with the strongest evidence base. However, evaluators can help bring attention to opportunities to expand the evaluation to include secondary outcomes that, though they aren't tied to success payments, can further build the evidence base for the program. Additionally, evaluators should make PFS partners aware of the importance of a process study, a key component to interpretation of PFS results discussed

below. If agreed upon, expanding the evaluation may require additional fundraising work by PFS partners.

**Support interpretation of results.** Although the evaluator's primary role is to determine the impact of the program, the evaluator will almost certainly be asked why the program did or did not achieve its expected outcomes. A process study is recommended to help understand the evaluation's results. A process study documents the implementation of a program, explores whether the program was implemented as intended, and identifies any barriers or improvements to program implementation. Absent a process study, little information would be available to help stakeholders understand the PFS results. If a PFS project does not achieve the agreed-upon outcomes, for example, that may suggest the program did not work, but a process study could reveal that the program was not implemented with fidelity to the original model. That suggests that the program may still be effective if implemented as intended. If the program achieves established performance metrics, process study findings provide information to those who want to replicate the program and its success. A full understanding of the data should inform whether the PFS achieves outcomes for the target population.

## Conclusion

Evaluators have a role in all four stages of PFS projects: determining PFS feasibility, negotiating of PFS structuring, implementing the project day to day, and reporting evaluation results. Evaluators work closely with each of the PFS partners along the way, including governments, intermediaries, service providers, and investors.

To ensure the evaluator can provide the most support and value to the PFS project, evaluators should be brought on board as early as possible and provided with the time and support necessary for the evaluation planning and pilot implementation phases. As with all other PFS partners, evaluators should be in close communication and coordination through project development and implementation.

PFS evaluations, and our own experiences as PFS evaluators, are still few enough that the roles and lessons discussed in this paper are not presented as best practices or models for the field. New lessons will certainly be learned in the coming years. We hope that by documenting and learning from our experiences thus far, future PFS projects will be strengthened by understanding the value of and providing space for a robust role for evaluators.

## Notes

1. These phases may (and frequently do) overlap, with one phase still in process even while parts of the next phase begin.
2. For information on evaluation design options and considerations, see Walsh et al. (2016)
3. For more information on success payments, see Hatry et al. (forthcoming).
4. For information on experimental designs, see Milner and Walsh (2016).

## References

- Azemati, Hanna, Michael Belinsky, Ryan Gillette, Jeffrey Liebman, Alina Sellman, and Angela Wyse. 2013. "Social Impact Bonds: Lessons Learned So Far." *Community Development Investment Review* 9 (1): 23–33.
- Berlin, Gordon L. 2016. *Learning from Experience: A Guide to Social Impact Bond Investing*. New York: MDRC.
- Gillespie, Sarah, Akiva Liberman, Janine Zweig, Devlin Hanson, Mary Cunningham, and Michael Pergamit. 2016. *Practical Considerations for Pay for Success Evaluations*. Washington, DC: Urban Institute.
- Hatry, Harry, Matthew Eldridge, Arden Kreeger, and Reed Jordan. Forthcoming. *Making Sound Cost Decisions in Pay for Success Projects*. Washington, DC: Urban Institute.
- Liebman, Jeffrey, and Alina Sellman. 2013. *Social Impact Bonds: A Guide for State and Local Governments*. Cambridge, MA: Government Performance Lab, Harvard Kennedy School of Government.
- Milner, Justin, and Matthew Eldridge. 2016. "From Evidence to Outcomes: Using Evidence to Inform Pay for Success Project Design." Washington, DC: Urban Institute. <http://www.urban.org/research/publication/evidence-outcomes-using-evidence-inform-pay-success-project-design>.
- Milner, Justin, and Kelly Walsh. 2016. "Measuring Success in Pay for Success: Randomized Controlled Trials as the Starting Point." Washington, DC: <http://pfs.urban.org/library/content/measuring-success-pay-success-randomized-controlled-trials-starting-point>.
- Walsh, Kelly, Rebecca TeKolste, Ben Holston, and John K. Roman. 2016. "An Introduction to Evaluation Designs in Pay for Success Projects." Washington, DC: Urban Institute. <http://pfs.urban.org/library/content/introduction-evaluation-designs-pay-success-projects>.

## About the Author



**Sarah Gillespie** is a research associate in the Metropolitan Housing and Communities Policy Center at the Urban Institute. Her research focuses on housing and homelessness, place-based initiatives, and performance measurement. Her work currently includes an evaluation of a supportive housing social impact bond for frequent users of the criminal justice system in Denver and a five-site national evaluation of supportive housing for families involved in the child welfare system.

# Acknowledgments

This paper was funded by the Laura and John Arnold Foundation. 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 author 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 [www.urban.org/support](http://www.urban.org/support).

This paper benefited greatly from the expertise of several colleagues who are also co-authors on the companion paper: Akiva Liberman, Janine Zweig, Mary Cunningham, Mike Pergamit, and Devlin Hanson. Thanks also to Janine Zweig, Mary Cunningham, and the Urban Institute PFSI team, including Matt Eldridge, Kelly Walsh, Justin Milner, and Rayanne Hawkins, for their valuable review and comments.



2100 M Street NW  
Washington, DC 20037  
[www.urban.org](http://www.urban.org)

## ABOUT THE URBAN INSTITUTE

The nonprofit Urban Institute is dedicated to elevating the debate on social and economic policy. For nearly five decades, Urban scholars have conducted research and offered evidence-based solutions that improve lives and strengthen communities across a rapidly urbanizing world. Their objective research helps expand opportunities for all, reduce hardship among the most vulnerable, and strengthen the effectiveness of the public sector.

Copyright © December 2016. Urban Institute. Permission is granted for reproduction of this file, with attribution to the Urban Institute.