

Using Social Impact Bonds to Spur Innovation, Knowledge Building, and Accountability

David Butler, Dan Bloom, and Timothy Rudd

MDRC

In this article, we propose a vision of a social impact bond (SIB) model that moves beyond just achieving cost-savings to spurring innovation, knowledge-building, rigorous evaluation, and, potentially, outcomes that go beyond cost savings. We discuss two of the key rationales for SIBs: securing new resources to expand programs more broadly and ensuring that government only pays for successful programs that save money. Both are important goals but are also limited. We therefore propose a more expansive vision of the SIB model.

We draw on our experience as the intermediary for a New York City SIB (NYC SIB) project that is attempting to lower recidivism and improve the lives of 16- to 18-year-olds in New York City's Rikers Island jail. This project is the first of its kind in the United States. Over the next year, we will be writing about the experience of designing, implementing, and beginning to test the potential of this program set within a complex and dynamic political and service environment. Other partners include the New York City Department of Correction, the Mayor's Office, Bloomberg Philanthropies, Goldman Sachs, the Osborne Association, Friends of Island Academy, and the Vera Institute of Justice.

What is a SIB?

SIBs are innovative financing arrangements that aim to increase the pool of money available for preventive services. In a SIB, investors provide financing to operate federal, state, or local-run programs that aim to achieve predetermined outcomes. Generally, these outcomes are expected to save government money, for example, by reducing the need for beds in prisons or homeless shelters. The government entity agrees in advance that, if the program meets its goals, it will use the savings to pay back the original investment, plus a return. Usually, an intermediary organization puts the pieces together—identifying appropriate interventions and service providers, making a match between government agencies and investors, helping to structure the financial deal, and monitoring the program as it operates. This is the role MDRC is playing in the NYC SIB. An independent evaluator will confirm that the program has achieved the pre-specified goals and determine whether the government is obligated to pay back the investors.

Using SIBs to Finance Replication of Proven Programs

SIBs have been described as an ideal vehicle for going to scale with proven prevention programs that currently operate on a small scale. Public agencies facing severe budgetary pressure often are caught in a vicious cycle: they must spend money on prisons, shelters, public assistance and other services, leaving less for programs that might reduce the need for such spending in the first place.

SIBs may offer a way out of this bind. However, it is important to consider some complicating factors. Profit-seeking investors will be most interested in social programs or models that are proven—and thus quite likely to produce savings—but identifying “good bets” is easier said than done. Most social programs, including many that are quite well known, have little or no solid evidence behind them. For many others, the available evidence is mixed, limited, or based on problematic evaluation designs. Even successful programs have not necessarily generated impacts of the magnitude necessary to pay for themselves and yield a return for an investor.

The federal government and private foundations have recently begun to articulate a system of tiers to describe the strength of evidence supporting particular social programs. Only those in the highest tier—validated by multiple randomized control trials—are considered “proven,” and there are very few such programs. Therefore, there is a very small pool of potential SIB deals. Perhaps more important, there is a long history of programs that have achieved strong results in small pilots but were not successful when replicated on a larger scale. Social programs and the problems they address are often complex and not well understood. Moreover, interventions have to be delivered within systems—for instance, criminal justice, foster care, or welfare systems—and the rules, regulations, and operating cultures of those systems often vary. When a program achieves positive results, the success may be attributable to a wide range of factors, and it is often difficult to identify exactly why the program worked. This makes it hard to replicate success and, in a SIB context, it puts tremendous pressure on the providers delivering the service and the intermediary responsible for overseeing program implementation. In the NYC SIB, we understood the marketing appeal of calling a program a “safe bet” in trying to attract private investors, but we were also aware of the replication challenges and the resulting risks of failure. Although there was considerable evaluation-based evidence supporting the program model we selected—some might have described it as “proven”—we were careful not to characterize the program in that way and were forthcoming about the implementation and scaling challenges.

Using SIBs to Ensure the Government Pays Only for Successful Programs

The idea of Pay for Success (PFS) is not new: government agencies have been writing performance-based contracts with social service providers for many years. In those contracts, rather than paying providers based on their costs, payment is determined by outputs (such as meeting program enrollment and participation targets) or on outcomes (such as achieving

job placement or job retention goals) or on some combination of the two.

The appeal of such arrangements is clear, particularly when payments are linked to outcomes. However, the record has been mixed, in large part because the outcomes a program achieves can be an unreliable measure of its impact—and impacts are a far more useful indicator of success. Measuring an impact requires accurate information about what would have happened to program participants if they had not received the program’s services. An example of this discrepancy was demonstrated in a study of the Job Corps program, which found little correlation between the impacts that individual programs achieved (as measured in a randomized control trial) and their performance relative to the Department of Labor’s performance measures.

Not addressing the “net impact” question is a significant weakness in the high-stakes world of performance-based contracting because these contracts create powerful incentives for capital-starved service providers to “cream” (target participants most likely to succeed) as a way to maximize revenue. Government agencies have become more sophisticated in trying to minimize the risks of creaming by, for example, referring service populations to contracted providers with “harder to serve” characteristics, or requiring outcomes that are more difficult for people to achieve on their own. While these responses can mitigate the risk of creaming, they seldom eliminate it.

SIBs potentially offer advantages over traditional performance-based contracts. For example, a SIB may provide upfront funding for nonprofit service providers, who cannot wait to be paid after the fact. At least in theory, SIBs also include an independent assessment of the program’s performance. But the SIB structure does not, in itself, address the problem of creaming or the potential mismatch between outcomes and impacts. The only way to ensure that government pays for success is to ensure that the evaluation or validation process compares the outcomes of program participants with those of a reliable comparison or control group.

Expanding the Vision of SIBs

Using SIBs to Spur Innovation and Build Knowledge

Given the dearth of models with strong evidence of effectiveness and the challenges of replicating success, it is important to consider whether SIBs or SIB-like arrangements can be used to spur innovation, build knowledge, and increase the number of truly effective programs. We believe they can.

One might think of tiers of SIBs, corresponding roughly to the evidence tiers described above. The top tier would focus on the small number of programs with strong evidence of effectiveness and demonstrated ability to achieve positive impacts at scale and in diverse settings. In those few cases, the primary purpose of the SIB would be to identify financing to support further replication and expansion. The risk of failure would still exist, but it would be the smallest of the tiers, which might be appealing to certain kinds of investors who are

interested in low-risk ventures and are willing to accept relatively modest rates of return. Commercial banks looking to fulfill Community Reinvestment Act requirements might fit the bill.

A middle tier would include programs that have limited, mixed, or incomplete evidence. Here, the risk of failure is greater, which might warrant a different mix of investors. For example, foundations might need to act as a “backstop” to limit the downside risk for for-profit investors. These might include equity investors, who are prepared to assume higher risks than commercial banks in return for larger returns. The intermediary role would require more detailed program knowledge, and more rigorous and in-depth evaluation would be needed in order to accurately measure program impacts and shed light on the replication/scaling process. Foundations or the federal government might support the evaluations, which would be more elaborate than what might be deemed sufficient for the narrow purpose of determining whether investors should be paid back. For example, it would be optimal to include robust implementation research to understand why replication succeeds in some places and fails in others.

The lowest and riskiest tier would focus on innovation. It would test programs that have a strong theoretical basis and/or promising results from very small-scale studies. In effect, these SIBs would resemble traditional demonstration projects, with tightly controlled implementation and rigorous, in-depth evaluation. The most likely investors for these initiatives would be foundations or the federal government, which have a history of promoting and testing innovation. However, certain kinds of profit-seeking investors might also play a role given that these programs attempt to improve outcomes that may save government money or may simply be something that government is willing to pay for. Rather than selling future promises to pay to investors, it may be feasible for government payments to be reinvested in additional SIBs that focus on innovation, or if the program is successful, the agencies might also agree to pick up the cost of the intervention moving forward so it can continue to run.

Omitting Impact Studies Could Imperil SIBs

In our conversations with potential SIB stakeholders across the country, we have grown concerned that support for high-quality evaluations is not a priority. The pressures to raise sufficient capital to cover the program investment can lead to underfunded evaluations. The political imperative to demonstrate the success of this new financing scheme can create incentives for weaker evaluation designs that are more likely to show positive results but that are spurious.

In this environment, SIBs may forgo plans for serious evaluation and replace them with limited third-party documentation audits. Such a strategy may identify intentionally false or inaccurate reporting, but it will not provide evidence that the program truly led to cost savings. Only in those few cases in which the SIB is replicating an intervention that has been reliably demonstrated to work at scale should SIB parties consider omitting an impact study. In such cases they could perhaps replace it with a combination of outcome measures and an

assessment of fidelity to the model. But even here, investors and government would be left with some uncertainty about effectiveness. Indeed, if there is no risk, no uncertainty, then why would it be in the best interest of government to use a SIB structure? After all, it would cost the government more money not to run a program that saves money.

Broadening the Definition of Success

SIBs have been proposed for programs that are intended to realize government savings in a relatively short time period. These kinds of projects are probably the right place to start in building support for SIBs. However, the goal of most social programs is not primarily to save money but to improve the lives of low-income and at-risk individuals and families. SIBs could be structured to encompass other socially desirable goals that do not lead to government budget savings but do lead to societal improvements, so long as government can decide what it is willing to pay to achieve specific goals. SIBs could be designed to finance a range of different outcomes from increasing high school graduation rates and persistence in college, to improved cognitive and behavioral skills for young children, or better mental health outcomes for adolescents. All of these areas have promising, and perhaps even some proven, interventions with the potential to be scaled up. And additional funding for these kinds of programs is in at least as short supply as funding for programs that may generate short-term savings. But thus far, we have not seriously asked ourselves what we are willing to pay for this kind of success. Whether that amount would be sufficient to cover program costs and pay an acceptable return to investors is an open question worth exploring.

Conclusion

SIBs, as currently described, are a new financing strategy with the potential to attract new money to pay for innovative social programs. At the same time, it is critical to consider how the strategy could be used to continue to build knowledge about what works. There are too few proven interventions, and too many difficulties in replicating even those few programs, to minimize the role of innovation and knowledge-building. Therefore, we have offered a different view of how SIBs could be structured to promote innovation. Unless we consider these and other alternatives, government is likely to end up paying for success that is never realized or the reservoir of SIB-ready ideas will run dry very quickly.

SIBs can also help ensure that government only pays for successful programs, and they are, potentially, a significant advance over earlier PFS approaches. Achieving that goal, however, will require continued support for rigorous evaluation. Finally, if we hope to realize the full potential of SIBs, we must expand our expectations for success beyond immediate government savings to explore how SIBs can be applied to accomplishing other socially desirable goals.

David Butler is a senior adviser at MDRC and has led the development and implementation of many of MDRC's demonstration projects and evaluations in the "health and barriers to employment" policy area. He is currently leading MDRC's ABLE project, the first social impact bond to be tested in the United States.

Dan Bloom directs MDRC's "health and barriers to employment" policy area, which designs and evaluates programs for groups seeking to gain a foothold in the labor market, including former prisoners, disconnected youth, low-income noncustodial parents (usually fathers), welfare recipients, and others. He is currently directing two large-scale federal projects that are testing subsidized and transitional employment programs.

Timothy Rudd is a research associate at MDRC and analyzes the cost and effectiveness of programs across MDRC's policy areas (jails, schools, higher education, employment, and pre-kindergarten) using a variety of frames including benefit-cost analysis. He played a valuable role in the negotiations of the financing for the first US social impact bond.