

Pay for Success in the U.S.: Are Viable and Sustainable Models Emerging?

Carolyn J. Heinrich

Patricia and Rodes Hart Professor of Public Policy, Education and Economics
Vanderbilt University
230 Appleton Place, Nashville, TN, USA; +001 615 3221169
Email: carolyn.j.heinrich@vanderbilt.edu

Sarah E. Kabourek

Doctoral candidate
Department of Leadership, Policy, and Organizations, Peabody College
Vanderbilt University
Email: sarah.e.kabourek@vanderbilt.edu

November 11, 2018

Preliminary; please do not quote or cite without permission from authors.

We thank the Peabody College of Education and Human Development at Vanderbilt University for support of this research, the U.S. Department of Education for providing documents for our analysis, and the Preschool Pay for Success feasibility study program applicants who participated in interviews and shared their time, information and perspectives on their feasibility pilot applications and preschool program expansion efforts.

Abstract

Social Impact Bonds, or Pay for Success (PFS) as known in the U.S., have attracted inordinate attention as a novel strategy for financing and providing preventative services to our most vulnerable populations. We undertake a systematic, qualitative analysis of the Preschool PFS feasibility pilot grant applications and projects initiated by the U.S. Department of Education in 2016 to encourage state and local exploration of the viability of PFS for implementing high-quality preschool programs. We examine the feasibility pilot features and investigate why grant applicants saw PFS as a promising strategy for achieving their short- and long-term preschool program goals. We also uncover challenges encountered, lessons learned, and the perceived viability and sustainability of fully executed PFS preschool programs. We find that few projects are advancing toward formal PFS arrangements, with many struggling to overcome constrained capacities, structural and political barriers, and inherent incentives to minimize risk and loss.

Introduction

Social Impact Bonds (SIBs), or Pay for Success (PFS) initiatives as they are more widely known in the U.S., have attracted outsized attention as a novel, “nothing to lose strategy” for addressing some of our most “intractable social challenges” (Gustafsson-Wright, 2015).

Although characterizations of SIBs/PFS as a social investment tool are not uniform in the literature, they are typically distinguished by: (1) their focus on preventative-oriented programs that are expected to have long-term payoffs; (2) performance-based contractual mechanisms for financing that pay only for specified program outcomes when attained; and (3) a configuration of partners that include the government, private sector investor(s), an intermediary that manages the arrangement, an independent evaluator, and the program implementer(s) (Arena et al., 2016; GAO, 2015; Warner, 2013). The first official SIB pilot was initiated in the United Kingdom (UK) in 2010, and some have pointed to the global recession as a catalyst for the emergence of SIBs (Williams, 2018). Indeed, Roy et al. (2017) describe SIBs as the “archetypal ‘solution looking for a problem’,” in that as governments were scaling back social welfare expenditures, SIBs presented an opportunity as a no- or low-risk instrument for financing and providing beneficial services to vulnerable populations.

Many SIBs and PFS initiatives are still in early, “feasibility” stages of development. Massachusetts was the first U.S. government entity to pass legislation authorizing funds for PFS (in January 2012), with the stated objective of contracting to improve government outcomes and lower the costs of services through public-private PFS partnerships (Mass. Gen. Laws ch. 10, § 35VV). By the spring of 2017, a total of eight states had enacted some form of PFS legislation, although as Curran (2017) explains, the structure and components of the legislation vary considerably, such that there is no “typical PFS” model guiding the contractual arrangements and funding mechanisms as they evolve. This contrasts notably with the more centralized approach to

SIB commissioning in the UK, where a cabinet-level office provides a range of tools and resources for SIB project development, including a SIB contract template, a unit-cost database, cost-benefit analysis guidance, and a government outcomes (GO) lab that assists government entities in assessing the feasibility of SIB projects and contracts.ⁱ In February 2018, four years after it was first proposed, the U.S. Congress passed federal legislation as part of the Bipartisan Budget Act of 2018 (PL 115-123) that provides \$100 million in federal funding to support PFS partnerships, feasibility studies, projects and evaluations, suggesting some progression toward a more integrated approach to guiding PFS initiatives in the U.S.

With a small number of fully launched PFS programs and even fewer that have advanced to the point where outcomes are assessed, the knowledge base on PFS is still primarily conceptual and descriptive (Curran, 2017; Maier & Meyer, 2017). Yet by examining the components and mechanisms of PFS arrangements as they are being structured and operationalized in feasibility studies and emerging programs, there is considerable learning and insights to be reaped. Furthermore, many core features of PFS and the motivations underlying their incorporation are not as novel as the hype around these new tools for government service provision would suggest (Warner, 2013). In fact, many of the challenges that are arising as feasibility studies advance have been the subject of substantial theoretical and empirical investigation in the literature on performance-based contracting and pay for results reforms.

We undertake a systematic, qualitative analysis of the Preschool PFS feasibility pilot grant applications and studies, both those funded and *not* funded by the U.S. Department of Education (DOE) in the 2016 Fiscal Year. The purpose of the feasibility pilot grants, awarded to eight entities (totaling approximately \$3 million), was to encourage state and local exploration of whether PFS is a viable and potentially effective strategy for implementing high-quality

preschool programs. We received the 20 applications that were submitted to the U.S. DOE, along with the reviewer ratings and comments, and we conducted interviews with 12 applicants (5 of 8 awardees, and 7 of 12 non-awardees) approximately 18 months following the funding award decisions. The objectives of this study are threefold: (1) to identify the feasibility study proposal features that the DOE saw as advantageous and aligned with anticipated PFS components; (2) to understand from the perspective of grant applicants why they saw PFS as a promising strategy for achieving their short- and long-term preschool program goals, and (3) to assess the status of the feasibility pilots, as well as the preschool program initiatives in sites that were not funded, and uncover the challenges encountered, lessons learned, and perceived viability and sustainability of fully executed PFS preschool programs in the U.S.

The following section begins by framing the launch of PFS and preschool feasibility pilots in the context of the broader literature on public management reforms and pay for performance initiatives, focusing on the comparable features of PFS and past reform efforts. Next, we describe the data used in this study—including instrumentation developed to collect original data and archival data made available by the U.S. DOE—as well as the qualitative methods employed in coding and analyzing the data. We then discuss the distinctive features of the preschool PFS applications and relate them to what theory and an “ideal” SIB/PFS model suggest would be desirable in a viable program. We pull into this discussion the perspectives of grant applicants and their understanding of the strengths, weaknesses and potential of the PFS model for achieving their program goals. Lastly, we draw together the insights and findings from the application analyses and interviews on the current status of the preschool programs and feasibility pilots to consider the viability and sustainability of emerging preschool PFS models, as well as the wider lessons for other government entities heading down the PFS path.

How Prior Research and Experience Inform PFS Efforts

PFS origins

There is strong consensus in the literature to date that a key innovation of PFS mechanisms for funding and delivering public services is the role of the private investor in providing gateway funding to launch programs. Other core components—including public-private collaboration, performance evaluation, performance-based contractual arrangements for social services delivery, and other accountability provisions—have been features of governance and vigorously studied in public and business administration literatures for years. In fact, over 40 years ago, Guttman and Willner (1976: xii) called attention to dramatic new ways the U.S. government was managing its responsibilities and accomplishing its goals through an “invisible bureaucracy” of private, for-profit and nonprofit firms that were increasingly contracted to “suggest, shape, and even implement much governmental policy.” The ensuing New Public Management (NPM) reforms focused on changing organizational and individual incentives in ways intended to mimic the private sector and reorient toward a pay-for-results culture, for example, rationally linking employee efforts to their pay and other incentives for improving organizational performance (Moynihan and Pandey, 2005).

Christopher Hood (1991:4-5), oft credited with coining the term NPM, described it as a “shorthand name” for “a set of broadly similar administrative doctrines” that set out to bring about a more “business-like” government through the development of new public-private partnerships; the devolution and contracting out of public service responsibilities to local “quasi-public” and private partners and networks of organizations; and expectations for more responsive and accountable public services delivery. At the same time, Hood (1991:4) posed a reflective question: “What exactly was the public management Emperor now wearing... did its novelty

[NPM] lie mainly in presentation or in content?” Hood and Jackson (1991) had identified no less than 99 different administrative doctrines, tracing the origins of their NPM tenets to as far back as Jeremy Bentham’s writings of the late 18th and 19th centuries. They also argued, however, that it was the interdependence of these elements that was key to the novelty of NPM and its evolution toward more highly developed, formal systems of accountability with a focus on outcomes produced.

Insights from public-private, pay for results and performance-based contracting initiatives

Warner (2013:304-6) has similarly suggested about SIBs/PFS that it may be in combining philanthropy and performance management with venture capital and social program finance that we have founded an “innovative new mix” for public services delivery. She argues that the reliance of current SIB/PFS projects on “the contract and performance management scheme to cover all aspects of governmental interest in the project” ignores complexities of public sector contracting and performance management that are well-versed in prior research and the public administration literature. In a synthesis of the theoretical and empirical literature on performance management, pay for performance and performance-based contracting, Heinrich and Marschke (2010) pointed out the significant costs and capacity demands associated with developing and implementing accurate and reliable performance measures and performance-based contracts that align diverse public and private interests and reflect the complicated nature and “technology” of public programs. This, they argued, has led many public sector arrangements to rely on simpler contracting schemes with incentives based on highly imperfect or incomplete measures of primarily shorter-term performance goals. Moreover, research has consistently shown that shorter-term performance outcome measures are often loosely or sometimes even negatively correlated with longer-term program goals and impacts (Barnow, 2000; Heckman et al., 2002;

Schochet et al., 2006). Thus, if an intended advance of SIB/PFS arrangements is to reorient the public sector's focus toward future gains (Lantz et al., 2018) and to pay for *impacts* (as made explicit in the recent U.S. legislation), then SIB/PFS designers should be wary of compromising with shorter-term contract performance measures or erring in assuming their positive correlation with longer-term impacts.

Indeed, we have neglected lessons from past research and experience showing that rather than “hollowing out” the role of public managers (Rhodes, 2004), introducing new contractual structures that are technically and operationally complex—even if managed by external partners or intermediaries—intensifies the need for strong government management capacity (Kettl, 1993; Romzek and Johnston, 2002; Hefetz and Warner, 2004). Furthermore, in designing and implementing public sector performance management reforms, we have attended more to their formal structures and less to informal aspects of organization that are critical to their implementation, such as culture, values and “craft” or “skilled practice” (Hill and Lynn, 2009; Hood and Peters, 2004). In their recent synthesis of the literature on SIBs, Fraser et al. (2018) identified competing public and private values as a major, cross-cutting theme that poses challenges in implementing SIBs. Concerns about the (in)compatibility of the “logics and normative assumptions of the private financial services sector” (p. 8) with values traditionally held in the public domain (e.g., equity, justice, representation, etc.) have long been a subject of public and scholarly debate, particularly concerning the role of private financial interests and profits in social services delivery (Salamon, 1989). For example, in welfare services delivery in Wisconsin, private sector providers were allowed to earn profits in delivering public welfare services, both to encourage them to bid in operating the program (an argument comparable to that made for investors in SIBs) and ostensibly to reward them for innovations and efficiencies

achieved in service delivery (Heinrich and Choi, 2007). When private contractors underspent and realized large profits, the concern was raised that the contract terms may have encouraged the provision of fewer services to keep costs down (Wisconsin Legislative Audit Bureau, 1999). As Tse and Warner (2018:2-3) found in their study of early childhood education PFS projects, the balance of power across public and private sector actors matters importantly for the values that are ultimately prioritized in these contractual arrangements, as well as what the public gets and pays for in services delivery.

As SIB/PFS arrangements develop and unfold over a period of years, it is also important to consider the dynamic aspects of contracting and performance management (Heinrich and Marschke, 2010). That is, even if the government and its private sector partners have designed contractual and performance management arrangements (SIB/PFS or other) that initially reflect shrewd attention to these details, over time their effectiveness may decline as problems emerge due to the inability to fully specify in advance relevant contingencies (e.g., changing economic circumstances, legislative mandates, etc.) or to avert gaming responses that frequently emerge when pressures to meet performance targets or goals intensify (Bevan and Hood, 2006; Courty & Marschke, 2004; Heinrich, 2003; Tse & Warner, 2018). With SIB/PFS projects requiring an upstream investment to be repaid in a series of (financially complex) downstream payments over time (contingent on impacts being realized), the potential for unforeseen complications and deviations from plans may be even greater. Hence, while Arena et al. (2016) suggest that the wider adoption and success of SIBs has been slow because the instrument “has not found its dominant design yet—i.e., a stable set of characteristics related to the specific configuration, in terms of users, financial structure, public accounting arrangements, and scope of application”—the dynamics inherent in public sector contracting and performance management relationships

may preclude stability or even favor an instrument for funding social programs that is continually evolving and varying across contexts.

In fact, in our exploration of the U.S. preschool PFS feasibility applications and pilot efforts to date, we have observed considerable variability in the PFS partnerships and plans for executing a PFS/SIB, as well as steep learning curves on which many of the proposed projects were on a gradual and painstaking climb. Some of the projects have encountered what they see as insurmountable barriers to PFS success, while others have hit roadblocks or contingencies that they are still working to overcome. Many of the complex challenges the PFS partners face were foreshadowed in past experiences with public sector performance-based contracting and pay for performance systems. Indeed, this seems to be a consistent theme in the burgeoning SIBs/PFS literature: that many cautionary tales or narratives emerging in the implementation of these new social investment tools could have well been anticipated based on our substantial experience with many of the core elements of these instruments (Fraser et al., 2018). This begs again the question Christopher Hood (1991) posed about NPM reforms: “What exactly [is] the public management Emperor now wearing...?” Below, after briefly describing our data and methods, we turn to consider the distinctive features of preschool PFS pilot projects underway in the U.S. and how they relate to what theory and the “ideal” of a SIB/PFS model suggests would be desirable in these arrangements.

Study Data and Methods of Analysis

The U.S. DOE released a request for proposals to fund Preschool PFS feasibility pilots (closing date: October 6, 2016) and made the 20 applications received publicly available (with some proprietary and budgetary information redacted). They also made available to us the ratings and comments of three reviewers for each submitted proposal. Appendix A lists the applicants

(award recipients and those that were not funded), their locations, and whether they participated in interviews we conducted for this research; additional information on the applications that were funded is available at this link: <https://www2.ed.gov/programs/pfs/awards.html>. We entered all of the ratings (scores) for each of the following major proposal features (shown in greater detail in Appendix B): (1) the magnitude of need in the target population; (2) quality of the preschool program design; (3) quality of the existing or proposed PFS partnership; (4) quality of the work plan; (5) quality of the project leadership and team; (6) budget (adequacy of resources), and (7) competitive preference or absolute “priority” of the project based on potential outcome measures. We also developed a coding scheme to extract other information from the feasibility study applications that SIB/PFS models and relevant literature indicated are critical components or building blocks of a successful PFS. These included the public-private partnership composition and setting; the program model and its design and quality elements; the target population, indicators of need and plans for serving them; program outcomes and the research design, and budgetary information (see Appendix C).

In addition, we developed a survey instrument, grounded in the SIB/PFS literature, to use in interviews with applicants (funded and unfunded). The IRB-approved copy of the instrument is shown in Appendix D and covers the following main topics: (1) motivation for pursuing a PFS arrangement; (2) the application process and perceived strengths and weaknesses of the application; (3) project partners; (3) the preschool program model; (4) the determination of project roles (if funded); (5) evidence used and project evaluation activities, and (6) prospects for success and plans for achieving the program goals. We reached out by phone and email to all 20 applicants, contacting the persons identified as the project leaders in the applications, to schedule interviews. We secured the participation of 12 applicants, five of eight who were awarded

feasibility pilot grants, and seven of 12 who did not receive an award (see again Appendix A). The timing of the interviews, June-August 2018, was approximately 18 months following the funding award decisions. The interviews typically lasted one hour and were taped and transcribed for analysis, although we do not name respondents in our presentation of the findings. Confidentiality was assured to encourage project leaders to reflect candidly on their experiences in implementing their preschool programs and pilot projects.

Our primary approach to the analysis involved triangulating information from the three main sources of qualitative data: the information extracted from the feasibility pilot applications, the reviewer ratings and comments on the feasibility pilot applications, and our notes and transcriptions from the interviews. For example, comparisons were made between more and less highly rated applications on different PFS project dimensions, as well as between funded and unfunded applications,ⁱⁱ and we then drew on information from the interviews and applications to better understand the nature and context of the observed differences and to illuminate the varying features, plans and experiences of the projects. Information was also drawn from the interviews and compared across projects with different features and plans to address our key study questions about how the observed PFS pilot components aligned with model “ideals,” what the grant applicants saw as promising about the PFS strategy for achieving their short- and long-term preschool program goals, and the prospects for success and sustainability of the feasibility pilots or the alternative models they were pursuing.

Do PFS Pilot Features “Live Up to” the Model “Ideal”?

The text of the U.S. Bipartisan Budget Act of 2018 (PL 115-123) that describes key provisions of PFS calls for: the use of federal funds for “vital social service programs” that “produce positive outcomes for both service recipients and taxpayers”; the use of “objective

data” to verify that programs are achieving “demonstrable, measurable results”; the incorporation of “outcomes measurement and randomized controlled trials or other rigorous methodologies for assessing program impact,” and the creation of public-private partnerships—bundling public spending with philanthropic or other private resources—to scale up effective social interventions already being implemented by government and private organizations. We relate these PFS legislative goals to the U.S. DOE’s review of major features of the Preschool PFS feasibility pilot grant applications, focusing primarily on the three most highly weighted components—the quality of the preschool program design, the PFS partnership, and work plan (25 of 105 total points each)—which were accorded more than 70 percent of the total possible points (see again Appendix B). The lowest weights (5 points each) were assigned to the quality of the project leadership and the priority designation (based on outcomes to be used in evaluation).

High-quality, evidence-based program design

As Tse and Warner (2018) point out, while it an overarching goal of SIBs/PFS to promote innovation and allow local flexibility, they require strong documentation or evidence of the likelihood to achieve results and high degrees of model fidelity to garner funding (Berlin, 2016). Indeed, the primacy of the quality of the preschool program design—reflecting federal intent to fund programs that are already known to be effective and thereby more likely to “produce positive outcomes for both service recipients and taxpayers”—came through resolutely in the feasibility pilot application reviews. All eight funded pilot applications specified that an evidence-based preschool curriculum would be used, with six of these eight naming the curriculum or required components. This contrasted with only half of the 12 non-funded pilot applications that specified the preschool curriculum would be research-based.

Furthermore, reviewer comments on the technical review forms indicated that they were reading carefully into the extent to which the curriculum components were already being enacted at the sites with fidelity or whether there was capacity to implement or expand them (e.g., teacher training and instructional staffing ratios, student screening and assessment, coordination with existing social services, culturally-sensitive programming and inclusive practices for students with disabilities, etc.). In addition, the reviewers were attentive to the proposed outcomes to be achieved by the preschool program and whether indicators for measuring changes in outcomes were clearly specified and measurable, satisfying “a requirement of data-driven analytical approaches” (reviewer comment). Every funded pilot application identified measures of student achievement, socio-emotional learning (SEL) and other student outcomes, whereas, only half of the non-funded applications described specific outcomes.

The guidance for the feasibility application reviewers indicated that the applicants should also explain *how* the measurable outcomes would be achieved by the program or intervention strategy, drawing on quantitative and qualitative or theoretical evidence. This aligns with the requirement of the U.S. Bipartisan Budget Act for the use of “objective data” in PFS projects to verify that programs are achieving “demonstrable, measurable results.” In the funded preschool PFS application that received full points on this dimension in all three reviews, reviewers commended the applicant’s high-quality benchmarks and identification of short, mid, and long-term outcome goals and measures, including kindergarten readiness (cognitive measures and their progression across three different time points), third grade reading proficiency scores, school attendance and on-time grade matriculation, as well as measures of children’s social and emotional development. Importantly, the applicant described the preschool program model components in considerable detail, along with the evidence base (and a logic model) linking the

use of specific, objective measures to each program goal (e.g., Peabody Picture Vocabulary Test, Behavior Rating Inventory of Executive Functioning – Preschool Version, etc.). Moreover, this and other applicant applicants were able demonstrate (via prior work) their capability with partners to match children’s pre-K records to school district data and develop a longitudinal database for tracking children and measuring their outcomes over time.

The preschool PFS applicants with more detailed and sophisticated descriptions of the program design and evaluation elements were also more likely to have solidified partnerships with intermediaries or evaluators and to have engaged in prior evaluation work with these partners. The top-rated applicant also had experience successfully leading a PFS feasibility project in another area of child welfare, and from our interviews, we learned they had received more than \$20 million in public and private funding to expand their preschool program (as they were preparing the application). This affirms the critical role of high levels of organizational knowledge and capacity for even developing an adequate application for a PFS feasibility pilot and demonstrating promise for executing it. It also raises the question of whether the resources provided in the recently passed U.S. Bipartisan Budget Act will be allocated to more vigorously support capacity building among new, as well as existing, organizational partners working toward a PFS, a role that the UK and other national governments have more actively assumed (Arena et al., 2016; Greenblatt & Donovan, 2012).

Creating public-private partnerships for PFS implementation

The 2018 Bipartisan Budget legislation specifically articulates the intent to *create* public-private partnerships that “scale up effective social interventions *already being implemented* by private organizations, nonprofits, charitable organizations, and State and local governments” [*emphasis added*]. In the feasibility pilot application reviews, reviewers were instructed to

evaluate the quality of an “existing Preschool PFS partnership, including the history of the collaboration, or, if a Preschool PFS Partnership does not exist, the quality of the plan to form a Preschool PFS Partnership,” as well as the extent to which partner roles and responsibilities were clearly described and appropriate. While the language of the legislation and review form clearly conveyed support for the formation of new PFS partnership structures and relationships, the pilot application reviewers identified as weaknesses the lack of an existing partnership and rated more poorly PFS partnership quality when elements of the PFS partnership arrangements (e.g., an investor, intermediary, evaluator or payor) were not identified.

For example, a feasibility pilot application with one of the lowest ratings for the partnership dimension described existing partnerships between the city, the early childhood program offered in its public schools, local and regional foundations, and a women’s caucus that had successfully sponsored legislation to establish an office of early childhood in the city. One of the reviewers noted that the partners had a history of working together on early childhood initiatives, and our interview with the applicant affirmed the city’s continuing, strong commitment and leadership in working toward the goal of universal pre-kindergarten (preK), including a successful tax referendum, a strategic plan for implementation, and new funding for quality improvement of area preschool programs. In explaining their low ratings of the application on this dimension, one of the reviewers suggested that the PFS partnership lacked detail on the roles of each of the partners, and another pointed out that the mayor’s letter did not commit the city to being the payor (or sharing the responsibility for paying for successful program outcomes). However, even among funded feasibility pilots (following more than a year of pilot work), it was rare that the payor had been identified.

In reality, among the eight feasibility pilot applications selected for funding, only three proposals identified a “likely” or prospective payor. One of these three funded applications identified a school district as a possible outcomes payor and indicated that the district’s commitment to being the payor would be contingent on a “rigorous” demonstration of the program’s effectiveness in improving child outcomes. The highest rated application did not identify a payor in advance, and in our interview, the pilot PFS director stated that she thought the PFS had a 5-10 percent chance of succeeding, with the biggest challenge determining who would be the payor. A team member of a funded feasibility pilot operating at the state level indicated that they likewise were struggling with who should be the payor—would the savings be state-wide or local? Given the highly regulated nature of funding mechanisms at the state level, it wasn’t clear that there would be a mechanism for the state to make payments to an investor. Alternatively, another project leader suggested that because of their state’s proportionately high level of contributions to special education funding, the state was the only logical (prospective) payor. Still, the director of one of the funded feasibility pilots that had identified a payor in its application communicated pessimism in an interview that the arrangement would work out. They were finding through their modeling that the anticipated outcomes payment would be too high for the intended payor (a school district).

In addition, only three funded feasibility pilots and none of the unfunded sites identified an intermediary in their applications, and of those that did, it was the same intermediary (organization) for each of the pilots. Where specified in the applications, it was generally expected that the intermediary would guide the design and management of the feasibility study and support its completion, providing expert consulting in implementing the PFS concepts and components. Thus, the intermediary is intended to play a central, coordinating role in the

development and execution of a PFS pilot, yet more than a year into the feasibility pilot grants, few funded pilots had yet confirmed or contracted with an intermediary. One of the funded applications stated that they would contract with an experienced intermediary to conduct the feasibility study through a “competitive procurement process,” yet the same pilot program leader commented in our interview on the lack of options or choice in experienced intermediaries (i.e., making it difficult for procurement to be competitive). A pilot program director that had considered the same intermediary used by the three pilots with a contracted intermediary indicated the quote for managing their PFS contract was “cost-prohibitive,” and they were struggling to find an affordable alternative. These circumstances prompt another important question about the U.S. federal government role in supporting SIBs/PFS—would the pilots and other prospective PFS projects be more successful if the government provided more of the capacities or technical assistance of an intermediary, such as the Government Outcomes (GO) Lab does for SIBs in the United Kingdom?

The other key partner in the essential infrastructure of a SIB/PFS is the independent evaluator. A hallmark of SIB/PFS evaluations (as stated in legislation) is the intention to employ the most rigorous methods possible to evaluate the “impact” of the intervention, preferably using randomized controlled trials (RCTs) or other rigorous methodologies that allow for causal attribution of the measured outcomes to the intervention. Five of the eight funded Preschool PFS applications had identified a likely organization to serve as the evaluator, while only two of the 12 that were not funded specified an evaluator. However, among both funded and unfunded applications, the references to the types of methods to be used were vague, with about half noting the aspiration to conduct a cost-benefit analysis, but only one referencing the potential to conduct

a RCT. The word “impact” was also used loosely in the applications and applicant interviews, including in reference to simple regression methods without a control or comparison group.

As Berlin (2016) pointed out, an important distinction or innovation of SIBs/PFS projects is that the payor only has to make payments if the intervention is determined to be successful by the independent evaluator. While this ostensibly reduces the risk for the government payor, the investor’s repayment is made contingent on the evaluation results, increasing the level of risk for the investor. As Davison and Heap (2013) deduced about SIB arrangements in the UK, the SIBs that are most attractive to investors are those that look more likely to generate large and *reliable* returns, rather than riskier, more innovative projects. And one strategy to realize more secure returns is to specify simpler, easier to achieve outcome measures along with methods that, in Berlin’s (2016, p. 2) words “may misleadingly provide only the illusion of benefits and savings to government entities.”

Plans to make the PFS work

As the intermediary plays a central role in implementing a PFS feasibility pilot, identifying an experienced intermediary and describing the intermediary responsibilities were key factors in reviewer ratings of the Preschool PFS work plans. The sum of the three reviewer scores on work plan ratings were on average more than 22 points higher (of 75 total) for applications that had named an intermediary (confirmed or prospective).ⁱⁱⁱ Similarly, work plan ratings were on average 15 points higher (also a statistically significant difference) if an independent evaluator had been identified in the application. Reviewer comments suggested that they were looking for a clear delineation of responsibilities between intermediaries, evaluators and other program partners, along with a well-developed timeline of proposed program activities. Having an intermediary and evaluator on board appeared to aid in solidifying these aspects of the

work plan; in fact, for the three funded applications using the same intermediary, the work plans and logic models included were nearly identical (suggesting “boilerplate” materials provided by the intermediary). Key program activities to be undertaken in the feasibility pilots (as articulated in the federal guidance and pilot applications) included: identifying a funder and funding streams to support the project; legal and regulatory assessments of the performance-based contract terms; determining the accessibility and quality of data for the project; aligning core program outcomes and the pricing/payment rates of the end payor; developing a rigorous evaluation methodology to determine if outcome(s) have been achieved; data analysis and modeling to estimate financial benefits and cost savings generated by the preschool program; assessing community needs, assets, and capacity; and specifying opportunities for stakeholders to have regular and meaningful input toward achieving the project goals.

Among the above tasks, perhaps most distinctive to the undertaking of a PFS are those that involve modeling the financial benefits and cost savings and defining the terms or pricing for the end payor in relation to program outcomes achieved, in part because this is typically projected for a longer term or stream of payments than a typical (e.g., annual or shorter-term) performance-based contract. As indicated in the U.S. Department of Education’s (2017) “Tool Kit” for guiding feasibility studies, these tasks are critical to determining if a PFS project is viable and to moving on to the next phase, where the transactions are structured and formal agreements to implement the project and the evaluation are negotiated. Project leaders working with an intermediary indicated that the intermediary assumed a major role in the economic modeling and pricing of outcomes, although none had progressed to the point of finalizing a financial model with outcome targets, the mechanism for assessing costs and benefits (in dollars) of achieving the outcomes, and the timing of the costs, benefits and payments. The U.S.

Government Accountability Office (2015) asserts that because the government makes payments based on the service provider meeting the outcome targets, all parties to the PFS contract should have a vested interest in the service provider's success. This should also be the case, however, for any performance-based contract where the government pays a service provider for outcomes achieved, which begs the question: what are the distinctive advantages of developing these formal PFS partnerships, undertaking the onerous planning, and negotiating these considerably more complex PFS contracts to execute projects that the evidence base already suggests are likely to generate large returns (Fox & Albertson, 2012)?

One potential advance of PFS arrangements is their encouragement of a focus on longer-term or future public sector gains (Lantz et al., 2018), which could help to resolve a long-standing public sector performance management problem known as the "time-shortening disease" (Bouckaert, 1993; Heinrich, 2003). This is the phenomena in which policymakers or public managers focus on shorter-term, more readily measurable (and achievable) performance goals to the neglect of intermediate or longer-term goals, in part because of political pressures to demonstrate immediate results, but also because it is often difficult to translate longer-term objectives into performance goals that can be readily tracked for performance monitoring and accountability. In multiple interviews with both funded and unfunded Preschool PFS applicants, we heard from program implementers about how the PFS opportunity spurred them to consider longer-term outcomes, such as 3rd grade reading scores, which one program leader speculated would be on the list of outcomes of all PFS pilot plans. Another pilot program representative indicated that they were considering an eight-year re-payment timeline for their program, but they were struggling to find studies on their evidence-based preschool model that had examined longer-term outcomes (to aid in developing the payment/pricing plan). A pilot applicant that was

not funded reported that they were still able to make progress toward examining outcomes throughout a student's whole public school career via new infrastructure they built to track children from age three years to 3rd grade; the new technology and assessments replaced a system where data previously "lived in different silos" and children's records had to be linked one by one. Still another pilot program leader indicated that even if the study determined that PFS financing was not feasible, the analysis they were undertaking with the intermediary would be useful for assessing what outcomes need improvement, how the program is advancing those outcomes, and how the program might be expanded to serve more children in need.

At the same time, across our interviews, there were no funded or unfunded program representatives who expressed strong confidence they would reach the point of successfully establishing a formal PFS agreement. A leader of one of the funded pilots—which was working with an experienced intermediary and evaluator and had progressed further in their feasibility study than most programs—described the need to evaluate, price and pay for outcomes over a longer time horizon as a major problem for both providers and investors. He explained the structure of the PFS payment arrangements created "cash flow issues" for providers and put investors' money at risk for a longer period. In addition, the handful of projects that had progressed to estimating program benefits and cost savings were not obtaining numbers that they believed were sufficient to attract or secure payors, where the anticipated payors were school districts or local or state government entities. One interviewee expounded, "if you don't have an end payor, you've got nothing." Four of the project leaders interviewed described state-local public school financing rules and mechanisms as a barrier to establishing viable payment plans and rates. "School districts tend to be quite conservative and quite risk averse," explained a project leader. Another lamented that in a program covering more than 30 school districts with

children that “go from district to district,” mobility is a challenge, “because no one wants to pay for a kid who isn’t going to be in their district.” Several project leaders also described the use of “braided” (state-local-private) funding streams for operating existing preschool programs that complicated or precluded efforts to develop PFS re-payment mechanisms using public funds.

In light of these findings, we turn now to consider: if not a formal SIB/PFS arrangement, will these program sites or partners still implement *something* like PFS? What have they learned from their experiences that they will take forward in their continuing efforts to expand high-quality preschool programming in their states or local areas?

Is There Promise for PFS in the U.S.?

In a July 2018 UK GO Lab report, Carter et al. illuminate how in the approximately 40 SIBs that have launched in the UK, “practitioners are pushing the boundaries of the SIB concept to try out new variants and new factors which might work better to improve outcomes”(p. 2). The report describes, for example, how SIBs vary in the extent to which they are preventative-focused and take advantage of the flexibility allowed to adapt, innovate or improve interventions, as well as in their access to data and the types of performance measurement they are able to support. Thus, even in the UK where the rollout of SIBs has been more centralized, standardized and better supported with technical assistance, they are finding that “local context matters greatly” for implementing this new public-private collaboration on outcomes-based contracting (p. 2). At the same time, Carter et al. point out those studying SIBs are failing to grasp these opportunities to go beyond questions (and rhetorical debate) about whether SIBs work to ask where and why they work well (or don’t work). While the U.S. has about half the number of SIB/PFS projects underway compared to the UK, we draw out some general and context-specific lessons that we see emerging from the U.S. DOE Preschool PFS feasibility pilots, as well as

from the applicants that, even without grant support, continued to work toward similar goals for expanding high-quality preschool programs through public-private collaborations.

Is PFS seen as a sustainable source of program funding?

The primary motivation for pursuing a PFS pilot among project leaders was to secure an additional source of financing for preschool expansion and quality enhancement efforts.

Consistent with speculation (Roy et al., 2017) that SIBs/PFS were spurred in part by financial constraints intensified by the deep recession, one project leader described his state's slow, incremental progress in increasing the number of preschool seats, stating: "we have finally kind of clawed our way back to where we were a decade ago before the Great Recession, and so we are looking for alternative ways to expand access... it [PFS] is a very complicated way of paying for a pilot program, but ultimately it's not going to be a source of sustained funding." The hope in this case was that they could demonstrate to school districts that the benefits of expanded preschool would outweigh the costs, and that PFS "would essentially be a transition step towards either a school district program or perhaps a school district contracting with a provider to operate a preschool program on either sort of a regular fee for service basis, or perhaps if they were feeling particularly innovative, on sort of an outcomes-based contract." In explaining why they did not see PFS as a sustainable source of program support, the project leader pointed to the "unavoidable leakage of money out of the system because you are paying a risk premium to the investors." He elaborated that with a simpler fee-for-service or outcomes-based contracting arrangement, the money "that's leaking out to the investors could be captured as a benefit to the school district." This insight resonates with Mulgan et al.'s (2011) argument that because the cost of capital is higher for private investors than government, this type of partnership may ultimately prove to be more expensive and harder to justify with adequate public returns.

Another project site that had prior experience operating a PFS similarly suggested that PFS did not hold “any particular allure other than that it was an option that we could pursue” for additional funds, adding that it was “way more complex” and “way more problematic” than the other public-private funding streams they were relying on to improve and expand public preschool options. Given the option, the project leaders would prefer more funding through their current public and private (foundation) collaboration, which they described as robust and not without its own accountability provisions. They also did not see that PFS would change the nature of their relationship with their current public-private partners, nor did they believe (as suggested in the literature) that PFS would “free service providers” to focus on program design and service delivery innovations and improvements (Fox & Albertson, 2011; Warner, 2013; Jackson, 2013). Another project leader that likewise indicated they were “just looking for additional sources” pointed out that in their district, they had to raise funds through community bond referendums, and that these efforts were more likely to raise funds for infrastructure (i.e., buildings) than for other needs associated with program expansion. As they were serving mostly special needs students, the attraction of applying for a PFS pilot was to expand the number of children they could serve with a program that they believed was already “the best intervention that a child with special needs can get before going to kindergarten.”

In our 12 interviews with preschool program leaders, we asked them to consider the level of confidence they had that their proposed feasibility pilot could lead to a successful PFS arrangement (from 0-100%), both at the time they submitted their proposal (earlier in the interview) and at the current time (toward the end of the interview). Across both funded and unfunded pilot PFS applicants, responses ranged from a low of 5-10 percent (funded applicant) to a high of 80-100 percent (unfunded applicant). Only one funded applicant felt more confident

after undertaking the pilot (an increase from 50% to 60%) that they would progress to a formal PFS arrangement. Two of the five funded interviewees did not think it would work, one stayed at the very low 5-10 percent probability, and the other felt confident that they would come up with some kind of feasible model, but they weren't sure what it was going to look like. On average, the Preschool PFS applications that were not chosen for funding were more optimistic about their prospects of ultimate PFS success at the time of application (average probability of 56%) than those selected (42%). The two applicants that conveyed the most knowledge and experience with PFS—both funded and with prior experience planning or implementing PFS projects—rated their chances of success the lowest at the onset (5-10% and 25-33%) and were not more confident of success 18 months later. These perceptions of the prospects of PFS success were consistent with Williams' (2018: 14) findings from interviews with 178 members of core groups of SIB actors in the U.S., Canada and the UK: "...many respondents expressed uncertainty, and in some cases downright skepticism, about the future of SIBs." Likewise, Ronicle et al. (2017) reported that of 62 projects that received grants in 2013 from the English Commissioning Better Outcomes (CBO) Fund, 37 had decided by 2016 that they would not continue to pursue a SIB.

Insights on the viability of PFS for U.S. programs

A comment frequently heard in our 12 interviews is that PFS is "very hard", particularly "to get all the players onboard" and to make a convincing case of the community benefits that would attract a public payor. One leader of a project that was set in a more rural setting explained that "when people are wearing multiple hats and trying to make things work as they currently work, it's more difficult to get people to kind of think outside the box and be able to devote significant time to a complicated [PFS] concept when you don't know if all that time and energy is going to be fruitful." The steering committee for this (funded) project implored: "if

we're going to move from a feasibility study to a pilot, we're not only going to need more money, we're going to need more people hours, staff hours, and so, where do you see this as being feasible?" This particular project is planning to move forward by scaling up their early learning program with area philanthropic organizations—*not* in a PFS arrangement—with the aim to convince other potential funders and school districts to come on board with the idea that “we don't have to reinvent the wheel entirely, but we just have to find a way to fund it and expand it here”. An urban-based project similarly described challenges in trying to bring the players and components to the table to make PFS work, but for the very different reason that “it's all fragmented among agencies and the school districts,” in terms of the infrastructure for implementation and data for tracking student outcomes over time. More generally, the experiences of the pilot applicants to date strongly confirm the wisdom gleaned from past pay for results and outcomes-based performance contracting efforts; that is, the need for government management capacity (and that of its partners and service providers) will intensify, not be relieved, in context of more complicated performance-based contracting (SIB/PFS) arrangements (Hefetz and Warner, 2004; Heinrich and Marschke, 2010).

While the PFS feasibility study applicants were encountering a steep learning curve for numerous project tasks, many saw the work involved as an invaluable opportunity for pushing forward on the challenging fronts of aligning program goals with longer-term outcome measures, developing rigorous evaluation methodologies, and attempting financial modeling of program benefits and costs. For example, one project leader described how in the process of developing the application, they dug into past years to see if they could get multiple data points on various child outcomes and realized how difficult it was (i.e., there was better data in some years, and some were hard to find at all). This led them to expand their assessments and generate new

higher-quality, more consistent data collection from age three through third grade. Another project leader described their team as being “completely submerged in the data analysis,” with the goal to use their analysis as a starting point in discussions with different stakeholders about the types of outcomes that they might track for students. A leader from one of the projects that was further along on the learning curve described a series of memos they had developed with their evaluator to guide their work on PFS work tasks, including a memo that outlined their “potential short, intermediate, and long term outcomes that may or may not be monetizable” and another that “looks at all of our identified potential outcomes and determines whether or not there is a monetary benefit, where that benefit exists, how big is the benefit, and kind of the timeline for a potential investor.” This process also illuminated concerns they had about the pilot’s viability, such as the holes in their logic model that sought to link anticipated changes in longer-term outcomes to the intervention and a lack of data available to measure some outcomes.

Another applicant that was not funded described the request for proposals for the Preschool PFS feasibility pilot as a “catalyst” for them to think about whether this might be a feasible financing option for universal preK. They felt that even though the pilot wasn’t going forward, the time allocated for strategic planning (around a 5-year plan for early childhood) and for exploring how they could scale the preK program added a lot of value. Like other project leaders, they also saw issues of political will and values arise as they worked to cultivate partnerships in the community and build a stronger commitment to preK. Among their biggest challenges was how to bring the school district leadership on board, a hurdle that was echoed in interviews with a number of project leaders. One project leader described how they had progressed very far in their feasibility study, only to see it fall apart because the school district refused to provide essential longitudinal data—“it was kind of like the project was dead at that

point.” Another project leader pointed to school districts’ “very vanilla” approach to contracting and lack of experience with performance-based contracting as a barrier to committing to the PFS partnership—“they tend to be extremely conservative; there really aren’t many rewards for being innovative.” Still another project leader described the potential for these issues to be “politicized”: “You know, there’s the question of should an investor benefit from societal ills, so to speak.” In several states, however, the school funding formula made it difficult for school districts to realize enough of a cost savings to pay for the program; thus, in the absence of a state funding formula change, there was no way for PFS to work in these school districts.

The leader of an unfunded project that similarly felt they were highly constrained financially to implement PFS due to the state’s public school financing model described how they persisted in their efforts to expand their coalition of supporters for public preK and to build and broaden the political will (across the political spectrum) to implement PFS. The program leaders actively cultivated the support of local and wider area business leaders; they continued to work with the local school district to expand preK funding and classroom seats; and they also sought to bring on new partners who could add capacity for evaluation, for developing financial models, and for securing PFS investors. They believed that as they increased the rigor of their methods and measures for tracking impacts over time and deepened and broadened their partnerships, they could convince the state (based on their results) to change its funding formula and make PFS viable. Another interviewee echoed the importance of “bold action,” exhibited by one of their partners in putting “political capital on the line” in a tight budget cycle to set aside funds for the project and challenging the private sector to match them. The “projects rise or fall on the political will that is behind them,” exhorted the intermediary for this project.

If not PFS, then what?

Even projects that encountered what they saw as insurmountable barriers to moving forward with a formal PFS arrangement still appreciated the opportunity for “the tools to maybe look at our investments in a different way,” to “hone in on best practices and activities that produce better child outcomes” and to make progress in tracking and linking outcomes in early childhood to later academic outcomes. While one project that faced challenges in getting the school district to commit went “outside of the district” to try to develop a viable PFS project, other projects focused on opportunities to make improvements within their school districts (e.g., newly linking datasets and exploring new collaborative approaches to financing preschool expansion). Another funded feasibility pilot that was not going to pursue the PFS described how they were working instead to do a pilot of their preschool design to show “proof of concept” and thereby attract philanthropic funding or to say to the state: “We’ve got proof. It works here... Let’s duplicate this design across the state... let’s utilize state monies.”

One project leader explained that while they had “a very strong picture of what the overarching concept and goals were [of a PFS],” they had not gotten into the minutiae. Going forward, they were likewise taking the concept of a PFS as a “blueprint” for a project they are now implementing with public and private partners to track students from preschool into their college years. Their plan includes providing college savings accounts for young children when they begin school and a year of free community college tuition when they graduate. As nearly all of the PFS feasibility study applicants had existing collaborations with other government, community or nonprofit organizations, most were continuing to build on these partnerships and to apply for additional resources. One grant recipient that was not intending to further pursue the PFS indicated that they were invited to submit an application to a foundation for what they

described as “bridge money,” which will help them to continue to expand their “early learning hub” model to school districts identified through the pilot as interested in expanding preschool opportunities.

Finally, a number of PFS feasibility study applicants described the benefits of going through the application process or feasibility study for prompting them to take a careful look at the features of their existing preschool program models and their strengths and weaknesses relative to other models or standards, and for thinking in financial terms how they could “sell” the benefits of their programs to potential funders. One project leader explained that they were not able to fully satisfy the requirements of their state’s quality rating system (QRS) because they “didn’t have a mechanism to gauge what was considered quality,” and thus, an important part of their proposal for the feasibility study was to create this infrastructure for gauging their program’s quality against the validated model. Similarly, a grant recipients described the support from “data scientists” hired with the funding as “priceless” for tying data on individual children to student readiness and other outcomes. In this project, the intermediary worked with the project leaders to establish a data sharing agreement with the state, while for other project efforts, access to the administrative data continued to be a limiting factor in making PFS work.

Conclusion

In a research conference discussion in November 2015, panelists described how PFS presents opportunities to help “governments to maximize the services provided to those in need, while ensuring taxpayer dollars are allocated in the smartest, most efficient way.”^{iv} In the course of the discussion, one panelist stated that “government needs to get out of the ‘investing’ business and get into the ‘buying’ business, just paying for results and not worrying about how we get them.” The government was also described as “highly risk-averse” and “the limiting

partner in advancing Pay for Success.” While our exploration of the U.S. DOE feasibility pilot applications and studies substantiated that government regulations, funding mechanisms, restrictions on access to data, and aversion to financial and political risks sometimes slowed or deterred progress toward establishing formal PFS contracting arrangements, it also affirmed the critically important role that the government plays at all levels—federal, state and local—in making a PFS partnership and project feasible. In fact, the majority of the preschool program leaders—whether funded or not with a feasibility pilot grant—were actively engaged in strengthening collaborations between the public sector, nonprofit organizations (local or community-based and state-wide), and private, for-profit partners (potential investors, business leaders, etc.). Our investigation also suggested that not only is the government a key partner in making a PFS contract viable, but it also carries the primary responsibility to ensure that any public-private partnership is politically accountable to the public for the delivery of services (to vulnerable populations) and for upholding values such as social equity (Tse and Warner, 2018), as well as for the results produced. Not surprisingly, like the private investor that contributes funding, the government wants to minimize the risks of loss, where the public sector (potential) loss in a PFS arrangement is less likely to be in program monies, but rather in a significant investment of time and resources in structuring collaborations and capacity-building, planning and analysis, and fostering political will and support (all of which have opportunity costs).

Our findings from interviews with those attempting to make preschool PFS pilots work suggested that the projects were frequently not “taking off” for a range of reasons, including capacity constraints (resources, personnel and expertise); the difficulties of financial modeling and projecting returns over time (amply convincing) to get investors and payors on board; and the inherent complexities of the contract arrangements and negotiations that need to account for a

host of contingencies (e.g., political leadership and legislative vagaries, economic fluctuations, changing population needs, etc.) over a longer (multi-year) period. They were also largely falling short on the expectation to undertake *impact* evaluations for assessing project returns (and determining payments for outcomes), primarily due to data access and methodological design limitations. Although many of the preschool program leaders aspired to conduct a cost-benefit analysis (indicated in nine of 20 applications)—which is predicated on a rigorous assessment of impacts—only one project mentioned the possibility of conducting a randomized controlled trial, and the interviewees primarily discussed pre- to post-intervention calculations of changes in outcomes (with only one mention of a comparison group). There is clearly a need for greater resources and capacity-building in evaluation if prospective PFS projects are going to come close to the “ideal” of measuring and paying for longer-term *impacts*—not simply changes in outcomes—given that research has long shown that changes in outcome levels do not always correlate with program impacts (Barnow, 2000; Heckman et al., 2002; Fraser et al., 2018).

Our qualitative analysis also suggested that the capacity-building of preschool PFS pilots was aided greatly through collaboration with experienced intermediaries and evaluators, but only a small number of the projects (even with the U.S. DOE grant funds) had adequate resources to afford their services. A majority of the projects were actively engaged in fundraising activities and efforts to expand coalitions of support, some with the goal to contract with intermediaries and evaluators that could accelerate their progress toward a formal PFS arrangement. This raises the issue of whether the federal government should be taking a more proactive and generous (in terms of funding support) role, as has been done in the UK, to assist public-private partnerships at the state and local levels in developing these capacities, and not only for those working toward formal PFS arrangements. Given that our research and that of many others suggests that public-

private collaborations striving to improve or implement public program management features such as data collection on longer-term outcomes, performance and impact evaluation, and performance-based contractual arrangements for services delivery could benefit greatly from additional technical assistance and resources for capacity-building, we believe the U.S. federal government could—consistent with the current goals and provisions of PFS in the 2018 Bipartisan Budget Act^v—extend these supports more broadly to public and private organizations espousing these goals, regardless of whether a PFS was the ultimate aim. While an allocation of \$3 million by the U.S. DOE for supporting these types of capacity building activities was commendable, this minimal level of funding falls way short of the need and demand that currently exists for this type of support. Furthermore, the U.S. DOE application reviews had high expectations for existing infrastructure, capacities and plans for conducting a PFS pilot, which as Dominey (2012) suggested, favors larger and more experienced organizations, or those with a track record that will be more likely to minimize the private investors’ risk.

Finally, ongoing research on SIBs/PFS suggests that we are unlikely to see a “dominant design” emerge (Arena et al., 2016), as local context and political, social and economic dynamics continue to shape how SIB/PFS projects are structured and implemented, as well as their prospects for success (Carter et al., 2018; Fraser et al., 2018; Tse and Warner, 2018). Among the lessons from decades past about the challenges of implementing pay for performance and performance-based contracts was the need to attend more to the political and strategic management demands of public-private partnerships and the informal aspects of organization, which are often localized but critical to effective governance and program implementation, including culture, shared norms and values, and ethics. Attaining the goal of a formal PFS contracting arrangement might not be the best outcome of all exploratory or feasibility pilot

efforts, and in fact, a key actor in U.S. PFS efforts has suggested that the PFS pipeline “has dried up significantly” (Williams, 2018:14). Alternatively, supporting more public-private collaborations in building basic capacities for more effectively serving vulnerable populations and holding public-private partnerships accountable for longer-term, outcomes and impacts could be a major public management advance of these efforts.

References

- Arena, Marika, Irene Bengo, Mario Calderini, and Veronica Chiodo. 2016. Social Impact Bonds: Blockbuster or Flash in a Pan? *International Journal of Public Administration* 39: 927–39.
- Barnow, Burt S. 2000. Exploring the Relationship between Performance Management and Program Impact: a Case Study of the Job Training Partnership Act. *Journal of Policy Analysis and Management*, 19 (1): 118–41.
- Berlin, Gordon L. 2016. Learning from experience: A guide to social impact bond investing. New York: MDRC.
- Bevan, Gwyn and Christopher Hood. 2006. What’s measured is what matters: Targets and gaming in the English public health care system. *Public Administration*, 84, 517–538.
- Bouckaert, Geert. 1993. Measurement and Meaningful Management. *Public Productivity and Management Review*, 17 (1): 31–43.
- Carter, Eleanor, FitzGerald, Clare, Dixon, Ruth, Economy, Christina, Hameed, Tanyah, and Mara Airoidi. 2018. *Building the tools for public services to secure better outcomes: Collaboration, Prevention, Innovation*, Government Outcomes Lab, University of Oxford, Blavatnik School of Government.
- Courty, Pascal and Gerald Marschke. 2004. An Empirical Investigation of Gaming Responses to Performance Incentives. *Journal of Labor Economics* 22(1): 23-56.
- Curran, Amy Cobb. 2017. Inroads to Innovation: State Adoption of Pay for Success Legislation. Chapman and Cutler, LLP, April.
- Fraser, Alec, Tan, Stefanie, Lagarde, Mylene, and Nicholas Mays. 2018. Narratives of promise, narratives of caution: A review of the literature on social impact bonds. *Social Policy & Administration*, 52, 4–28.
- Fox, Chris and Kevin Albertson. 2011. Payment by results and social impact bonds in the criminal justice sector: New challenges for the concept of evidence-based policy? *Criminology and Criminal Justice*, 11, 5: 395–413.
- Greenblatt, Jonathan and Annie Donovan. 2012. The Promise of Pay for Success. *Community Development Investment Review*. Federal Reserve Bank of San Francisco, pp. 19-22.
- Gustafsson-Wright, Emily, Gardiner, Sophie, and Vidya Putcha. 2015. The potential and limitations of impact bonds: Lessons from the first five years of experience worldwide. Washington, DC: Global Economy and Development at Brookings.
- Guttman, Daniel and Barry Willner. 1976. *The Shadow Government*. New York: Pantheon Books.

Heckman, James J., Heinrich, Carolyn J. and Jeffrey Smith. 2002. "The Performance of Performance Standards." *Journal of Human Resources*, Volume 37, Number 4: 778-811.

Hefetz, Amir and Mildred Warner. 2004. Privatization and Its Reverse: Explaining the Dynamics of the Government Contracting Process. *Journal of Public Administration Research and Theory* 14(2): 171-190.

Heinrich, Carolyn J. 2003. "Measuring Public Sector Performance and Effectiveness." In *Handbook of Public Administration*, Guy Peters and Jon Pierre, (eds.), London: Sage Publications, Ltd., pp. 25-37.

Heinrich, Carolyn J. and Youseok Choi. 2007. "Performance-based Contracting in Social Welfare Programs." *The American Review of Public Administration* 37(4): 409-435.

Heinrich, Carolyn J. and Gerald R. Marschke. 2010. "Incentives and their Dynamics in Public Sector Performance Management Systems." *Journal of Policy Analysis and Management* 29(1): 183-208.

Hill, Carolyn J. and Laurence E. Lynn, Jr. 2009. *Public Management: A Three-Dimensional Approach*. Washington, DC: CQ Press.

Hood, Christopher. 1991. A Public Management for all Seasons? *Public Administration*, 69 (1): pp. 3-19.

Hood, Christopher and Michael W. Jackson. 1991. The New Public Management: A Recipe for Disaster? *Canberra Bulletin of Public Administration*, 64: pp. 16-24.

Hood, Christopher and Guy Peters. 2004. The Middle Aging of New Public Management: Into the Age of Paradox?, *Journal of Public Administration Research and Theory*, 14(3): pp. 267-282.

Jackson, Edward T. 2013. Evaluating social impact bonds: questions, challenges, innovations, and possibilities in measuring outcomes in impact investing. *Community Development*, 44:5, 608-616.

Kettl, Donald F. 1993. *Sharing power: Public governance and private markets*. Washington, D.C.: The Brookings Institution.

Lantz, Paula, Miller, George, Rhyon, Corwin, Rosenbaum, Sara, Ku, Leighton, and Samantha Iovan. 2018. "Pay for Success" Financing and Home-Based Multicomponent Childhood Asthma Interventions: Modeling Results From the Detroit Medicaid Population. *The Milbank Quarterly*, Vol. 96, No. 2, pp. 272-299.

Maier, Florentine and Michael Meyer. 2017. Social Impact Bonds and the Perils of Aligned Interests. *Administrative Sciences*, 7 (3). pp. 1-10.

Moynihan, Donald P. and Pandey, Sanjay K. (2005) 'Testing How Management Matters in an Era of Government by Performance', *Journal of Public Administration Research and Theory*, 15(1):421-439.

Mulgan, Geoff, Reeder, Neil, Aylott, Mhairi and Luke Bo'sher. 2011. *Social Impact Investment: The Challenge and Opportunity of Social Impact Bonds*, London: Young Foundation.

Romzek, Barbara and Jocelyn Johnston. 2002. Effective Contract Implementation and Management: A Preliminary Model. *Journal of Public Administration Research and Theory*. 12(3): 423-453.

Ronicle, James, Alec Fraser, Stefanie Tan and Catie Erskine. 2017. *The LOUD SIB Model: Four factors that determine whether a social impact bond is launched*. England: Big Lottery Fund report.

Roy, Michael J., Neil McHugh, and Stephen Sinclair. 2017. Social Impact Bonds: Evidence-based Policy or Ideology? In *Handbook of Social Policy Evaluation*. Edited by Bent Greve. Cheltenham: Edward Elgar Publishing, pp. 263–78.

Salamon, Lester. 1989. *Beyond Privatization: The Tools of Government Action*. Washington, DC: The Urban Institute Press.

Schochet, Peter, Burghardt, John, and Sheena McConnell. 2006. National Job Corps study and longer-term follow-up study: Impact and benefit-cost findings using survey and summary earnings records data. Washington, DC: Employment and Training Administration, U.S. Department of Labor.

Tse, Allison E. and Mildred E. Warner. 2018. The razor's edge: Social impact bonds and the financialization of early childhood services. *Journal of Urban Affairs*, DOI: 10.1080/07352166.2018.1465347.

U.S. Government Accountability Office. 2015. Pay for success: Collaboration among federal agencies would be helpful as governments explore new financing mechanisms (GAO-15-646), Washington, DC.

Williams, James W. 2018. Surveying the SIB Economy: Social Impact Bonds, 'Local' Challenges, and Shifting Markets in Urban Social Problems. Working paper, York University, Department of Social Science.

Warner, M. E. (2013). Private finance for public goods: Social impact bonds. *Journal of Economic Policy Reform*, 16, 303–319.

Wisconsin Legislative Audit Bureau. 1999. *A Review: Wisconsin Works (W-2) Expenditures*. Madison: Legislative Audit Bureau.

Appendix A: U.S. Preschool Pay for Success Applications

Preschool Pay for Success Project Location	Applicant/Project Leaser	Award Status (2016)
Austin, Texas	Austin Independent School District	Not funded
Clatsop County, Oregon	Clatsop County and Northwest Oregon Kinder Ready Collaborative	Funded
Cuyahoga County, Ohio	Cuyahoga County Office of Early Childhood	Funded
Durham, North Carolina	Durham County	Not funded
Greenville, South Carolina	Legacy Charter School	Funded
Las Vegas, Nevada	City of Las Vegas Department of Youth Development and Strong Start Academy	Not funded
League City, Texas	Clear Creek and Hitchcock Independent School District	Not funded
Mecklenburg County, North Carolina	Mecklenburg County Government and Charlotte-Mecklenburg Schools	Funded
Napa Valley, California	Napa Valley Unified School District and Napa County Office of Education (NCOE)	Funded
New York State	New York State Office of Children and Family Services and Council on Children and Families	Not funded
Pittsburgh, Pennsylvania	Office of Early Childhood, and Citiparks	Not funded
Racine County, Wisconsin	Higher Expectations for Racine County and Racine County Public Schools	Not funded
Rio Rancho, New Mexico	Shining Stars Preschool	Not funded
Santa Clara County, California	Santa Clara County Office of Education	Funded
State of Colorado	Colorado Department of Human Services, Office of Early Childhood	Not funded
State of Hawaii	Office of Hawaiian Affairs and Institute for Native Pacific Education and Culture	Not funded
State of Minnesota	Minnesota Department of Education and school districts	Funded
State of Oklahoma	Oklahoma Department of Education	Not funded
Ventura County, California	Ventura County Office of Education and First 5 Ventura County	Funded
West Sacramento, California	Early Learning Services for the City of West Sacramento and Universal Preschool for West Sacramento	Not funded

Appendix B: Categories and Guidelines for Scoring PFS Feasibility Study Applications

- *Need for Project*
 - Applicants should clearly state and demonstrate the extent of the problem facing the Target Population using data and other relevant information.
- *Quality of the Preschool Program Design*
 - In responding to this criterion, applicants should identify clearly specified and measurable outcomes for the preschool program and explain how these outcomes can be achieved by the program.
- *Preschool PFS Partnership*
 - The quality of an existing Preschool PFS Partnership, including the history of the collaboration, or, if a Preschool PFS Partnership does not exist, the quality of the plan to form a Preschool PFS Partnership.
 - The extent to which the roles and responsibilities of members or proposed members of a Preschool PFS Partnership are clearly described and are appropriate and sufficient to successfully implement a PFS project.
- *Quality of the Work Plan*
 - The adequacy of the work plan to achieve the objectives of the proposed Feasibility Study project on time and within budget, including clearly defined responsibilities, timelines, and milestones for accomplishing project tasks on time.
 - The adequacy of procedures for ensuring stakeholder feedback in the operation of the proposed Preschool PFS Feasibility Pilot. If the Feasibility Study includes the reduction in special education placement as a Financial Benefit, the extent to which the work plan includes outreach to and involvement of the representatives from the State and local special education community or individuals with special education expertise, including groups representing families.
 - The extent to which the time commitments of the project director and team and other key project personnel are appropriate and adequate to meet the objectives of the proposed project.
- *Quality of the Project Leadership and Team*
 - The Secretary will consider the quality of the project leadership and team. The Secretary will consider the extent to which the applicant has the project and financial management experience necessary to manage the Preschool PFS Feasibility Pilot.
- *Budget Narrative*
 - The Secretary will consider the adequacy of resources necessary to complete the Feasibility Study, including any philanthropic or other resources that may be contributed toward the project. In determining the adequacy of resources, the Secretary will consider the extent to which the budget will adequately support program activities and achieve desired outputs and outcomes.
- *Competitive Preference Priority*
 - To meet this priority, an applicant must propose a Feasibility Study to evaluate if PFS is viable that would evaluate social and emotional or Executive Functioning Outcome Measures, or both. These potential outcome measures may be predictive of future school success, cost savings, cost avoidance, and other societal benefits, and may be appropriate to include in a PFS project.

Appendix C: Researcher Codebook for Qualitative Analysis of Feasibility Study Applications

- *Team Composition*
 - Applicant (name)
 - Project leader affiliation
 - Program delivery operator
 - Delivery method (public, private, or both)
 - Intermediary
 - Other Stakeholders
- *Setting Characteristics*
 - Setting (urban, suburban, rural)
 - Currently serving (number of preschool seats)
 - Anticipated need (anticipated number of new seats)
 - Existing resources: other
- *Target Population*
 - Eligibility criteria: age, poverty, English as a Second Language
- *Quality Elements*
 - Existing program expansion (Y/N)
 - Plan to develop new program (Y/N)
 - Research based curriculum (Y/N and name)
 - Research based quality indicators (Y/N and name)
- *Outcomes*
 - Finalized measures (Y/N)
 - Achievement measures (Y/N)
 - Social-emotional measures (Y/N)
 - Other school outcomes
- *Research Design*
 - Data sources
 - Methodology identified (Y/N)
 - Independent evaluator identified (Y/N and name)
- *Budget*
 - Amount requested
 - Amount received

Appendix D: Interview Guide

Section I: Questions for both winning and losing SIB feasibility pilot applications

Part A: Motivation

1. When did your organization first consider pursuing a Social Impact Bond as a means of financing your program and activities?
2. What reasons motivated you to develop a proposal for a Social Impact Bond feasibility pilot through the U.S. Department of Education?
3. Do you see the Social Impact Bond approach as a less risky approach to financing your program operations?
 - a. How do you see your level of risk in comparison to the SIB investor in this project? In comparison to the SIB intermediary?

Part B: Grant Application and Planning

4. How much confidence did you have at the time you submitted your proposal that the proposed feasibility pilot could lead to a successful Social Impact Bond working arrangement? Please choose among the following:
10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
5. What did (or do) you see as the strongest aspects of your Social Impact Bond feasibility pilot application?
6. What did (or do) you see as the weakest aspects of your Social Impact Bond feasibility pilot application?
7. What do you think is the most innovative or transformative feature of Social Impact Bonds?
8. How did you choose your evidence-based model for this Social Impact Bond project?
9. What other models did you consider? Are you confident in the evidence underlying your chosen model?
10. Do you plan to continue with this evidence-based model?

Part C: Partnerships

11. How did you identify your proposed project partners for the Social Impact Bond proposal?
12. Do the types of partners you considered or identified look any different than the collaborations you have formed in the past to conduct your work?
13. Did you identify a potential investor for your Social Impact Bond proposal? How did you go about this, or did an investor approach you to develop the proposal?
14. What do you see as most innovative about the Social Impact Bond approach to public-private partnerships (if anything)?
15. In what ways does (or would) a Social Impact Bond change the nature of your relationship to your project partners, compared to the typical ways you arrange contracts for service delivery or other program operations?
16. Did you select an independent evaluator for the Social Impact Bond project?
17. Have you used independent evaluators previously to assess the effectiveness of the work in your organization?

18. (If yes to both .16 and .17): Have you worked with this particular evaluator previously at your organization?

Part D: Preschool Program

19. Do you currently offer a public preschool program (aside from Head Start)?
- a. Where are preschool programs offered? (i.e., center-based, school-based, etc)
20. Of the students enrolling in public kindergarten in your area, what is the approximate proportion of students who have attended public preschool, Head Start, or private care?
21. Was there an existing public-private partnership to support/provide preschool prior to the Social Impact Bond project?
- a. What were the goals of that partnership?
22. Some national public preschool evaluations find that preschool intervention effects “fade out” by third grade.
- a. At this point, have you been able to track or determine any medium- or long-term outcomes?
 - b. What do you see as the biggest challenge facing your current preschool program, in terms of its effectiveness towards long-term outcomes?
23. What is the estimated capacity of your current preschool program, in terms of classroom space, staff, and expenses?
- a. How was local capacity taken into account when undertaking the SIB feasibility proposal/study?

Section II: Questions for winning proposals only

Part E: Determining and Assigning Roles

24. Have you settled on an intermediary for carrying out your Social Impact Bond?
25. Have you settled on an investor for financing your Social Impact Bond?
26. Are you reaching out to populations that are otherwise less likely to be served (or more costly to serve)? How is this reflected in repayment terms for the SIB?
27. What types of roles are the intermediary and/or investor playing in executing the Social Impact Bond pilot activities?
28. What strategies have you used in the Social Impact Bond partnership to make key decisions, for example, in balancing stakeholder interests and authority over different aspects of the project?
29. How much influence does each Social Impact Bond project partner or stakeholder have in determining the following (note: please also indicate *which* partner(s) or stakeholder(s) have a role in these tasks):
- a. Which outcomes to measure
 - b. The target population/eligibility criteria and number to serve
 - c. Measures and methodologies for evaluation
 - d. Terms of re-payment of investor(s) and timeline for re-payment
 - e. Budget items
 - f. Project deliverables?
30. Have any conflicts of interest or related problems emerged among the implementing partners?

Part F: Evidence and Evaluation

31. What types of evidence show that the proposed intervention could lead to government savings large enough to repay investors?
32. Did you conduct a cost-benefit analysis (CBA) to evaluate the viability of a SIB project?
 - a. How did you determine the key cost/benefit components for the CBA?
 - b. Is there a publicly available copy of the CBA that you could share with us?
33. Does the expectation for employing an evidence-based model limit (or support) the testing of new innovations?
34. Will you measure SIB program impacts with an experimental or quasi-experimental research design? How are the project deliverables linked to impact measures?
35. How will the intermediary/arbitrator determine how much should be repaid to the investor based on the results?
36. Have you established the terms of repayment in the contract, or have these provisions been discussed in the feasibility pilot?
37. At this point in the feasibility project, how much confidence do you have that your Social Impact Bond project will succeed? Please choose among the following:
10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

Section III: Questions for denied proposals only

Part G: Working Outside of the USDOE Grant

38. Having not been selected for the U.S. Dept of Education Social Impact Bond feasibility pilot, are you still pursuing a Social Impact Bond arrangement?
39. Are you pursuing your project/program goals through other means of collaboration or funding? (If so, what are you doing instead?)
40. Are you working to implement the evidence-based model included in your application? (Why or why not)?
41. How likely do you think it is that you will undertake a Social Impact Bond project in the future? Please choose among the following:
10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

ⁱ <https://www.gov.uk/guidance/social-impact-bonds>

ⁱⁱ Simple descriptive statistics were utilized in some of these comparisons (e.g., t-tests of mean differences in ratings or scores) to investigate associations, but given the small sample size (with no claims made about representativeness), we do not interpret these as systematic or causal relationships.

ⁱⁱⁱ This difference was confirmed as statistically significant (at $\alpha=0.05$) with a two-sample t-test.

^{iv} Text from the roundtable description: <https://appam.confex.com/appam/2015/webprogram/Session6214.html>.

^v Text on the goals of PFS in the Bipartisan Budget Act of 2018 (PL 115-123), enacted February 9, 2018:

(1) To improve the lives of families and individuals in need in the United States by funding social programs that achieve real results. (2) To redirect funds away from programs that, based on objective data, are ineffective, and into programs that achieve demonstrable, measurable results. (3) To ensure Federal funds are used effectively on social services to produce positive outcomes for both service recipients and taxpayers. (4) To establish the use of social impact partnerships to address some of our Nation's most pressing problems. (5) To facilitate the creation of public-private partnerships that bundle philanthropic or other private resources with existing public spending to scale up effective social interventions already being implemented by private organizations, nonprofits, charitable organizations, and State and local governments across the country. (6) To bring pay-for-performance to the social sector, allowing the United States to improve the impact and effectiveness of vital social services programs while redirecting inefficient or duplicative spending. (7) To incorporate outcomes measurement and randomized controlled trials or other rigorous methodologies for assessing program impact.