PAR PUBLIC ADMINISTRATION REVIEW

Carolyn J. Heinrich
Vanderbilt University

Sarah E. Kabourek
NORC at the University of Chicago

Pay-for-Success Development in the United States: Feasible or Failing to Launch?

Research Article

Abstract: *Social impact bonds, known as pay-for-success (PFS) initiatives in the United States, have attracted attention as a novel strategy for financing and providing preventive services to the most vulnerable populations. This article provides an exploratory qualitative analysis of the Preschool PFS Feasibility Pilot grant applications and projects initiated by the U.S. Department of Education to encourage state and local exploration of PFS for implementing high-quality preschool programs. Drawing on the public administration evidence base that informs PFS design, the authors examine the feasibility pilots' features and investigate why grant applicants saw PFS as a promising strategy for achieving their preschool program goals. The challenges encountered, lessons learned, and perceived viability of fully executed PFS preschool programs are also discussed. Findings show 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.*

Evidence for Practice

- Government actors play a key role in making pay-for-success (PFS) contracts viable and carry the primary responsibility to ensure that these public-private partnerships are politically accountable to the public.
- Most of the Preschool PFS Feasibility Pilot projects funded by the U.S. Department of Education did not advance to PFS contract arrangements because of capacity constraints, difficulties in financial modeling and getting investors and payers on board, and other complexities of negotiating multiyear contracts.
- Even if PFS pilot efforts do not progress to full-scale projects, important capacities are being built (and should continue to be supported with federal funding) for more effectively serving vulnerable populations through public-private partnerships and for tracking longer-term outcomes.
- A greater commitment is needed on the part of PFS partners to ensure that impact evaluations undertaken to assess project returns and determine payments for outcomes are methodologically rigorous and based on long-term outcomes that cannot be easily "gamed."

Social impact bonds (SIBs) have attracted outsize attention as a novel, "nothing to lose" strategy for addressing some of our most "intractable social challenges" (Brady 2016; Gustafsson-Wright, Gardiner, and Putcha 2015). Characterizations of SIBs as a social investment tool are not uniform in the literature, but SIBs are typically distinguished by (1) their focus on preventive-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 includes 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 in 2010, and some have pointed to the global recession as a catalyst for the emergence of SIBs (Williams 2018). Indeed, Roy, McHugh, and Sinclair (2017) described 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, or pay-for-success (PFS) initiatives as they are more widely known in the United States, are still in the early "feasibility" stages of development. The Social Innovation Fund, a program of the federal Corporation for National and Community Service (authorized by the Edward M. Kennedy Serve America Act in 2009), invests in evidence-based programs and interventions at state and local levels to find "new ways to solve old problems that are faster, cost-effective, data-driven and lead to better results," some of which have progressed to PFS projects (see https://www.nationalservice.gov/programs/social-innovation-fund).

Carolyn J. Heinrich is Patricia and Rodes Hart Professor of Public Policy, Education and Economics in the Peabody College at Vanderbilt University. Her research focuses on education, workforce development, social welfare policy, program evaluation, and public management. She also works directly with federal, state, and local governments and nongovernmental organizations in her research to improve policy design and program impacts.
E-mail: carolyn.j.heinrich@vanderbilt.edu

Sarah E. Kabourek is a research scientist at NORC at the University of Chicago. She uses quantitative and qualitative methods to study early childhood education, public preschool, and issues related to equity in public education finance.
E-mail: sarah.e.kabourek@vanderbilt.edu

Public Administration Review, Vol. 79, Iss. 6, pp. 867–879. © 2019 by The American Society for Public Administration.
DOI: 10.1111/puar.13099.

In January 2012, Massachusetts was the first U.S. government entity to pass legislation authorizing funds for PFS initiatives, 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) explained, the structure and components of legislation vary considerably, such that there is no "typical PFS" model guiding the contractual arrangements and funding mechanisms as they evolve. This contrasts with the more centralized approach to SIB commissioning in the United Kingdom, where a cabinet-level office provides a range of tools and resources for SIB project development, including an SIB contract template, a unit-cost database, and cost–benefit analysis guidance, and fund managers and advisory firms provide technical assistance to government entities in assessing feasibility and implementing SIB projects and contracts.

In 2016, the U.S. Department of Education (US ED) allocated funds for PFS feasibility studies, and in February 2018, four years after it was first proposed, the U.S. Congress passed federal legislation, the Social Impact Partnerships to Pay for Results Act (SIPPRA), that provided $100 million in federal funding to support PFS partnerships, feasibility studies, projects, and evaluations in a more integrated approach to guiding PFS efforts in the United States.

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

Therefore, we undertake an exploratory, qualitative analysis of the Preschool PFS Feasibility Pilot grant applications (funded and unfunded) and studies that were launched by US ED in the 2016 fiscal year. The purpose of the Preschool PFS Feasibility Pilot grants, which were awarded to eight entities in amounts ranging from $335,000 to $397,000 (totaling approximately $3 million), was to encourage state and local exploration of whether PFS is a viable and effective strategy for implementing or expanding high-quality preschool programs. We received 20 applications that were submitted to US ED, along with the reviewer ratings and comments. We conducted interviews with 12 applicants (5 of 8 awardees and 7 of 12 non-awardees) approximately 18 months after the funding award decisions were made.

The objectives of this study are threefold: (1) to identify the proposed feasibility pilot features that US ED and applicants saw as advantageous and critical to the success of PFS; (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 challenges encountered, lessons learned, and perceived viability of fully executed PFS preschool programs in the United States.

The following section begins by describing the basic SIB/PFS model and 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 US ED—as well as the qualitative methods employed in coding and analyzing the data and their limitations.

We then turn to our analysis, which examines the proposed Preschool PFS Feasibility Pilot features and US ED ratings of them to illuminate what elements were viewed as critical to PFS feasibility, pulling 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 of emerging preschool PFS models, and we conclude with a discussion of key lessons pertinent to governments and their partners heading down the PFS path.

How Prior Research and Experience Inform the Preschool PFS Efforts

PFS Basics and Origins

SIBs/PFS are financing mechanisms designed to reduce the risk to government of paying for programs that do not produce public benefits. A private investor provides money up front to finance the rollout or expansion of the public program, and the government

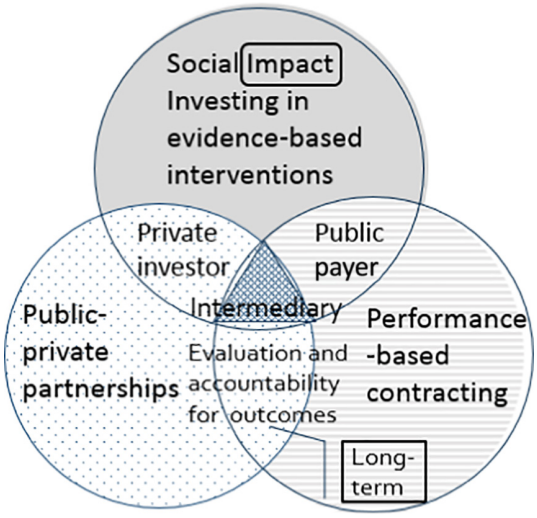

Figure 1 Conceptual Depiction of the Pay-for-Success Model and Components

agency (or public payer) is required to repay the investor *only if* specified (contracted) outcomes are achieved. Governments and their public-private partners work together to implement research-based preventive interventions and to define specific, measurable *longer-term* outcomes of the programs on which the terms of repayment to the investor are based. Ideally, a rigorous evaluation is conducted to establish a causal link between the intervention and the measured outcomes and to monetize the value of the program impacts. An intermediary serves as the arbiter of the contractual arrangement and manages the transfer of funds from the public payer to the private investor if it is determined that the outcomes are met (see figure 1).

While the roles of the private investor in providing gateway funding to launch programs and the intermediary as a financial broker between the public and private partners are key innovations of the SIB/PFS design, other core components shown in figure 1—including public-private collaborations in services delivery, performance-based contracting, and evaluation and accountability outcomes—have been features of governance and vigorously studied in public and business administration literatures for years. In fact, more than 40 years ago, Guttman and Willner (1976, xii) called attention to dramatic new ways in which 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 (Moynihan and Pandey 2005). Hood (1991, 4–5) described NPM as "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. Tracing the origins of NPM to Jeremy Bentham's writings of the late eighteenth and early nineteenth centuries, Hood and Jackson (1991) argued 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) 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 found an "innovative new mix" for public service delivery. She argued 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 known in the public administration literature. In a synthesis of theoretical and empirical literature on pay-for-performance systems 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. Williams (2018), in fact, suggests that many SIB intermediaries are already pulling back from measuring impacts and recommending outcomes-based evaluations.

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, Heinrich, and Smith 2002; Schochet, Burghardt, and McConnell 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 U.S. legislation), then SIB/PFS designers should be wary of compromising with shorter-term outcome 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 1994), 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 (Hefetz and Warner 2004; Kettl 1993; Romzek and Johnston 2002). Furthermore, in designing and implementing 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 a synthesis of the literature on SIBs, Fraser et al. (2018) identified competing public and private values as a major crosscutting theme that poses challenges in implementing SIBs. Concerns about the incompatibility of the "logics and normative assumptions of the private financial services sector" (Fraser et al. 2018, 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).

In welfare services delivery in Wisconsin, for example, private 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, concerns were 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 service delivery.

As 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 (PFS or other) that initially reflect shrewd attention to these details, over time their effectiveness may decline as problems emerge because of 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 goals intensify (Bevan and Hood 2006; Courty and Marschke 2004; Heinrich 2003; Tse and Warner 2018). With 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) suggested 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 may preclude stability or even favor an instrument for funding social programs that is continually evolving and adaptable across contexts.

Preschool Programs: A Ripe Prospect for PFS?

Preschool has been valued by the public as an important social intervention for decades, and rigorous cost–benefit analyses have identified social investment returns ranging from $7 to $10 for every dollar invested in high-quality programs (i.e., Heckman et al. 2010; Karoly 2016). The solid evidence base provides parties to a PFS arrangement with a more predictable, monetizable return on investment and the expectation that benefits will continue to accrue over the long term (Karoly 2016). Despite this strong support for public preschool as an effective strategy for promoting positive outcomes of children and reducing gaps in school readiness, program capacity is one of the greatest barriers to preschool enrollment (Bailey et al. 2017; Bassok, Gibbs, and Latham, 2018; Chaudry et al. 2017). After growing steadily in the early 2000s, state budgets for public preschool declined during the Great Recession and have not recovered (Haskins 2017). Consequently, preschool providers often rely on the blending of federal, state, and local funding streams and public-private partnerships to create a single classroom, which in some ways makes preschool a ripe prospect for PFS (Chaudry and Datta 2017; Hustedt and Barnett 2011). That said, a qualitative study of efforts to expand early childhood services through SIBs in three cities argued that governments "walk a razor's edge" in attempting to establish a cost structure that maximizes public investment over private investor profits (Tse and Warner 2018).

In our exploration of US ED's Preschool PFS initiative, we observed considerable variability in the PFS partnerships and plans for executing a PFS, 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, whereas others have hit roadblocks or contingencies that they are still working to overcome. Many of the complex challenges the PFS partners face were foreshadowed by past experiences with public sector performance-based contracting and pay-for-performance systems. Indeed, this seems to be a consistent theme in the burgeoning SIB/PFS literature: many cautionary tales or narratives emerging in the implementation of these new social investment tools could have been anticipated based on our substantial experience with many of the core elements of these instruments (Fraser et al. 2018).

Next, after briefly describing our data and methods, we consider the distinctive features of Preschool PFS Feasibility Pilot projects underway in the United States and how they relate to what theory and experiences with SIB/PFS models to date suggest would be desirable in these arrangements.

Study Data and Methods of Analysis

US ED released a request for proposals (closing date October 6, 2016) with the stated goal to fund Preschool PFS Feasibility Pilot projects to "determine if this model is an effective strategy to implement preschool programs that are high-quality and yield meaningful results" (US ED 2016). US ED made the 20 applications received publicly available (with some proprietary and budgetary information redacted). Appendix A in the Supporting Information online lists the applicants (award recipients and those not funded), their locations, and whether they participated in the interviews we conducted for this research; additional information on the funded applications is available at https://www2.ed.gov/programs/pfs/awards.html.

Each submitted proposal was independently rated by three reviewers, and we compiled all reviewer ratings and comments in a database. The following major proposal features (shown in greater detail in appendix B) were scored: (1) the magnitude of need in the target population, (2) the quality of the preschool program design, (3) the quality of the existing or proposed PFS partnership, (4) the quality of the work plan, (5) the quality of the project leadership and team, (6) the budget (adequacy of resources), and (7) the 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 have indicated are critical components or building blocks of a successful PFS. These include 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 Institutional Review Board–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 partner, (4) the preschool program model, (5) the determination of project roles (if funded), (6) evidence used and project evaluation activities, and (7) prospects for success and plans for achieving the program goals.

We reached out by phone and email to all 20 applicants, contacting the people identified as the project leaders in the applications, to schedule interviews. We secured the participation of 12 applicants: 5 of 8 who were awarded feasibility pilot grants and 7 of 12 who did not receive awards (see appendix A). Interviewees included executive program staff at the school (1), city (3), district (1), county (5), and state level (2). The timing of the interviews, June to August 2018, was approximately 18 months after the funding award decisions were made. 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. Confidentially 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 exploratory 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,[1] 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. In analyzing the qualitative data, we used an inductive (grounded theory) approach to coding, following initial line-by-line coding procedures suggested by Charmaz (2014). A more detailed discussion of these steps, along with the final categories and sample focus codes, are shown in appendix E.

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 features aligned with anticipated, distinctive components of the PFS financing model; 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 of the feasibility pilots or the alternative models they were pursuing. We integrated the responses of both grant recipients and unfunded project leaders in our analysis and discussion, because they were pursuing similar goals and experiencing analogous challenges in their work. However, given the 60 percent overall participation rate for our interviews and the fact that it is not possible to assess how representative the applicants were of organizations nationwide that may have been considering SIB-financed programs, we do not make any claims about the external validity of our findings, nor are we undertaking any causal analysis in this exploratory study.

What Proposed PFS Pilot Features Are Perceived to Presage Feasibility?

In our analysis of US ED's review of Preschool PFS Feasibility Pilot applications to better understand what pilot features were seen (from the reviewers' and applicants' perspectives) as presaging PFS success, we focus primarily on the three most highly weighted components of the application: the quality of the preschool program design, the PFS partnership, and the work plan (25 of 105 total points each, or more than 70 percent of the total possible points). 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) pointed out, while it is an overarching goal of SIBs/PFS to promote innovation and allow local flexibility, these projects 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 preschool program design—consistent with federal (Office of Management and Budget, OMB) intent to fund programs that are already known to be effective and more likely to produce positive impacts (GAO 2015)—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 nonfunded pilot applications that specified the preschool curriculum would be research based.

Furthermore, reviewers' 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, socioemotional learning, and other student outcomes, whereas only half of the nonfunded 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 OMB (2013) prescriptions for evidence use that have now been incorporated into SIPPRA requirements for 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. Moreover, this and other applicant applicants were able demonstrate (through prior work) their capability with partners to match children's prekindergarten 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 touted their 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 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 Bipartisan Budget Act of 2018 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 United Kingdom and other national governments have more actively assumed (Arena et al. 2016; Greenblatt and Donovan 2012).

Public-Private Partners and Payers for PFS Implementation

The U.S. Government Accountability Office's (2015) review of early PFS projects suggested that they worked better when they built on or expanded existing, effective interventions, and this was reflected in the feasibility pilot application reviews and the subsequent SIPPRA legislation, which articulated 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). The pilot application 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. The reviewers, in turn, identified as weaknesses the lack of an existing partnership and rated more poorly PFS partnership quality when key partners or elements of the PFS arrangements (e.g., investor, intermediary, evaluator, or payer) 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 prekindergarten, including a successful tax referendum, a strategic plan for implementation, and new funding for quality improvement of area preschool programs. In explaining the low ratings of the application on this dimension, one reviewer suggested that the PFS partnership lacked detail on the roles of the partners, and another pointed out that the mayor's letter did not commit the city to being the payer (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 payer was identified.

In reality, among the eight feasibility pilot applications selected for funding, only three proposals identified a "likely" or prospective payer. One of these three funded applications identified a school district as a possible outcomes payer and indicated that the district's commitment to being the payer would be contingent on a "rigorous" demonstration of the program's effectiveness at improving child outcomes. The highest-rated application did not identify a payer in advance, and in our interview, the pilot PFS director stated that she thought the PFS had a 5 percent to 10 percent chance of succeeding; the biggest challenge was determining who would be the payer.

A team member of a funded feasibility pilot operating at the state level indicated that they were also struggling with who should be the payer—would the savings be statewide or local? Given the highly regulated nature of funding mechanisms at the state level, it was not clear that there would be a mechanism for the state to make payments to an investor. Alternatively, another project leader suggested that because of the state's proportionately high level of contributions to special education funding, the state was the only logical (prospective) payer. Still, the director of one of the funded feasibility pilots that had identified a payer 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 payer (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. When 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 was 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 the project 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 who had considered the same intermediary used by the three pilots with a contracted intermediary indicated that 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?

The other key partner in the essential infrastructure of a SIB/PFS is the independent evaluator. A hallmark of SIB/PFS evaluations is the

intention to employ the most rigorous methods possible to evaluate the "impact" of the intervention, preferably using randomized controlled trials or other rigorous methodologies that allow for causal attribution of the measured outcomes to the intervention (GAO 2015). Five of the eight funded Preschool PFS applications 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 randomized controlled trial. 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 payer 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 payer, 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 United Kingdom, 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 (Tse and Warner 2019). One strategy to realize more secure returns is to specify simpler, easier to achieve outcome measures along with methods that, in Berlin's (2016, 2) words, "may misleadingly provide only the illusion of benefits and savings to government entities."

Putting the PFS Components Together: Contract Arrangements and Capacities to Make PFS Work

As shown in figure 1, the intermediary plays a key role as a conduit between partners in implementing a PFS feasibility pilot. Thus, identifying an experienced intermediary and describing the intermediary's responsibilities were critical factors in US ED reviewer ratings of the Preschool PFS work plans. The sum of the three reviewer scores on work plan ratings were more than 22 points higher (of 75 total) on average for applications that named an intermediary (confirmed or prospective). Similarly, work plan ratings were 15 points higher on average if an independent evaluator was identified in the application.[2] 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 time line 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 payer; 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 foregoing tasks, perhaps most distinctive to the undertaking of an SIB/PFS are those that involve modeling the financial benefits and cost savings and defining the terms or pricing for the end payer 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 US ED's (2017) "tool kit" for guiding feasibility studies, these tasks are critical to determining whether 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 GAO (2015) asserted 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 in which the government pays a service provider for outcomes achieved, which raises 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 and Albertson 2011; Tse and Warner 2019)?

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 resolve a long-standing public sector performance management problem known as the "time-shortening disease" (Bouckaert 1993; Heinrich 2003). This is a phenomenon in which policy makers 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 Feasibility Pilot applicants, we heard from program implementers about how the PFS opportunity spurred them to consider longer-term outcomes, such as third-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 repayment time line 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 third 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 that 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 that 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 payers, where the anticipated payers were school districts or local or state government entities. One interviewee expounded, "If you don't have an end payer, 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 repayment mechanisms using public funds.

In light of these findings, we now consider: if not a formal 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?

What Is the Potential for PFS in the United States? Are Viable Models Emerging?

In a Government Outcomes Lab report, Carter et al. (2018, 2) explained that in approximately 40 SIBs launched in the United Kingdom, "practitioners are pushing the boundaries of the SIB concept to try out new variants and new factors which might work better to improve outcomes." The report described, for example, how SIBs vary in the extent to which they are preventive and take advantage of the flexibility in their ability 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 United Kingdom, where the rollout of SIBs has been more centralized and better supported by advisory firms (e.g., Social Finance UK) and fund managers, it has been found that "local context matters greatly" for implementing this new public-private collaboration on outcomes-based contracting (Carter et al. 2018, 2).

At the same time, Carter et al. pointed out that those studying SIBs fail to grasp these opportunities to go beyond questions (and rhetorical debate) about whether SIBs work to ask where and why they work well (or do not work). While the United States has about half the number of SIB/PFS projects underway compared with the United Kingdom, we draw out some general and context-specific lessons emerging from US ED's 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 that SIBs/PFS were spurred in part by financial constraints intensified by the deep recession (Roy, McHugh, and Sinclair 2017), 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 toward 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 and Albertson 2011; Jackson 2013; Warner 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 percent to 100 percent), 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 percent to 10 percent (funded applicant) to a high of 80 percent to 100 percent (unfunded applicant). Only one funded applicant felt more confident after undertaking the pilot (an increase from 50 percent to 60 percent) 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 percent to 10 percent probability, and the other felt confident that they would come up with some kind of feasible model, but they were not 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 percent) than those selected (42 percent). 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 percent 10 percent and 25 percent to 33 percent) and were not more confident of success 18 months later.

These perceptions of the prospects of PFS success were consistent with Williams's (2018, 14) findings from interviews with 178 members of core groups of SIB actors in the United States, Canada, and the United Kingdom: "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 Fund, 37 had decided by 2016 that they would not continue to pursue a SIB.

Insights on the Viability of PFS for Preschool Programs in the United States

A comment frequently heard in our 12 interviews was 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 payer. 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 the 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 whether 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 looked "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 time line 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 prekindergarten. They felt that even though the pilot wasn't going forward, the time allocated for strategic planning (around a five-year plan for early childhood) and for exploring how they could scale the prekindergarten 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 prekindergarten.

Among the 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 prekindergarten 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 prekindergarten 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 data sets 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 because they "didn't have a mechanism to gauge what was considered quality." 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 recipient 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

Through SIPPRA, the Every Student Succeeds Act, the Workforce Innovation and Opportunity Act, and other federal, state, and local initiatives, public funds have been authorized to continue exploring the potential for PFS—an innovative social investing

tool—to produce measurable, long-term social benefits while also generating cost savings for the public sector. Our exploration and analysis of US ED's first Preschool PFS Feasibility Pilot efforts to support public preschool program expansion identified critical elements of PFS projects that US ED saw as foundational to PFS success, including high-quality, evidence-based preschool program designs; existing public-private partnerships with histories of strong collaboration that also involved or identified a public payer, intermediary, and evaluator; and a clear plan that delineated responsibilities between the program partners and specified a detailed time line for execution of key project tasks (e.g., the modeling of financial benefits and cost savings and definition of terms or pricing for the end payer in relation to program outcomes achieved). In effect, US ED had high expectations for applicants' existing infrastructure, plans, and capacities for conducting a PFS pilot, which Dominey (2012) suggested is likely to favor larger and more experienced organizations or those with a track record that will be more likely to minimize the private investors' risk.

Our findings from interviews with program leaders attempting to make Preschool PFS Feasibility 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 payers 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 (multiyear) period. They were also largely falling short on the expectation to undertake *impact* evaluations for assessing project returns (and determining payments for outcomes), primarily because of data access and methodological design limitations. Although many of the preschool program leaders aspired to conduct a cost–benefit analysis (indicated in 9 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 one rare mention of a comparison group). This raises the specter that in the absence of greater resources and capacity-building in evaluation, simpler, gameable or misleading measures and methods for assessing program impacts will undercut a core goal of SIPPRA to improve the longer-term impact and effectiveness of vital social services programs (Berlin 2016; Fraser et al. 2018).

Our qualitative analysis also suggested that the capacity building of Preschool PFS Feasibility Pilots was aided greatly by collaboration with experienced intermediaries and evaluators, but only a small number of the projects (even US ED grant funds) had adequate resources to afford their services. While an allocation of $3 million (or between $300,000 to $400,000 per project) by US ED for supporting capacity-building activities was commendable, this minimal level of funding falls short of the need and demand that currently exists for this type of support. A majority of the projects (funded and unfunded) were actively engaged in fund-raising activities and efforts to expand coalitions of support—for example, strengthening collaborations between the public sector, nonprofit organizations (local or community-based and statewide), and private, for-profit partners (potential investors, business leaders, etc.)—some with the goal to contract with intermediaries and evaluators that could accelerate their progress toward a formal PFS arrangement.

More generally, our study findings raise the question of whether, going forward, the federal government should take a more proactive and generous role (in terms of technical and funding support) to assist public-private partnerships at the state and local levels in developing essential capacities for executing PFS. In any PFS feasibility effort, public-private partners will be required to put resources at risk—for example, significant investments of time and personnel in structuring and managing collaborations, building technical capacities, planning and analysis, and fostering political will and support—all of which have opportunity costs.

To make it more likely that these efforts will pay off, government partners at federal, state, and local levels could make more concerted efforts to review regulations that present barriers to contractual arrangements, reduce restrictions on access to data, and ease political conflicts that have frequently slowed or deterred progress toward establishing formal PFS partnerships and securing a public payer. Indeed, our investigation suggested that not only is the government a key partner in making a PFS contract viable, but also it 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). In addition, we believe the federal government could—consistent with the current goals and provisions of SIPPRA—extend these supports more broadly to public and private organizations espousing the goals of performance-based contractual arrangements and longer-term impact evaluation, regardless of whether a PFS was the ultimate aim.

At the same time, it is important to reiterate that this is an exploratory study of one federal agency's recent efforts to assess the feasibility of PFS in one social program area (preschool programs) and that limitations of the study design and methods greatly restrict the generalizability of the findings to other PFS initiatives in progress or getting underway. We suggest that more independent research—outside of the evaluation work undertaken by contracted (PFS) evaluation parties—should be planned and supported with SIPPRA funds across varying political, social and economic contexts that will undoubtedly shape both how PFS projects are structured and implemented and their prospects for success (Carter et al. 2018; Fraser et al. 2018; Tse and Warner 2018). We also suggest that attaining the goal of a formal PFS contracting arrangement should not be the only measure of success of PFS feasibility efforts, and in fact, it might not be the most desired outcome of these efforts. Rather, supporting more public-private collaborations in building basic capacities for more effectively serving vulnerable populations and in holding public-private partnerships accountable for longer-term outcomes and impacts could be a major public management advance of PFS initiatives that has long been a goal of public management reforms such as NPM and its predecessors.

Lastly, the research agenda for public administration scholars as PFS moves forward under SIPPRA is brimming with weighty questions that will be important to address, particularly as PFS projects advance and more data become available to study them.

These include questions about viable contracting designs and how they balance financial, political, and social risks and benefits, as well as the types of performance outcome metrics adopted to assess PFS success and determine payouts for results. In addition, it will be important to investigate the implications of PFS for the nature and types organizational capacity created (or drained) among the public-private partners in implementing PFS projects.

Relatedly, how much will we realize in public sector savings relative to public sector investments in making PFS work? Researchers should also examine how much new investment in social programs PFS brings about over the longer term and how it is allocated, that is, attending to the social equity implications of who benefits from public services newly offered or expanded through PFS. These are only a handful of research questions in what could be a burgeoning public administration research agenda, depending in part on the public will and capacity to continue down the PFS pathway.

Acknowledgments

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 Pilot program applicants who participated in interviews and shared their time, information, and perspectives on their feasibility pilot applications and preschool program expansion efforts. We also thank the anonymous referees of this manuscript and Mildred Warner for helpful feedback on our study.

Notes

1. Simple descriptive statistics were used 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.
2. These differences were confirmed as statistically significant (at $\alpha = 0.05$) with a two-sample *t*-test.

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(12): 927–39.

Bailey, Drew, Greg J. Duncan, Candice L. Odgers, and Winnie Yu. 2017. Persistence and Fadeout in the Impacts of Child and Adolescent Interventions. *Journal of Research on Educational Effectiveness* 10(1): 7–39.

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.

Bassok, Daphna, Chloe R. Gibbs, and Scott Latham. 2018. Preschool and Children's Outcomes in Elementary School: Have Patterns Changed Nationwide between 1998 and 2010? *Child Development*. Published electronically on April 17. https://doi.org/10.1111/cdev.13067.

Berlin, Gordon L. 2016. Learning from Experience: A Guide to Social Impact Bond Investing. MRDC, March. https://www.mdrc.org/sites/default/files/Learning_from_Experience_SIB.pdf [accessed August 13, 2019].

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(3): 517–38.

Bouckaert, Geert. 1993. Measurement and Meaningful Management. *Public Productivity & Management Review* 17(1): 31–43.

Brady, Dale. 2016. Taxpayers' Nothing to Lose Strategy to Fund Novel Approaches to Huge Problems. *Observer*, November 2. http://observer.com/2016/11/social-impact-bonds/ [accessed August 14, 2019].

Carter, Eleanor, Clare FitzGerald, Ruth Dixon, Christina Economy, Tanyah Hameed, and Mara Airoldi. 2018. Building the Tools for Public Services to Secure Better Outcomes: Collaboration, Prevention, Innovation. Evidence Report, Government Outcomes Lab, University of Oxford, Blavatnik School of Government, July. https://golab.bsg.ox.ac.uk/documents/164/BSG-GOLab-EvidenceReport-20190730.pdf [accessed August 14, 2019].

Charmaz, Kathy. 2014. *Constructing Grounded Theory*. 2nd ed. Thousand Oaks, CA: Sage Publications.

Chaudry, Ajay, and A. Rupa Datta. 2017. The Current Landscape for Public Pre-Kindergarten Programs. In *The Current State of Scientific Knowledge on Pre-Kindergarten Effects*, 5–18. Washington, DC: Brookings Institution, Pre-Kindergarten Task Force. https://www.brookings.edu/wp-content/uploads/2017/04/duke_prekstudy_final_4-4-17_hires.pdf [accessed August 14, 2019].

Chaudry, Ajay, Taryn Morrissey, Christina Weiland, and Hirokazu Yoshikawa. 2017. *Cradle to Kindergarten: A New Plan to Combat Inequality*. New York: Russell Sage Foundation.

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. https://www.chapman.com/media/publication/755_Chapman_Inroads_to_Innovation_Pay_for_Success_Legislation_0417.pdf [accessed August 15, 2019]

Davison, Robbie, and Helen Heap. 2013. *Can Social Finance Meet Social Need?* Liverpool: Tomorrow's People. June. http://www.huckfield.com/wp-content/uploads/2013/06/13-Heap-Davison-Soc-Fin-Soc-Need-U26.pdf [accessed August 16, 2019].

Dominey, Jane. 2012. A Mixed Market for Probation Services: Can Lessons from the Recent Past Help Shape the Near Future? *Probation Journal* 59(4): 339–54.

Fraser, Alec, Stefanie Tan, Mylene Lagarde, and Nicholas Mays. 2018. Narratives of Promise, Narratives of Caution: A Review of the Literature on Social Impact Bonds. *Social Policy & Administration* 52(1): 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. 2013. The Promise of Pay for Success. *Federal Reserve Bank of San Francisco Community Development Investment Review*, April, 19–22.

Gustafsson-Wright, Emily, Sophie Gardiner, and Vidya Putcha. 2015. The Potential and Limitations of Impact Bonds: Lessons from the First Five Years of Experience Worldwide. Global Economy and Development at Brookings, July. https://www.brookings.edu/wp-content/uploads/2016/07/Impact-Bondsweb.pdf [accessed August 13, 2019].

Guttman, Daniel, and Barry Willner. 1976. *The Shadow Government*. New York: Pantheon Books.

Haskins, Ron. 2017. Financing Early Childhood Programs. In *The Current State of Scientific Knowledge on Pre-Kindergarten Effects*, 85–92. Washington, DC: Brookings Institution, Pre-Kindergarten Task Force. https://www.brookings.edu/wp-content/uploads/2017/04/duke_prekstudy_final_4-4-17_hires.pdf [accessed August 14, 2019].

Heckman, James J., Carolyn J. Heinrich, and Jeffrey Smith. 2002. The Performance of Performance Standards. *Journal of Human Resources* 37(4): 778–811.

Heckman, James, J., Seong H. Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. The Rate of Return to the HighScope Perry Preschool Program. *Journal of Public Economics* 94(1–2): 114–28.

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): 171–90.

Heinrich, Carolyn J. 2003. Measuring Public Sector Performance and Effectiveness. In *The Sage Handbook of Public Administration*, edited by B. Guy Peters and Jon Pierre, 25–37. London: Sage Publications.

Heinrich, Carolyn J., and Youseok Choi. 2007. Performance-Based Contracting in Social Welfare Programs. *American Review of Public Administration* 37(4): 409–35.

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

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.

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

Hood, Christopher, and Michael W. Jackson. 1991. The New Public Management: A Recipe for Disaster? *Canberra Bulletin of Public Administration* 64: 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): 267–82.

Hustedt, Jason T., and W. Steven Barnett. 2011. Financing Early Childhood Education Programs: State, Federal, and Local Issues. *Educational Policy* 25(1): 167–92.

Jackson, Edward T. 2013. Evaluating Social Impact Bonds: Questions, Challenges, Innovations, and Possibilities in Measuring Outcomes in Impact Investing. *Community Development* 44(5): 608–16.

Karoly, Lynn A. 2016. The Economic Returns to Early Childhood Education. *The Future of Children* 26(2): 37–55.

Kettl, Donald F. 1993. *Sharing Power: Public Governance and Private Markets*. Washington, DC: Brookings Institution.

Lantz, Paula M., George Miller, Corwin N. Rhyan, Sara Rosenbaum, Leighton Ku, and Samantha Iovan. 2018. "Pay for Success" Financing and Home-Based Multicomponent Childhood Asthma Interventions: Modeling Results from the Detroit Medicaid Population. *Milbank Quarterly* 96(2): 272–99.

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

Moynihan, Donald P., and Sanjay K. Pandey. 2005. Testing How Management Matters in an Era of Government by Performance. *Journal of Public Administration Research and Theory* 15(1): 421–39.

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

Office of Management and Budget. 2013. Memorandum to the Heads of Departments and Agencies: Next Steps in the Evidence and Innovation Agenda. M-13-17, July 26. https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2013/m-13-17.pdf [accessed August 14, 2019].

Rhodes, Richard A. 1994. The Hollowing Out of the State: The Changing Nature of the Public Service in Britain. *The Political Quarterly* 65(2): 138–51.

Romzek, Barbara S., and Jocelyn M. Johnston. 2002. Effective Contract Implementation and Management: A Preliminary Model. *Journal of Public Administration Research and Theory* 12(3): 423–53.

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. https://www.tnlcommunityfund.org.uk/media/research-documents/social-investment/loud_sib_model.pdf?mtime=20190215124541 [accessed August 14, 2019].

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, 263–78. Cheltenham: Edward Elgar.

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

Schochet, Peter Z, John Burghardt, 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. Report prepared by Mathematica Policy Research for the Employment and Training Administration, U.S. Department of Labor, August. https://wdr.doleta.gov/research/FullText_Documents/National%20Job%20Corps%20Study%20and%20Longer%20Term%20Follow-Up%20Study%20-%20Final%20Report.pdf [accessed August 14, 2019].

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.* Published electronically on May 14. https://doi.org/10.1080/07352166.2018.1465347.

———. 2019. A Policy Outcomes Comparison: Does SIB Market Discipline Narrow Social Rights? *Journal of Comparative Policy Analysis*. Published electronically on May 24. https://doi.org/10.1080/13876988.2019.1609789.

U.S. Department of Education (US ED). 2016. U.S. Department of Education Announces New Pay for Success Grant Competition for Preschool. News release, August 19. https://www.ed.gov/news/press-releases/us-department-education-announces-new-pay-success-grant-competition-preschool [accessed August 16, 2019].

———. 2017. Pay for Success Feasibility Tool Kit. https://www2.ed.gov/about/inits/ed/pay-for-success/pay-for-success-toolkit.pdf [accessed August 16, 2019].

U.S. Government Accountability Office (GAO). 2015. *Pay for Success: Collaboration among Federal Agencies Would Be Helpful as Governments Explore New Financing Mechanisms*. Washington, DC: U.S. Government Printing Office. GAO-15-646.

Warner, Mildred E. 2013. Private Finance for Public Goods: Social Impact Bonds. *Journal of Economic Policy Reform* 16(4): 303–19.

Williams, James W. 2018. Surveying the SIB Economy: Social Impact Bonds, "Local" Challenges, and Shifting Markets in Urban Social Problems. *Journal of Urban Affairs.* Published electronically on September 28. https://doi.org/10.1080/07352166.2018.1511796.

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

Supporting Information

A supplemental appendix can be found in the online version of this article at http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1540-6210.