Innovation in public service delivery: Essays on the Social Impact Bond



Christina Lynn Economy

St. Catherine's College

University of Oxford

Thesis submitted in partial fulfillment of the requirements of the degree of DPhil in Public Policy in the Blavatnik School of Government at the University of Oxford.

22nd April 2021

64,589 words

Abstract

Following the requirements of an article-based dissertation, this thesis offers three separate articles that explore the effect of the social impact bond (SIB) model on processes of public service delivery. SIBs are an outcomes-based approach to funding social services, where typically a private investor gives upfront capital to a service delivery organization and is then paid back by the government based on the achievement of outcomes. Despite over ten years of experimentation, there is a critical gap in our understanding of how SIBs work in practice and the effects of the model on public service delivery. This thesis draws on theories of new public management, i.e. the incorporation of private sector mechanisms in public service delivery, to investigate various aspects of the 'SIB effect' related to financial incentives, performance targets, risk transfer, multi-sector collaboration, and provider flexibility.

In the first paper, I offer a conceptual framework (based on a review of the SIB literature coupled with themes from new public management) for how SIBs might impact public service delivery. I then apply this framework in a case study of a SIB launched in the UK, using in-depth interviews, document review, and observation of front-line staff. I find that the SIB had mixed effects in this case; it enabled rigorous monitoring of performance, availability of additional resources for the service, and outcome goal alignment across stakeholders. However, the SIB model also generated stakeholder conflict, feelings of resource burden and disproportionality by the government partner, loss of democratic accountability, and the pressuring presence of an underlying financial driver.

In the second paper, I investigate whether using a prescribed and replicable (i.e. high-fidelity) intervention insulates providers and front-line staff from potential negative effects of the SIB model (such as perverse incentives for outcome gaming). After laying out a framework for how the SIB model might affect providers and front-line staff, I draw on a case study (using in-depth interviews, participant observation, and document review) to analyze how these effects change under a high-fidelity intervention. I find that while

using a high-fidelity intervention with the SIB model mitigates incentives to game outcome payments, it also restricts the ability of the provider to innovate and deliver flexible, context-driven services. Using a high-fidelity intervention does not protect providers from pressure to meet output targets, which could potentially create other disincentives in service delivery.

In the third paper, co-authored alongside Drs. Eleanor Carter and Mara Airoldi, we investigate how SIB projects have been designed in relation to two of the model's original promises: shifting the focus of public service delivery to achieving impact and transferring risk of poor performance to external investors. By conducting qualitative content analysis on the SIBs launched to date in the US and UK, we analyze how the design of SIB projects varies across two key dimensions: degree to which payment is linked to impact and nature of the working capital. We find that SIB design in practice has deviated from the model's original intent in several important ways. In the UK, but not the US, outcomes that projects pay out for are generally not subject to rigorous validation methods, calling into question whether payments are meaningfully linked to impact. In both the US and the UK, risk mitigation strategies in the repayment structures may limit the intended transfer of risk to external investors.

Through these three papers, I make original and substantive theoretical and empirical contributions to the debate on how social impact bonds affect processes of public service delivery. I do this by: 1) providing an original conceptual framework from which to analyze the effect of SIBs on systems of public service delivery, 2) laying out the conceptual arguments for how the SIB model could affect provider and front-line staff, and extending this conversation to include differences between types of interventions, 3) furthering the field's empirical understanding of how the SIB model impacts public service delivery and front-line staff through an in-depth case study, surfacing a unique and original discussion of the risks of disproportionality, and 4) offering the first systematic investigation of SIB design over time, comparing the SIB model in practice to its original promises. These contributions have practical implications for policymakers deciding how best to deliver public services.

Acknowledgments

This thesis would not have been possible without the guidance and support of so many.

A huge thanks to my supervisors Mara Airoldi and Peter Kemp. Mara, your unending encouragement and mentorship has made my time as a doctoral student incredibly rewarding. I am forever grateful for your generosity with both your time and brainpower; your insights have improved this thesis immeasurably. Peter, your guidance, always delivered with thoughtfulness, has been invaluable throughout this process and for my training as a researcher.

Much gratitude to my colleagues, mentors, and friends at the GO Lab. Elle, Clare, Nigel, Andreea, Tanyah, Franziska, Grace, Leigh, Mehdi, and Ruth - thank you for all of the encouragement, guidance, and much needed tea breaks that have made my years at Oxford exceptionally special.

Thanks are due to the kind and clever minds who have provided me feedback on this work, including Anne Davies, Carolyn Heinrich, Alec Fraser, David Humphreys, Martin Williams, Kent Weaver, and Gwyn Bevan. Thanks also to Gloria Gong, for being a supportive mentor, thought partner, and friend over the past several years.

I must also of course thank the individuals involved with the Essex case study, who generously gave their time and insights as part of this thesis.

To my lovely friends, including Alice, Helen, and Sophie, thank you for always being there for me with laughs and adventures. Thanks also to my fellow DPhil students, who have provided feedback, advice, and helped me shape my ideas throughout this process.

A huge thank you to my family in Hawaii for their unwavering love and support. Mom, thank you for always being there with a hug. Dad, thank you for imparting all your pearls of wisdom. My sisters, Louise, Jackie, Jamie, and Mary, thank you for surrounding me with such an incredible force of love and strength. My nieces and nephews, Kade, Taetum, Hercules, and Autumn (the last two of which were born during this process!), thank you for all of the giggles. And to my brother Koa, during this time you navigated your cancer battle with more grace, strength, and humor at just 15 years old than I would have ever thought possible. Thank you for all the joy that you bring into my life. You are indeed Perfect (featuring Beyoncé).

Finally, I could not have completed this thesis without my partner and my person, Mike. In fact, I may not have even started. Mike, thank you for believing in me. Every day you make me feel so loved and supported. My face lights up thinking about the rest of our lives together.

Table of Contents

ABSTRACT ..... **I**

ACKNOWLEDGMENTS ..... **III**

TABLE OF CONTENTS ..... **IV**

List of Figures VI
List of Tables VI
List of Abbreviations VII

INTRODUCTION **1**

Policy Context 2
Hybridity in social service delivery *2*
The rise of the Social Impact Bond *6*
Impact of the SIB Model on Public Service Delivery: Evidence to date & knowledge gaps 10
Theoretical Background 11
Aims and Objectives 12
Research Questions and Motivations 13
Methods 15
Case study *17*
Qualitative content analysis *20*
Significance 22
Contents of Thesis 23

LITERATURE REVIEW **26**

Introduction 26
Roots in New Public Management 27
The 'SIB Effect': An ongoing debate 30
On the system of public service delivery *32*
On providers and front-line staff *46*
On service users *50*
Intended Contribution to the Literature 53

ARTICLE ONE - MORE THAN JUST A CLASH OF CULTURES: USING SOCIAL IMPACT BONDS TO DELIVER PUBLIC SERVICES **57**

Abstract 57
Introduction 57
Conceptual Framework: How a SIB might impact public service delivery 61
Methods: Using the case of Essex 67
Findings: How the SIB impacted service delivery in Essex 73
Multi-sector partnership: facilitating collaboration or creating conflict? *75*
Data monitoring system: actively managing the project, but at what cost? *80*
External investment: flexible funding or locked into a rigid contract? *83*
Payment for outcomes contract: aligned to outcomes or to generate returns? *85*
Discussion: What can we learn from Essex 88
Conclusion 93

ARTICLE TWO - THINKING OUTSIDE THE BLACK-BOX: USING A SIB WITH A HIGH-FIDELITY INTERVENTION **95**

Abstract: 95
Introduction: 96
Conceptual Framework: SIB effect on front-line staff 101
Methods: Case study of a SIB with a high-fidelity intervention 105
Findings: How the Essex SIB affected front-line staff 110
Performance Management *110*

Support Networks ... *112*
Flexibility ... *114*
Incentives ... *115*
DISCUSSION: WHEN TO USE THE SIB MODEL? ... 117
CONCLUSION: ... 122

ARTICLE THREE - HAVE WE 'STRETCHED' SOCIAL IMPACT BONDS TOO FAR? AN EMPIRICAL ANALYSIS OF SIB DESIGN IN PRACTICE ... 124

ABSTRACT: ... 124
INTRODUCTION: PROMISES OF THE SIB MODEL ... 125
CONCEPTUAL FRAMEWORK: A CODING FRAME TO ANALYZE VARIATION IN SIB DESIGN ... 129
Payment made for impact ... *130*
Independent/at-risk capital ... *133*
METHODS: QUALITATIVE CONTENT ANALYSIS OF UK AND US SIBS ... 136
FINDINGS: SIB DESIGN IN PRACTICE ... 138
Payment for Impact ... *138*
Nature of the capital ... *143*
DISCUSSION: HAVE WE STRETCHED SIBS TOO FAR? ... 147
CONCLUSION: ... 152

CONCLUSION ... 154

RESEARCH AIMS ... 154
SUMMARY OF FINDINGS ... 156
SIB effect on the system of public service delivery ... *156*
SIB effect on front-line staff ... *160*
Evolution of SIB design over time ... *163*
STRENGTHS AND LIMITATIONS ... 168
CONTRIBUTION OF THE THESIS ... 173
POLICY IMPLICATIONS AND FURTHER RESEARCH ... 176
Implications for the SIB model as a policy option ... *176*
Implications for SIB design ... *179*
Future research ... *181*

REFERENCES ... 184

APPENDICES ... 198

APPENDIX 1: CO-AUTHORSHIP DECLARATION ... 198
APPENDIX 2: ETHICS APPROVAL, ESSEX CASE STUDY ... 199
APPENDIX 3: PARTICIPANT INFORMATION SHEET, ESSEX CASE STUDY ... 200
APPENDIX 4: WRITTEN CONSENT FORMS, ESSEX CASE STUDY ... 203
APPENDIX 5: LIST OF KEY STAKEHOLDERS, ESSEX CASE STUDY ... 206
APPENDIX 6: INTERVIEW TOPIC GUIDES, ESSEX CASE STUDY ... 207
APPENDIX 7: NVIVO THEMATIC CODING NODES, ESSEX CASE STUDY ... 210
APPENDIX 8: CODING FRAME DEFINITIONS, QUALITATIVE CONTENT ANALYSIS ... 212
APPENDIX 9: SENSITIVITY ANALYSIS, QUALITATIVE CONTENT ANALYSIS ... 214

List of Figures

Figure 1: Three-way relationship of a social impact bond 6
Figure 2: Impact Bond Map, March 2021 8
Figure 3: Improving and Impeding Mechanisms 66
Figure 4: Impact of SIB features in Essex 74
Figure 5: Dimensions of variation in SIBs (adapted from Carter 2020) 129
Figure 6: Payment for Impact Scores, All Projects 138
Figure 7: Payment for Impact Scores, by Country 140
Figure 8: Payment Model Scores, by Country 141
Figure 9: Validation Method Scores, by Country 141
Figure 10: Payment for Impact Scores, UK 142
Figure 11: Payment for Impact Scores, US 142
Figure 12: Nature of Capital Scores, All Projects 143
Figure 13: Nature of Capital Scores, by Country 144
Figure 14: Primary Investor Scores, by Country 145
Figure 15: Repayment Structure Scores, by Country 145
Figure 16: Nature of Capital Scores, UK 146
Figure 17: Nature of Capital Scores, US 147
Figure 18: Payment for Impact Scores, All Projects 165
Figure 19: Payment for Impact Scores, by Country 166
Figure 20: Nature of Capital Scores, All Projects 166
Figure 21: Nature of Capital Scores, by Country 167

List of Tables

Table 1: Four key features of a SIB 8
Table 2: Comparing contracting models 9
Table 3: Overview of themes explored across three articles 13
Table 4: Literature on SIB effect across levels 32
Table 5: Reflections on SIB literature across levels 54
Table 6: Four key features of a SIB 62
Table 7: How particular features of SIBs might impact public service delivery 65
Table 8: Summary of findings from the Essex SIB 74
Table 9: How the SIB structure might affect front-line staff 104
Table 10: SIB effects on front-line staff in Essex 118
Table 11: Payment for Impact Operationalization 132
Table 12: Payment Linked to Impact Composite Score 133
Table 13: Nature of the Capital Operationalization 135
Table 14: Nature of the Capital Composite Score 136
Table 15: Overview of themes explored and high-level findings across three articles 156
Table 16: How particular features of SIBs might impact public service delivery 158
Table 17: Summary of Findings from the Essex SIB 159
Table 18: How the SIB structure may affect front-line staff, according to existing evidence 161
Table 19: Effects on front-line staff of high-fidelity SIB intervention in Essex 162
Table 20: Payment for Impact Operationalization 163
Table 21: Nature of the Capital Operationalization 164

List of Abbreviations

CNCS	Corporation for National and Community Service
DCLG	Department for Communities and Local Government
ECC	Essex County Council
FFS	Fee for Service
GLA	Greater London Authority
GO Lab	Government Outcomes Lab, Blavatnik School of Government
GPL	Government Performance Lab, Harvard Kennedy School of Government
HMP	Her Majesty's Prison
INDIGO	International Network for Data on Impact and Government Outcomes
KPI	Key Performance Indicator
LCF	Life Chances Fund
LGA	Local Government Authority
MST	Multi-Systemic Therapy
NPG	New Public Governance
NPM	New Public Management
OECD	Organization for Economic Cooperation and Development
PbR	Payment by Results
PFS	Pay for Success
RCT	Randomized Control Trial
SIB	Social Impact Bond
SIPPRA	Social Impact Partnerships to Pay for Results Act
SLB	Street-Level Bureaucrats
SPV	Special Purpose Vehicle
UK	United Kingdom of Great Britain and Northern Ireland
US	United States of America

Introduction

This thesis contributes to debates in public management on how to best organize the delivery of public services. Policymakers and academics alike have long struggled with questions around how to more effectively, efficiently, and equitably support vulnerable populations. An often-recurring narrative in tackling these challenges is that the public sector would benefit from private and third sector help, thus giving way to a series of reforms characterized by hybridity in social service delivery (Skelcher 2007). New Public Management (NPM), one such hybrid approach focused on incorporating private sector elements to improve the efficiency of public services, gained traction in the 1980s and persists today (Hood 1991). However, questions remain around these types of hybrid instruments and the impact of NPM strategies on public service delivery; under what contexts do they actually work to effectively and equitably support vulnerable populations? This thesis makes empirical and theoretical contributions to this literature by focusing on the specific model of the social impact bond (SIB).

The SIB model is an outcomes-based approach to funding social services, where the government commissions services and agrees to pay on specific outcomes, a service provider delivers an intervention, and private investors provide the upfront working capital and are repaid based on the achievement of outcomes. Although the SIB model has been used in over 200 projects worldwide over the past ten years (INDIGO 2021), there are critical gaps in our understanding of how the model impacts service delivery. What is the impact of the SIB model on the system of public service delivery, on provider organizations and front-line staff, and on service users?

Following the requirements and format of an article-based dissertation, this thesis explores the impact of the SIB model on various levels of service delivery through three separate articles. The thesis focuses on the UK and the US, the two countries with the greatest number of SIB projects who have paved the way in exploration of the model. The first two articles are based on a case study of a social impact bond launched

in Essex, United Kingdom. The first article examines the effect of the SIB at a systems level, defined as the network of interacting organizations and individuals that contribute to service delivery (Johnson, Fortune, and Bromley 2017). The second article examines the effect of the SIB at an organizational level on the provider and front-line staff. The third and final article, co-authored alongside Eleanor Carter and Mara Airoldi, relies on qualitative content analysis of all SIBs launched in the UK and the US from 2010-2020 (n=114) to investigate how the model has evolved over time, and whether it follows through on some of its key promises for altering the structure of service delivery.

The following introduction chapter lays out the policy context in which social impact bonds were developed and introduced. It then describes what we know about the impact of SIBs so far, situated in the related theoretical and empirical literature on new public management. Next, the research questions and objectives of the thesis are discussed. I then lay out the methods employed, including a description of the case study design and the qualitative content analysis. This is followed by a discussion of the thesis' significance to policymakers and academic scholars. This *Introduction* chapter concludes by laying out the contents in each remaining chapter of the thesis.

Policy Context

Hybridity in social service delivery

Hybrid configurations are defined as "organizational arrangements that use resource and/or governance structures from more than one existing organization" (Borys and Jemison 1989). In social service delivery, hybridity refers to the "indistinct boundary" between the interests of the public and private sectors that results from "close engagement of business and not-for profits in the governmental process" (Skelcher 2007). In recent history, hybridity has been a central tenet of various policy platforms around the world. Two such platforms, New Public Management and Big Society, paired with austerity measures, led to

further exploration of outcome-based commissioning and ultimately the development of the SIB model in the UK.

New Public Management and 'Big Society'

To fully understand the context in which the SIB was developed, it is helpful to discuss two hybrid-centric policy manifestos, the first global and unremitting (NPM), and the second contained to the UK and short-lived (Big Society). Following economic recession in the early 1980s, New Public Management (NPM) - the idea of applying private-sector mechanisms to increase the efficiency of public services - came to the forefront of the reform agenda (Hood 1991). NPM aimed to improve the quality of public services, while simultaneously saving the government money by introducing market discipline and aligning public and private sector goals (Pollitt and Bouckaert 2004). These NPM ideals manifested as strategies to increase market competition, rely more heavily on hands-on management, and focus on explicit measurement of performance (Ferlie et al. 1996). As a result of these reforms, strongly associated with the private sector, government agencies around the world relied more heavily on private actors to "design, manage, and deliver public policy" (Skelcher 2007).

In 2010, the UK's Prime Minister David Cameron announced a political agenda coined "Big Society," which focused on reducing the role of central government and giving more power to local communities, the voluntary sector, and private actors (Kisby 2010). Similar in nature to the ideals of NPM, the term 'Big Society' referred to a set of policies aimed at reducing government spending and relying more on third sector and private actors; including by opening up more public service contracts to the private and voluntary sector (Kisby 2010). The policy also emphasized new ideas related to social investment, encouraging a new social investment market focused on shifting the responsibility of social programs away from the state (Dowling and Harvie 2014). Logic behind these policies seemed to suggest that local communities, including voluntary and private sector providers, would be able to deliver services more efficiently and effectively than the public sector by giving more power to the people (Kisby 2010). Simultaneously to the

political agenda of "Big Society," which already focused on reducing government spend, the UK government entered into an aggressive era of austerity. As more responsibility was transferred from central to local governments a la "Big Society" policies, local governments also saw their budgets dramatically slashed (Lowndes and Pratchett 2012).

Austerity and outcome-based commissioning

Austerity, characterized by "sustained and widespread cuts to government budgets," has been the defining feature of British public policy since 2010 (Gray and Barford 2018). From 2010-2020, spending on welfare payments, housing subsidies, and other key social services has been cut by over £30 billion (Mueller 2019). Local government budgets were hit hard, with a nearly 60% cut in central government funding that has created tremendous challenges in providing key services, most acutely for adult social care, children services, and homelessness support (LGA 2019). To make matters worse, cuts have not been evenly spread across the country and increasing inequality between local authorities has resulted (Gray and Barford 2018). For example, the most deprived local authorities have faced the greatest cuts to children's services funding, particularly to preventative family support and early intervention services (Webb and Bywaters 2018). With austerity measures resulting in shrinking public budgets, local governments have been pushed to find new ways to fund vital social services (Neyland 2018). Paired with a 'Big Society' political agenda, including incentives to pursue forms of social investment and involve non-public actors, strategies of outcome-based commissioning came to the fore.

In the context of austerity-driven pressure to "do more with less" with government funding, outcome-based commissioning strategies gained traction (Albertson et al. 2018). Outcome-based commissioning refers to a method of government financing that involves purchasing social programs by paying for outcomes, rather than paying for services delivered as with a traditional fee-for-service arrangement (Albertson et al. 2018). For example, with a traditional fee-for-service contract, a government agency may pay a homelessness provider to open a shelter with a certain number of beds and provide support to individuals experiencing

homelessness to find permanent housing. In an outcomes-based commissioning arrangement, however, the government agency may pay the provider based on the number of individuals who exit the shelter and remain stably housed for a certain length of time (thus paying for outcomes achieved, rather than services delivered). The logic of outcomes-based commissioning is that it could save government money by only spending resources on "what works" by paying for outcomes after they have been achieved, rather than paying for programs or services upfront (Bovaird and Davies 2011).

In recent years, government agencies have been increasingly experimenting with outcomes-based commissioning tools (Albertson et al. 2018). Proponents of outcome-based commissioning strategies see it as "an approach to public service management that seeks to improve value and impact" through a focus on the tangible change created by a public service or an intervention (GO Lab 2020a). Paying for outcomes is often seen as part of an effort to move to more strategic commissioning practices to improve effectiveness and efficiency of services delivered (Bovaird and Davies 2011). Central and local levels of UK government committed to a "highly ambitious course of steering its interventions through their expected impact on outcomes, not simply through programmes of action" (Bovaird and Davies 2011). While there are difficulties with this outcomes focus, in particular challenges around specifying outcomes, poor measurement systems, and difficulty attributing outcomes to the actual program (Pollitt and Bouckaert 2004), the central and local levels of UK government have forged ahead testing new types of this outcome-based approach.

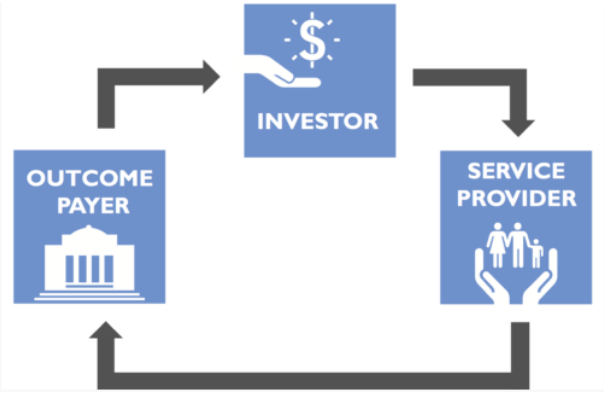
One common outcome-based commissioning approach is called payment by results (PbR). Payment by results refers to the process of paying for services based only on the achievement of certain outcomes, rather than paying for services delivered (GO Lab 2020a). In 2010, the Coalition Government accelerated policies using PbR, implementing the model to fund a range of public services in areas including employment, health, criminal justice, and children services (Moss 2010). Two of the most well-known UK PbR initiatives are the Work Programme, launched by the Department for Work & Pensions in 2011 to help over 2 million

people gain employment, and the Troubled Families Program, launched in 2012 by the Department for Communities and Local Government to provide support to over 500,000 families with a range of needs (Albertson et al. 2018). However, PbR was criticized for favoring large providers who have the funds to provide interventions without upfront financing (Dowling 2017). In response, pairing the idea of PbR with desires to increase social investment and enable providers regardless of size, the social impact bond (SIB) model was developed.

The rise of the Social Impact Bond

As a result of various policy manifestos (New Public Management, Big Society, and austerity) colliding, alongside experimentation with outcomes-based commissioning, the Social Impact Bond (SIB) was born in the UK in 2010. In the SIB model, generally a government commissions services and agrees to pay on specific outcomes, a service provider delivers an intervention and undergoes an evaluation to determine the impact, and private investors provide the upfront capital and are repaid based on the achievement of outcomes. In its most basic form, a SIB can be understood as a three-way relationship between a government commissioner, social sector service provider, and outside investor.

Figure 1: Three-way relationship of a social impact bond

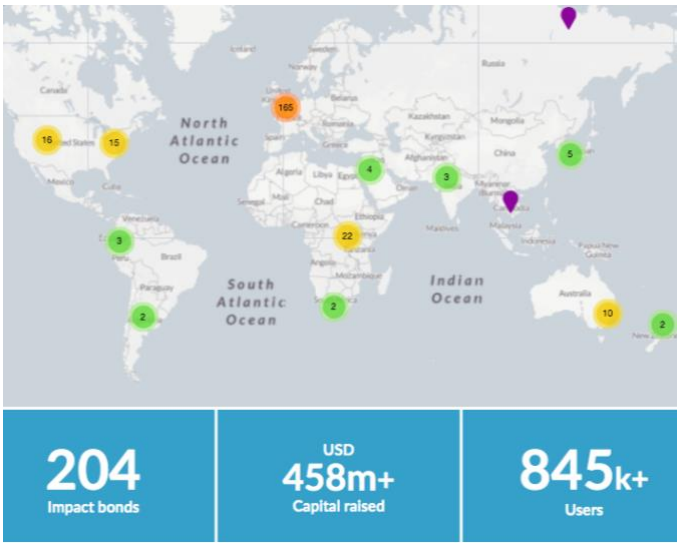


(Instiglio 2021)

Social impact bonds were launched as a way to incorporate private-sector ideals into service delivery through outcomes-based contracts, whilst allowing for the participation of a third-sector provider. For market discipline enthusiasts, SIBs are an embodiment of the potential for public-private hybrids, bringing in private sector techniques to achieve public good without risking public money. Hands-on professional management, private-sector involvement, explicit performance standards, and greater discipline in use of resources are all features of the NPM movement (Hood 1991) that are reiterated in the SIB model. By relying on a social investment-driven partnership between the state, private sector, and third sector, the SIB model also caters to the ideals of Cameron's 'Big Society' agenda. Finally, by only paying for agreed upon outcomes while bringing in the private sector to fund services upfront, the SIB was seen as a promising model to deal with the repercussions of local government budget cuts due to austerity measures.

The UK launched the first SIB in 2010, which aimed to reduce recidivism, and has since implemented over 85 SIB projects across a wide range of issue areas (INDIGO 2021). In 2012, following experimentation in the UK, a large market for SIBs also began to emerge in the US. To date, the US has launched 27 SIB projects (commonly called Pay for Success projects) totaling over $190m in upfront capital raised (INDIGO 2021). While in the UK motivations for developing the SIB model involved funding social services in an era of austerity, the US emphasized the potential for the model to test and develop new innovative programs (Albertson et al. 2018; Neyland 2018). In both countries, however, the SIB model was presented as a new way to incorporate private sector strategies (and resources) in to public service delivery, and through this shift the focus of service delivery to achieving impact while protecting taxpayer money by only paying for what works (Fraser et al. 2016; Albertson et al. 2018). Across the world, over 200 SIBs have been launched to tackle many challenging social problems, including recidivism, early childhood health and education, homelessness, and substance misuse (INDIGO 2021).

Figure 2: Impact Bond Map, March 2021



(INDIGO 2021)

So, what, exactly, characterizes a social impact bond? While the precise definition of a SIB varies across stakeholders, based on a review of the literature and the existing 206 SIB projects (INDIGO 2021) I suggest four basic features characteristic of projects considered to be a SIB: multi-sector partnership, data monitoring system, payment for outcomes contract, and external investment.

Table 1: Four key features of a SIB

SIB Feature	Description
1. Multi-sector partnership	Generally, at least three stakeholders (typically government, provider, and investor) work together to launch a SIB project. There may be multiple government agencies, providers, and/or investors involved with the deal. There is often also an intermediary and an evaluator involved.
2. Data monitoring system	Systems of data collection and analysis are set up in order to track outcomes achieved in SIB projects. These systems are used to hold provider(s) accountable for meeting performance metrics and are used to determine payments made.
3. Payment for outcomes contract	A contract is signed by relevant parties that establishes a series of payable outcomes. The government commissioner (or other paying agency) only pays for services delivered if and when these pre-specified outcomes are achieved.
4. External investment	A private investor provides upfront at-risk capital to the service provider to fund the project and is repaid based on the achievement of outcomes.

The table below aims to clarify the distinctions between SIBs, PbR, and fee-for-service contracting models. In typical fee-for-service arrangements, payments to providers are made based on the services delivered. However, in both PbR and SIB arrangements, payments are made based on outcomes, at least in part. A key promise of the SIB model is a high-degree of performance management, spear-headed by the presence of the external investor bringing in private-sector management mechanisms (Fraser et al. 2016). Typical government contracts financed through fee-for-service arrangements or PbR do not necessarily include a high-degree of performance management, but this might depend on the culture of the government agency or provider organization. In the SIB model, upfront capital is provided by an actor other than the government and provider partner, namely an external investor. This upfront capital is promised to be at-risk, i.e. it is only repaid if the project is successful in achieving the agreed upon outcomes (Instiglio 2021). In both fee-for-service arrangements and SIBs, the upfront costs of service provision are secured within the contracting model, provided by the government agency in the case of fee-for-service and provided by the investor in the case of SIBs.

Table 2: Comparing contracting models[1]

	Fee-for-service	Payment by Results*	Social Impact Bonds*
Payment made for outcomes, not services delivered	x	✓	✓
High degree of performance management	?	?	✓
Data collection systems to track outcomes	?	✓	✓
Upfront capital from external investor is at-risk based on performance	x	x	✓
Service provider costs are funded upfront	✓	x	✓

[1] While this chart is original, guidance was taken from a similar chart in (Loder 2011, 44)
* Payment by results and social impact bonds are both types of outcomes-based commissioning approaches.

Impact of the SIB Model on Public Service Delivery: Evidence to date & knowledge gaps

Despite ten years of experimentation with the model, existing research does not paint a clear or complete picture of how SIBs impact public service delivery in practice. And, as one might expect when core public services are financed with the potential for private-sector returns, the SIB model is steeped in debate. The "strategically ambiguous" nature of the model (Tan et al. 2019) has given rise to a number of paradoxical claims on the SIBs impact on public service delivery (Maier, Barbetta, and Godina 2018). Proponents of the SIB argue that the model shifts the focus of service delivery to measuring and achieving impact, introduces market discipline, encourages innovation, and creates multi-sector partnerships that improve service delivery (Fraser et al. 2016). Critics of the SIB model in turn argue that this approach may actually impede service delivery by introducing profitization values, encouraging perverse incentives, and necessitating high transaction costs (McHugh et al. 2013; Warner 2013; Rowe and Stepheson 2016; Roy et al 2017).

Determining the 'SIB effect,' how the model impacts public service delivery, is both a challenging and nuanced task. In a review of empirical evaluations of SIB projects (n=16), Fox and Morris (2019) find that existing studies "tend to conflate the outcome based commissioning mechanism with the intervention or services that are funded," and therefore are unable to "assess the contribution of these separate elements" on processes of service delivery or outcomes for service users. The authors highlight the need for more high-quality evaluations, particularly those that focus on the effects of the SIB model itself (Fox and Morris 2019). Furthermore, Lowe et al. (2019) point out that most studies to date have focused on "high-level outcomes" of SIBs rather than "examining the detailed complex set of interactions" between stakeholders. The SIB can have an effect on a number of different levels, as they are multifaceted policy tools that involve complex relationships and dynamics across various sectors (Lowe et al. 2019). As a result, the 'SIB effect' can vary greatly depending on what level and from whose perspective it is analyzed.

More research is needed on the impact that the SIB model has on public service delivery, at various levels of implementation. At the macro-level, what is the impact of the SIB on the system of public service delivery? At the meso-level, how does the SIB impact particular organizations, including service providers and front-line staff? At the micro-level, what are the effects of using a SIB model on service users?

The following thesis aims to contribute to our working body of knowledge related to the first two of these key questions, investigating the impact of the SIB model at the levels of the system of public service delivery and front-line service provider staff.[2]

Theoretical Background

Investigating questions related to the impact of the SIB structure on public service delivery has wider relevance to the field of public management. Involving multiple actors in the delivery of social services, SIBs are one model of hybrid governance that can be viewed as a public-private partnership. The literature on public-private partnerships and hybrid governance arrangements elevate debate on the cost and quality impacts of such arrangements, questions around accountability and transparency, and concern over subverting public sector goals for private gain (Skelcher 2007; Seidman 1988; Moe 2001). The involvement of a non-profit service provider brings in additional questions related to how these public management reforms affect the nature of third sector organizations and their role in relation to the state (Bode and Brandsen 2014).

In order to interrogate these debates in the context of the SIB model, it is helpful to turn to lessons from New Public Management. NPM, as introduced earlier in this chapter, suggests that incorporating private-sector mechanisms and increased market discipline will improve the efficiency and effectiveness of the

[2] Time and resource constraints of the thesis meant that I could not explore all three questions. Whilst there is very little known on the impact of the SIB model at the level of service users and intended beneficiaries, marking a large gap in the existing literature, this question is beyond the scope of the thesis.

public sector (Hood 1991; Pollitt and Bouckaert 2004). SIBs can be seen as an expression of NPM ideals, particularly with the SIB's focus on bringing private-sector actors and new forms of market discipline to the public sector (Dayson, Fraser, and Lowe 2020). One of the primary narratives used to justify the development and experimentation with the SIB model, that private sector management techniques can help improve public sector service delivery (Fraser et al. 2016), closely mirrors the values of the NPM theoretical framework (Hood 1991). Proponents of NPM would suggest that the SIB model introduces new accountability and oversight by tying funding to explicit measures of performance, and furthermore improves performance by incorporating more hands-on management techniques (Balk 1996; Minogue 2001; Pollitt and Bouckaert 2004). NPM critics, in turn, would suggest that the same features of the SIB model, tying funding to explicit measures of performance and introducing more hands-on management, lead to corrupt practices and impedes front-line staff capacity to provide quality services (Nethercote 1989; Barberis 1998; Doig 1997; Minogue 2001).

New Public Management provides a helpful theoretical underpinning for analyzing the impact of the SIB model on public service delivery. The roots of the SIB model in the NPM framework are discussed more thoroughly in the *Literature Review* chapter.

Aims and Objectives

The aim of this research is to offer unique theoretical and empirical contributions to the debate on how social impact bonds affect processes of public service delivery. In the form of three stand-alone articles, I seek to investigate different aspects of the hybrid structure of the SIB model and the consequences to various actors involved.

The objectives of this thesis are:

1. To better understand how the different features of the SIB arrangement - a multi-sector partnership, data monitoring system, payment for outcomes contract, and external investment - combine to affect systems of public service delivery
2. To better understand whether and how the SIB model implemented with a high-fidelity intervention affects the behavior and perceptions of front-line staff in regard to support networks, performance management, flexibility, and incentives
3. To better understand the implications of variation in the design of SIB projects across two key dimensions: payment made for impact and nature of the upfront capital

Following on from these objectives, the table below provides an overview of the themes explored across the thesis' three articles.

Table 3: Overview of themes explored across three articles

SIB Features	Multi-sector partnership	Data monitoring system	External investment	Outcomes contract
Article	**Themes explored**			
Article 1: SIB effect on system	Stakeholder relationships	Performance management & resource allocation	Flexibility & democratic accountability	Outcome alignment & incentives
Article 2: SIB effect on front-line staff	Support networks	Performance Management	Flexibility	Incentives
Article 3: Variation in SIB design			Nature of upfront capital (risk transfer)	Payment linked to impact

Research Questions and Motivations

The driving motivation behind this research is to explore how government agencies can more effectively (and efficiently) support vulnerable populations. Over the past ten years, the SIB model has been put

forward as one potential way to achieve this goal. Beginning in 2015, I worked directly with state and local government agencies in the US who were exploring the SIB model. As a Fellow for the Harvard Kennedy School Government Performance Lab (GPL), I provided hands-on technical assistance to government leaders who were pursuing the SIB model to provide more equitable access to early years education for children and families in their jurisdictions. Working firsthand on SIB projects, embedded within government agencies, provided me with a unique perspective on the inner workings on the SIB model and insight on government-side trade-offs, motivations, and priorities when considering these approaches. Reflecting on these experiences with SIBs, I felt many of the issues facing government policymakers were not new, and that there was much to learn from the wider field of public management. While most of my work was project-specific and focused on the details of the particular context, I wondered from a more holistic perspective what the impact of the SIB model was on systems of service delivery and I desired to synthesize my understanding of the model in a way that could contribute to a growing body of knowledge. In general, I wondered what value-add does the SIB bring to the various partners (government, provider, and service user) otherwise involved in service delivery? Under what contextual circumstances might it be a good idea for government agencies to pursue the SIB model? And conversely, what are the drawbacks of the SIB model? Does using the SIB approach have negative consequences for government, provider, or service user partners?

The following thesis seeks to answer some of these questions. Bringing a practice-informed perspective, I seek to better understand the impact of the SIB model - which brings together a multi-sector partnership, data monitoring infrastructure, outcomes contract, and external investment - on different levels of service delivery. The 'SIB effect' merits further examination in the scholarly literature as it relates to a number of wider public management questions, including the use of financial incentives through outcomes-based contracting, collaborative governance structures between private, public and voluntary actors, and levels of discretion afforded to front-line staff. Exploring the 'SIB effect' is also of critical importance to

practitioners considering the model, to aid in their efforts to provide effective and equitable social supports to the constituents they serve.

Through three standalone articles, this thesis investigates the following research questions:

1. What is the effect of using the SIB model on the system of public service delivery?
2. What is the effect of using the SIB model on service providers and front-line staff?
3. How has the design of the SIB model evolved over time in respect to two key promises: payments made for impact and independent, at-risk capital?

Methods

Given the nature of the research questions above, qualitative methods are best suited to fully unpack the various relationships, mechanisms, and components of the SIB model and tease out what occurs in practice on the ground. The specific methods of this thesis include an in-depth case study of a SIB project launched in Essex, and qualitative content analysis of all the SIB projects launched to date in two countries. These methods are appropriate in order to investigate the little-known phenomena of complex interactions characteristic of the SIB model, in a naturalistic setting that allows for data gathering and analysis to occur simultaneously (Harris et al. 2009). This qualitative framework, rather than demonstrating causality, aims to increase understanding of a situation and gather rich data from which ideas are primarily induced (Bryman 2015). The investigation does rely on an integrated approach combining both inductive and deductive reasoning, beginning with theory derived from observations of the social phenomenon that is the SIB model (deductive), but allowing space for the empirical evidence to inform the development of the theory as data was collected (inductive) (Bryman 2015).

In conducting this research, it is important to address my own epistemological and ontological positions. Ontology is defined as the theory of what exists, or the form and nature of reality (Guba and Lincoln 1994).

Epistemology relates to sources of knowledge, and the relationship between what can be known and the researcher (Guba and Lincoln 1994). When studying a complex public management model such as the SIB, with varied partners and perspectives, assuming a single objective "reality" would fall short of capturing many of the nuances and pieces of knowledge that are available. Aligning with a constructivist epistemological framework, I would argue that reality is subjective, context is important, and data interpretation is necessary to draw meaningful conclusions (Fosnot 2013). In line with this framework, qualitative methods can help to uncover knowledge by providing rich details and a deeper understanding of the SIB phenomena (Fosnot 2013).

My previous work experience providing technical assistance to government agencies pursuing SIBs, as well as my current affiliation with the Government Outcomes Lab (GO Lab) at the University of Oxford (which conducts research and provides technical guidance on SIB projects) influences the way I view SIBs and the assumptions I hold about the model. Acknowledging this through a lens of reflexivity suggests that my own beliefs and values have likely influenced the research process and conclusions reported in this thesis (Berger 2015). Efforts were taken to mitigate my potential biases and ensure that I challenged any preconceptions of the SIB model that I brought from my own experiences, for example by triangulating data from various sources and creating multiple opportunities for experts in the field to interrogate the research methods and findings of each paper. My supervisors, colleagues at the GO Lab and fellow PhD students with a diverse range of experiences, SIB investors and government commissioners, and academics at other institutions all provided feedback throughout the course of this thesis and challenged me to ensure all claims were well-evidenced. In addition, findings from articles one and two were presented and discussed at a learning event at Essex County Council to all partners in the SIB project and other SIB stakeholders.

Case study

The first two articles of the thesis rely on an in-depth case study of a SIB launched in Essex, United Kingdom to provide Multi-Systemic Therapy (MST) to adolescents at-risk of being placed in out-of-home care. While a detailed description of the case and methods are provided in each article, the following provides an overview of the case study used in articles one and two.

In response to a growing number of children entering statutory (out-of-home) care, the Essex SIB was launched in 2012 and ran until December 2018. During this time, over 380 young people received MST through the SIB model, in an effort to preserve family stability and prevent young people from entering out-of-home care. Essex County Council agreed to pay investors based on the number of days children receiving the MST service spent outside of care. As a result, the council paid approximately £7.1m to investors in outcome payments for a program that cost £5.8m (ECC 2012; GO Lab 2019). Approximately 19% of the young people who received MST entered care after the services ended[3] (Drew et al. 2019). Stakeholders assumed a rate much higher rate than 19% of children accessing MST would enter care after services ended, and as a result, the SIB far surpassed its outcome targets and investors were repaid in full plus a return.

At the time of this research, the Essex SIB was in its final stages of delivery, making it a useful case study through which to investigate the impact of the SIB model throughout the entire process of service delivery. This timing allowed the dual benefit of a) being able to actively observe and conduct research while services were still being delivered, while b) being able to analyze the entire SIB process as it completed service delivery at the end of 2018. A case study design has the advantage of providing a rich and holistic picture of a complex case, allowing for in-depth examination of events, processes, and relationship dynamics

[3] At the time of analysis, 302/388 young people served by MST had at least one year of post-MST data. 56 of these children entered care after exiting MST services (19%).

(Bennett and Checkel 2014; Cartwright 2017). Other scholars who have studied the SIB model argue that single case study designs are useful when studying such new phenomena (Lowe et al. 2019), that small n designs are well-suited to evaluating policy instruments such as SIBs (Fox and Morris 2019), and qualitative case studies can help to explore questions around SIB implementation (Fraser and Mays 2020). Detailed and contextualized case studies allow for the exploration of questions relating to how, why, and for whom a complex intervention, such as the SIB model, has had impact (Yin 2013; Pawson and Tilley 1997).

The case study was structured with guidance from the framework of a process evaluation. A process evaluation seeks to determine how successfully an intervention or model was implemented (Steckler and Linnan 2002). Typically, process evaluations are applied to interventions with a clear theory of change and attempt to measure implementation of program activities, for example whether a program was delivered as intended (Steckler and Linnan 2002). The SIB model doesn't have a clearly defined theory of change, so it is difficult to apply the traditional lens of the process evaluation to the SIB mechanism itself. In article one, I lay out a conceptual framework that attempts to clarify the mechanisms through which a SIB may impact service delivery. These mechanisms can be improving or impeding, similar in concept to barriers and facilitators of successful intervention implementation (Schneider et al. 2009). The case study attempts to measure how the SIB model was implemented throughout the project, and the impact that the different features of the SIB had on service delivery. Using interviews, document review, and participant observation, the analysis was designed to uncover how the SIB model altered processes of service delivery (i.e. through which mechanisms), what the effect of those alterations were, and for whom (Pawson and Tilley 1997; Steckler and Linnan 2002).

Semi-structured in-depth interviews were conducted with all available project stakeholders (n=31) aside from service users.[4] Interviewees included Essex County Council leadership and employees (n=11), SIB

[4] Service users were kept separate from the SIB element of the project, and so were interviewed in a separate study (Drew et al. 2019) that aimed to unpack their experiences with the MST service.

investors (n=3 investors, 1 advisor), intermediary leadership and staff (n=3), provider leadership and staff (n=10), and others involved with the project (n=3). Using a semi-structured approach (Bryman 2015), interviews were conducted flexibly, allowing informants to answer open-ended questions and shifting the emphasis of the conversation based on the informant's role and previous responses. To help facilitate an environment where respondents could openly reflect on their experience and opinions of the SIB project, interviews were conducted individually with each informant.

Interviews were supplemented by comprehensive documentary review on existing board papers and presentations, meeting minutes, contracts, evaluation papers, and project reports (n=229 documents, 2028 pages). The documents provided insight in to how the project was developed, from inception through completion, including evidence of project challenges and how they were dealt with, progress towards meeting outcome goals, data monitoring processes and subsequent actions taken, points of disagreement between various actors, and approaches to project governance. The stakeholder interviews and documents were coded thematically, with help from the NVivo 10 software, using an integrated approach (Weston et al. 2001) combining both inductive and deductive approaches by beginning with the initial themes set out in the conceptual framework but allowing space for the data to inform the theory (Bryman 2015).

To further supplement the data collected, participant observation was conducted through three days of shadowing front-line staff while the service was still being delivered. I was able to silently observe front-line teams conduct their daily activities. While this data collection strategy was agreed to be used strictly for background purposes, participant observation can help to provide a richer understanding of the project and enhance and inform further data collection and analysis (Musante (DeWalt) and DeWalt 2010). Participant observation is a useful tool to help the researcher understand context and also critically examine any assumptions of the research (Musante (DeWalt) and DeWalt 2010).

The qualitative methods conducted are subject to a number of limitations. Interviewee responses and participant behavior can be biased by the presence of the researcher (Bryman 2015). For example, participants may have felt the need to self-censor when discussing sensitive subjects (including relationship tensions, pressures to meet outcome goals, financial incentives, etc.). In addition, as interviewees were asked to reflect on past events, responses are subject to recall bias (relying on memories of events instead of events themselves) (Bryman 2015). Case study designs also have a number of disadvantages, as they are based on the subjective interpretation of the researcher, calling in to question replicability and objectivity (Berg and Lune 2012). Furthermore, conclusions from single case studies are often difficult to generalize outside of the specific context (Gerring 2004). Depending on the selected case there may be different relationships between the variables or phenomena being studied (Berg and Lune 2012). Efforts were taken to mitigate potential biases, for example by conducting interviews individually, offering all respondents to be anonymized, triangulating data from three different sources, and validating processes and findings with experts in the field.

Qualitative content analysis

The third and final article of the thesis relies on qualitative content analysis of all SIB projects launched in the US and UK from 2010-2020 (n =114). While a more detailed description of the methods is included in article three, the following discussion provides an overview of the content analysis conducted.

In this article, co-authored alongside Drs. Eleanor Carter and Mara Airoldi, we investigate the variation of SIB design across two quintessential dimensions originally promised by the SIB model: 1) payments are linked to impact, and 2) up-front capital is independent and at-risk. We plot the variation of SIB projects across these two dimensions to help determine how the SIB model differs from other commissioning approaches, such as fee-for-service or payment by results. This analysis represents the first systematic attempt to map variation in design of SIBs against the espoused model, and allows us to draw conclusions

on the model's ability to live up to its original promises. If SIBs have stretched too far from what makes it a distinctive, innovative approach this could undermine its potency as a tool for service delivery improvement (Carter 2020).

Qualitative content analysis is an appropriate method for which to investigate this research question as it allows for a systematic classification process of coding qualitative data in order to identify patterns (Hsieh and Shannon 2005). At the heart of the method is building a coding frame, which enabled us to focus on select aspects of SIB design relevant to our research question, rather than sifting through huge amounts of material available for each project (Schreier 2014). To create an appropriate coding frame to analyze variation in the design of SIB projects, we used a data-driven and iterative process to break down the two dimensions of interest (payment linked to impact and nature of the capital) in order to operationalize them. As part of this we tested the coding frame on a subset of the material, in a pilot phase, to ensure that the operationalizations captured elements of SIB project design that we aimed to analyze. As a result, we created a composite coding score for each dimension. For payment linked to impact, this included a composite score combining the SIB project's payment model and validation method. For nature of the capital, this included a composite score combining the SIB project's primary provider of upfront capital and the repayment structure. For definitions and more information, see the *Conceptual Framework* of article three (pg.129).

SIB projects in our sample included all those projects identified as a SIB by the International Network for Data on Impact and Government Outcomes (INDIGO) that were launched in the UK and the UK between 2010-2020 (n=114). These projects represent over 60% of all SIBs worldwide, providing a rich set of data points for analysis that allowed us to investigate longitudinal trends in how SIBs have 'stretched' over time within and across the two countries. The inclusion of both countries enables a comparative lens to see how SIBs have been designed across the US and UK. To populate the coding frame for these projects, we conducted content analysis of publicly available documents, including evaluation reports, press releases

and project announcements, project contracts, and government website pages. To validate these findings and confirm unclear project-level information, we called on the expertise of an advisory group to share project-specific data and validate our scoring (n=5).

This qualitative content analysis is subject to a few limitations. As with many types of qualitative analysis, the construction of a new coding frame raises concerns around objectivity. How we decided to construct the coding frame has implications for interpretation of the data and subsequent findings. We attempted to mitigate this concern by soliciting critique and validation of the coding frame by other experts in the field (n=11). In addition, qualitative content analysis is narrowing in nature, and as such might not provide the full context of variation in SIB design across countries or over time. For example, the coding frame only focuses on two dimensions of SIB design, although the authors argue that these dimensions are the most critical to the SIB's original purpose. Additionally, the research sample only includes projects launched in the US and the UK, creating a narrow focus in which the findings might not translate to other countries experimenting with the SIB model. Finally, for some of the SIB projects in the sample, public materials were insufficiently detailed to populate the coding frame's dimensions.

Significance

This thesis seeks to better understand the impact of the SIB model, comprising of a multi-sector partnership, data monitoring infrastructure, outcomes contract, and external investment, on different levels of service delivery. Investigating this impact contributes to the wider academic debate on New Public Management reforms, including on topics related to the use of financial incentives in public service delivery, accountability in hybrid governance structures, and varying levels of discretion to front-line staff.

Further exploring the implications of the SIB model on different levels of service delivery also has practical implications for policymakers and service providers. Over the past ten years, over 200 SIBs have been

launched around the world as practitioners continue to experiment with the model. Equipping government and third sector leaders with practical, contextual knowledge about the various positive and negative implications of the SIB model will aid them in making decisions on how to most effectively support vulnerable populations.

The SIB model is subject to heated debate, primarily because it involves private investors funding social services with the potential to generate returns for achieving outcomes. Despite this debate, and the model's increasing application to policy areas including homelessness, child welfare, and employment support, it is unclear how the model impacts processes of public service delivery. In this thesis, I offer three separate articles that aim to address a critical gap in our understanding of how SIBs work in practice and the impact they have on service delivery systems. Through these three articles, I make unique theoretical and empirical contributions to the debate on how social impact bonds affect processes of public service delivery.

Contents of Thesis

In alignment with the requirements for an article-based thesis, the contents of the thesis following this introduction include: a literature review chapter, three stand-alone articles, and a conclusion chapter.

In the literature review chapter, I further discuss the roots of the SIB model in New Public Management, which foreshadows much of the current debate around SIBs. Then, I lay out the current debate over SIBs by reviewing the literature on the 'SIB effect'. In synthetizing the literature, I develop a conceptual framework which identifies three distinct levels of analysis: (i) the system of public service delivery, (ii) the provider organizations and front-line-staff, and (iii) the service users. For the SIB effect on systems of public service delivery, I categorize the major promises and critiques of the model hypothesized by scholars in the field. For the SIB effect on providers and front-line staff, I discuss existing evidence and how it ties into wider literature on street-level bureaucrats. For the SIB effect on service users, I highlight the lack of

empirical evidence on how the SIB model affects service user outcomes. Finally, I conclude with a discussion of key gaps in the literature and situate the intended contribution of my dissertation.

In the first article, I offer a conceptual framework for how SIBs might impact public service delivery that stems from the SIB model's four key features and is based on a review of the literature. I then apply this framework in a case study of a SIB launched in Essex, using in-depth interviews, document review, and observation of front-line staff. I find that the SIB had mixed effects on public service delivery in Essex. The SIB model enabled rigorous monitoring of performance, availability of additional resources for the service, and outcome goal alignment across stakeholders. However, the SIB model also generated stakeholder conflict, feelings of resource burden and disproportionality by the government partner, loss of democratic accountability, and the pressuring presence of an underlying financial driver. I conclude with a discussion of policy implications, including the importance of safeguarding democratic accountability, developing positive relationships, and mitigating disproportionality (high levels of resources spent on a small cohort) if government agencies continue to pursue SIBs.

In the second article, I investigate whether using a prescribed and replicable (i.e. high-fidelity) intervention insulates providers and front-line staff from potential negative effects of the SIB model. SIBs often fund 'black-box' interventions, where the process of delivering a social intervention is not prescribed nor monitored and only the outcomes matter. This causes concerns amongst public servants and scholars around the potential of gaming outcomes contracts, for example by focusing on service users most likely to generate outcome payments at the expense of other, harder-to-reach service users. In this article, I draw on the in-depth case study of the Essex SIB to analyze how using a high-fidelity intervention with the SIB model affects providers and front-line staff. I find that while using a high-fidelity intervention with the SIB model mitigates incentives to game outcome payments, it also restricts the ability of the provider to innovate and deliver flexible, context-driven services. Using a high-fidelity intervention does not protect providers from

pressure to meet output targets, which could potentially create other disincentives in service delivery (for example, inappropriate referrals).

In the third and final article, co-authored alongside Drs. Eleanor Carter and Mara Airoldi[5], we investigate whether SIB projects have been designed to accomplish two of the model's primary original intentions: shifting the focus of public service delivery to achieving impact and transferring risk away from the government partner. By conducting qualitative content analysis on the SIBs launched to date in the US and UK, we plot and analyze how the design of SIB projects vary across two key dimensions: degree to which payment is linked to impact, and the independent, at-risk nature of the working capital provided. We find that SIB design in practice has deviated from the model's original intent in several important ways. In the UK, outcomes that projects pay for are generally not subject to rigorous validation methods, calling in to question whether payments are meaningfully linked to impact. In both the US and the UK, risk mitigation strategies in the repayment structures may limit the intended transfer of risk to external investors.

The conclusion chapter reflects back on the objectives of the thesis and ties together a discussion of the main findings across the three articles. I highlight the strengths and limitations of the research and discuss the contribution of the thesis to the wider academic literature. Finally, I suggest policy implications of the research findings and conclude the thesis with suggestions for future research.

[5] See Appendix 1 for co-authorship declaration and contribution statement.

Literature Review

Introduction

Recently, scholars have attempted to identify the 'SIB effect,' hypothesizing the potential benefits and problems that the model brings compared to alternative commissioning tools (e.g. payment by results, a fee-for-service contract, or in-house delivery). However, determining the 'SIB effect' is both a challenging and nuanced task. First, it is difficult to separate the effect of the SIB from the effects of the funded intervention. Second, the SIB can have effects at multiple levels of service delivery: the system-level, and organization-level, and an individual-level. As a result of this complexity, questions remain. By using the SIB model to fund social services, what is the impact on the system of service delivery? On provider organizations and front-line staff? And on service users themselves? This chapter aims to review the existing literature on the SIB effect, and highlight critical gaps in our understanding of how SIBs work in practice.

The literature review chapter proceeds as follows: First, I discuss the roots of the SIB model in New Public Management (NPM), which foreshadows much of the current debate around SIB projects. Then, I lay out the current debate over SIBs by reviewing the literature on the 'SIB effect'. In synthetizing the literature, I develop a conceptual framework which identifies three distinct levels of analysis: (i) the system of public service delivery, (ii) the provider organizations and front-line-staff, and (iii) the service users. For the SIB effect on systems of public service delivery, I categorize the major promises and critiques of the model hypothesized by scholars in the field. For the SIB effect on providers and front-line staff, I discuss existing evidence and how it ties into wider literature on street-level bureaucrats. For the SIB effect on service users, I highlight the lack of empirical evidence on how the SIB model affects service user outcomes. Finally, I conclude with a discussion of key gaps in the literature and situate the intended contribution of my thesis.

Roots in New Public Management

The idea of improving public service delivery performance by incorporating private-sector techniques, one of the primary rationales behind the SIB model, is not new. Following economic recessions in the late 70s and early 80s, the idea of applying private-sector mechanisms and involving the private and not-for-profit sectors to increase the efficiency of public services came to the forefront of public sector reform, opening the way to budget cuts, personnel management, performance auditing, and privatization in public services (Mongkol 2011). These strategies were first used in the UK under Prime Minister Margaret Thatcher and in some municipal U.S. governments, then in New Zealand and Australia, and were followed by most Organization for Economic Cooperation and Development (OECD) countries (Gruening 2001). Around the world, public sector services were undergoing similar reforms modeled from private-sector techniques. In 1991, Christopher Hood coined the term 'New Public Management' (NPM) to refer to this movement (Hood 1991).

NPM typically includes seven key features: 1) hands-on professional management, 2) explicit standards of performance, 3) emphasis on output controls, 4) disaggregation of public sector units, 5) increased competition, 6) private-sector styles of management, and 7) greater discipline in the use of resources (Hood 1991). In practice, NPM often meant more discretionary control to managers, resources that were linked to achievement of targets, use of shorter-term contracts, greater flexibility in hiring and the use of incentives, and cutting costs (Ibid). These reforms were meant to improve the quality of public services, while saving the government money by increasing efficiency (Balk 1996; Minogue 2001; Pollitt and Bouckaert 2004).

Described by Nagel (1997) as modeling the delivery of public services off of the economic market, NPM faced criticisms in the wake of ongoing ideological debates on the efficacy of markets for publicly provided services (Ferlie et al. 1996; Thynne 2003). Many believe that the adoption of private-sector practices poses a risk to public services because of the inherent differences between the two sectors (Pollitt 1990;

Armstrong 1998; Painter 1988; J. O'Flynn 2007). In particular, the public sector has "more complex objectives, more intricate accountabilities, and a more turbulent political environment than the private sector," making it difficult to meet multi-faceted public service goals using the ethos of the private sector (Mongkol 2011). There were worries that private sector management strategies, including the focus on measurable performance, would lead to corrupt practices or distort behavior (Barberis 1998; Doig 1997; Minogue 2001). There was also concern that these new methods were hurting public services, as they led to the ballooning of middle level management at the expense of front-line capacity (Nethercote 1989).

Despite the concerns with NPM, many elements of the movement are still seen in public service reform today. Justification of how SIBs could improve public service delivery revolved around concepts stemming from NPM ideals, including bringing private-sector actors and new forms of market discipline to the public sector. According to an early review of the literature on SIBs (Fraser et al. 2016), a consistent narrative used to justify SIBs closely mirrored NPM logic: that the private sector management techniques brought through the SIB model would be able to improve public sector service design, delivery, and accountability. Proponents of the SIB model argued that mechanisms of accountability and oversight would be introduced in ways not typical to public services because private sources of funding are tied to explicit measures of service provider performance (Burand 2013; Lagarde et al. 2013; Stoesz 2014). By only paying for specified outcomes, the SIB model could yield a greater discipline in the use of resources and help minimize inefficiencies characteristic of typical government spending (Bafford 2012; Edmiston and Nicholls 2018). Hands-on professional management, private-sector involvement, explicit performance standards, and greater discipline in use of resources are all features of the NPM movement (Hood 1991) that are echoed in the SIB model. Furthermore, the assumption that incorporating private sector actors and techniques would lead to improvements in public service delivery and efficiency stems firmly from NPM logic (Osborne 2006).

Despite strong ties to NPM logic, the SIB model may not fit neatly into all components of the NPM framework in practice. Three major goals of NPM have been characterized by the "three M's" - markets, managers, and measurement (Ferlie et al. 1996). While SIBs emphasize the role of management and call for an increase in measurement, it is not clear what role SIBs have played in fostering the market (or increasing market competition for public services). To the contrary, Neyland (2018) argues that SIBs have served as an anti-market device by ruling out competition and "guaranteeing a local monopoly" by fixing in place the intervention, the providers, and the investors. However, initial arguments from SIB proponents claimed that these projects could increase competition, by fostering private sector-led innovation and "encourag[ing] new market entrants" in the world of public services (Fox and Albertson 2011). A lack of competitive market goes against NPM ideals, but this may not have been an intentional design of the SIB model and rather an outcome of the context in which SIBs have grown (i.e. few social investors willing to invest or provider-led projects).

Some scholars have argued that perhaps New Public Governance (NPG) (Osborne 2006) may be a more appropriate lens with which to analyze SIBs, with its emphasis on the role of multiple stakeholders and recognition that social outcomes are "emergent properties of complex systems" (Fox and Morris 2019). NPG is a framework that highlights the reliance on preferred suppliers, rather than a competitive market, and emphasizes neo-corporatist values of co-designed interventions and collaborative, trust-based relationships (Osborne 2006). It is helpful to draw on the NPG framework when analysing the SIB's emphasis on collaboration amongst multiple stakeholders (Lowe et al. 2019), but for most components of SIB design and implementation the underpinning logic is more deeply rooted in NPM (Dayson, Fraser, and Lowe 2020). For example, private sector strategies of performance management and data monitoring and measurement are heavily emphasized in the SIB approach (Tan et al. 2015). SIBs purport to use market-based approaches to tackle challenging social problems that have typically been seen as government responsibility (Nicholls 2010), in part by importing "ideals, values and tools" from private sector management (Lowe et al. 2019). When SIBs were first being experimented with, one of the overarching

narratives in favor of the model was that private sector management techniques, explicit outcome measures, and more efficient use of resources (leading to cost-savings) would improve public service design and delivery (Fraser et al. 2016). Criticisms of the model, in addition, also largely stem from an NPM lens. Critics argue the SIB model is an extension of privatization of public agencies, and increases the risk of financializing social services with a focus on cost-reduction in an era of austerity (Dowling 2017; Tse and Warner 2018; Cooper, Graham, and Himick 2016). Scholars argue that this financializaton runs the risk of turning service users into potential sources of private sector revenue and risks losing democratic accountability over governance structures (Dowling 2017; Sinclair, McHugh, and Roy 2019; Tse and Warner 2018). These arguments, both for and against SIBs, strongly support the applicability of the NPM framework for analyzing SIBs. The next section provides an original synthesis of the debate around how SIBs might impact public service delivery, which has roots in the wider arguments for and against NPM strategies.

The 'SIB Effect': An ongoing debate

After ten years of experimentation, what has been the effect of using the SIB model on the delivery of public services? Unfortunately, this question largely remains unanswered, as determining the 'SIB effect' is a challenging and nuanced task. First, separating the SIB effect from the intervention effect is difficult. Most existing studies "tend to conflate the outcome based commissioning mechanism with the intervention or services that are funded," and therefore are unable to "assess the contribution of these separate elements" on processes of service delivery or outcomes for service users (Fox and Morris 2019). Second, the classification of 'SIB' is an umbrella term, with large variation in implementation along key features of the model. This heterogeneity, however, can be used for comparative analysis and unpack how the various features of the model can affect service delivery. Third, the SIB can have an effect on a number of different levels, as they are multifaceted policy tools that involve complex relationships and dynamics across various sectors (Lowe et al. 2019). The 'SIB effect,' therefore, can vary depending on what level and from whose

perspective it is examined. The following section describes learning from the literature on the potential 'SIB effect' on 1) the system of public service delivery, 2) the provider and front-line staff, and 3) the service user.

Methods and Conceptual Framework for Literature Review

In conducting the literature review, I started by reviewing the references included in Fraser et al.'s (2016) comprehensive literature review of social impact bonds, which drew on 38 academic papers and 63 grey literature sources. While analyzing these papers, I created a running list of the suggested effects of the SIB model and categorized them as either positive impacts (promises) or negative impacts (critiques) and at on what level (macro, meso, or micro) the effects were relevant. I then expanded my analysis by incorporating relevant papers published in 2016 or later, that included the term 'social impact bond' or 'pay for success' in the title or abstract.[6] I then sifted through these recent papers looking for additional claims about how SIBs impact service delivery, and incorporated them in to my running list. As a result, I have categorized my findings across three levels: macro (system), micro (service provider), and meso (service users). At the macro level, I include findings that reflect on the promises and pitfalls of the SIB model on the overall system of public service delivery. At the meso level, I include findings on how the SIB model affects provider organizations and front-line staff. This literature is supplemented by insights from the literature on street-level bureaucrats, which describes different policy conditions (relevant to SIBs) that can have negative or positive effects on those delivering a service. At the micro-level, I include findings on the SIB effect on the outcomes of service users. This original approach to synthesizing the evidence, summarized below in Table 4, helped me to identify relevant and interesting questions to investigate as part of this thesis.

[6] I utilized the following search databases: Google Scholar, JSTOR, ProQuest, Social Science Citation Index (Web of Science), and Social Science Research Network.

Table 4: Literature on SIB effect across levels

Level	Unit of analysis	Literature Reviewed
Macro	System of public service delivery	Journal articles, grey literature, and project evaluations reflecting on promises and pitfalls of the SIB model on public service delivery
Meso	Service provider and front-line staff	Journal articles, grey literature, and project evaluations reflecting on SIB effect on service provider Insights from street-level bureaucracy literature
Micro	Service users/intended beneficiaries	Journal articles, grey literature, and project evaluations reporting on SIB effect on service users

On the system of public service delivery

How might a SIB impact the system of public service delivery? The system of public service delivery is defined as the entire network of interacting organizations, individuals, structures, and processes that contribute to service delivery (Johnson, Fortune, and Bromley 2017). Many scholars studying SIBs have made theoretical or empirical claims regarding the potential promises and pitfalls of the SIB model on public service delivery. This section attempts to organize these claims, in order to better understand how a SIB might potentially improve or undermine public service delivery. As detailed below, I find that the key promises revolve around the improvement of public service delivery through five purported mechanisms: shifts the focus of public service delivery to achieving impact, transfers risk away from government, introduces market discipline, creates flexibility for innovation, and fosters cross-sector collaboration. These are counterbalanced by core criticisms that the SIB model undermines public service delivery through five potential mechanisms: high transaction costs, financializing of social services, creation of perverse incentives, narrow scope of services, and failure to foster innovation or transfer risk.

SIB promises: improving public service delivery

The existing literature lays out a wide range of *potential* advantages of using the SIB model. The section below lays out detail on the five major promises of SIBs from this literature:

1. Shifts the focus of public service delivery to achieving outcomes and impact
2. Transfers risks of unsuccessful programs away from government to external investors
3. Introduces market discipline, with increased focus on management and measurement
4. Creates flexibility for innovation
5. Fosters cross-sector collaboration on a pressing social issue

It is worth noting that are that not all of these promises are specifics to SIBs, as many overlap with outcome-based commissioning arrangements, of which SIBs are a special category.

1. Shifts focus of public services to achieving outcomes and impact

SIB projects are meant to shift the impetus of service delivery to measuring and achieving impact, as the government only pays for services delivered if specific outcome metrics are achieved (Fraser et al. 2016; Albertson et al. 2018; Chiapello and Knoll 2020). Typically, these outcomes are agreed by all parties (government, provider, investor) prior to the project's implementation. This process, facilitated by the SIB structure, can align outcome goals across participating parties and yield a consistent focus on achieving those goals (Liebman 2011; Fraser et al. 2016). Robust counterfactual-based evaluation methods can help to give confidence that SIB projects enable government agencies to pay only for meaningful improvements in the lives of intended beneficiaries (Liebman 2011; Cohen 2011; Corrigan 2011; Mulgan et al. 2011; Fox and Albertson 2012; Albertson et al. 2018; Fraser et al. 2018; Chiapello and Knoll 2020). Additionally, by "allocating resources to where they will achieve the most impact," the use of SIBs can help to achieve impact with greater efficiency than other forms of service provision (Fox and Albertson 2012).

SIBs could also foster an impact focus by altering the incentive structure for service providers. It is argued that the payment for performance component of the SIB incentivizes higher quality service provision, because the risk of underperformance is higher than with a traditional contract (Baliga 2013). Provider incentives are altered by the payment mechanism, shifting their behavior to focus on improving outcomes

(Battye 2015). Instead of focusing on inputs and outputs, common practice among providers, they instead focus on "quantifiable social outcomes that dictate the rate and amount of payment" they receive (Edmiston and Nicholls 2018). Discussed further in the flexibility promise, this outcome focus is meant to enable service providers with a flexibility to innovate that can improve service delivery (Leventhal 2012; Fox and Albertson 2012; Fraser et al. 2016). Ronicle and Smith (2020) found that one of "the strongest effects" of the SIB model was that "the SIB increased the focus on achieving outcomes," and enabled service providers to adapt in efforts to better achieve those outcomes.

2. Transfers risks of unsuccessful programs away from government and on to external investors

A common rationale for the use of SIBs is the transfer of risk from government and taxpayers to the private sector (Fox and Albertson 2011; Mulgan et al. 2011; Edmiston and Nicholls 2018; Fraser et al. 2018). By using SIBs, government commissioners only pay for the achievement of agreed-upon outcomes, and so in theory should be shielded from the risk of project failure (Liebman 2011; Loder 2011; Fox and Albertson 2012; Chiapello and Knoll 2020). In this way, the SIB model can be defined as "a contractual arrangement intended to transfer the financial risks of complex and innovative services away from government commissioners and direct service providers" (FitzGerald et al. 2019). If a service is unsuccessful, the external investors stand to lose their investment and the government commissioner is free to spend public dollars elsewhere. This risk transfer theoretically protects government agencies from the costs of failure, and thus "provide(s) space for risk-taking innovation" within social service delivery (Loder 2011). The degree to which this risk transfer has taken place in practice is debatable, and discussed more in the next section on SIB critiques.

Some scholars hypothesize that the transfer of risk can increase the accountability, effectiveness, and efficiency of public service provision (Fox and Albertson 2011; Fraser et al. 2016; Edmiston and Nicholls 2018). It is argued that accountability and value for money can be increased by paying only for specified outcomes rather than for services delivered (Edmiston and Nicholls 2018). In addition, when the private

sector holds the risk of project failure, this could "promote more market testing" and "encourage new market entrants" in the world of public services (Fox and Albertson 2011). Baliga (2013) adds that in many other types of privatization or outsourcing the government pays for services regardless of project success, and that SIBs offer a more beneficial transfer of risk to the public sector.

3. Introduces market discipline, with increased focus on management and measurement

Fraser et al. (2016) identified a common narrative in the SIB literature that private sector management techniques, such as those in the SIB mechanism, can improve public sector shortcomings in service design, delivery, and accountability (Mulgan et al. 2011; Liebman 2011). A number of "ideals, values and tools from private sector management have been imported" into the SIB model, by using revenue maximization techniques typical to private businesses in order to tackle social problems (Lowe et al. 2019). Proponents argue that the SIB model may lead to greater accountability and clearer expectations of the results government commissioners seek from service providers, and more transparency of government expenditure (Lagarde et al. 2013; Stoesz 2014; Schinckus 2017). Through the structure of a SIB, profits are tied to socially valuable results, so as to align public and private incentives (Cohen 2011; Liebman 2011; Nicholls and Murdock 2012; Baliga 2013; Liang, Mansberger, and Spieler 2014).

In particular, because private sources of funding are tied to service provider performance, mechanisms of accountability and oversight are introduced in ways not typical to public services (Burand, 2013). More active performance management of contracts, whereby stakeholders meet regularly to discuss progress, on-the-ground learning is incorporated in to next steps, and implementation challenges are identified and responded to quickly, is one such example of the type of oversight brought by this 'market discipline' (GPL 2017; Ronicle and Smith 2020). The enhanced use of data has been reported as one of the central advantages of the SIB structure, particularly using performance data to guide decision-making, review practice, and make program changes (Fraser et al. 2020; Disley et al. 2015; Griffiths, Thomas, and Pemberton 2016). Lowe and Wilson (2017) have found that in some cases, increased performance management and the use

of performance targets can lead to improved practice. These strategies can potentially help to minimize inefficiencies found in typical government spending (Bafford 2012).

4. Creates flexibility for innovation

Another common promise of social impact bonds is that they provide an opportunity to innovate (Griffiths and Meinicke 2014; Fraser et al. 2016; Carter et al. 2018). It is argued that the outcome focus, rather than a process focus, allows service providers flexibility for program adaptation and leads to greater personalization of services (Leventhal 2012; Fraser et al. 2016; Ronicle and Smith 2020). It can allow providers greater latitude to determine how to obtain results, which some argue leads to improvement in service delivery (Fox and Albertson 2012; Maier, Barbetta, and Godina 2018). Liang et al. (2014) argue that it is the dual mechanisms of 1) the investor's aim for higher returns and the 2) provider's requirement to deliver effectively, which can drive innovation. Arena et al. (2016) pose a different argument, that it's the SIB's "new network of relationships" that yields flexibility for innovation:

> *"SIBs can foster innovation by leaving service providers free to design new initiatives to achieve the expected social outcome, by exploiting synergies between different actors, by reconfiguring the structure of service delivery, by creating opportunities of cross-fertilization between different sectors of the public administration."*

Apart from innovation and flexibility for service delivery, others argue that SIBs can lead to innovation in the selection and implementation of new programs. Butler et al. (2013) suggest the potential for SIBs to test new and innovative programs that have "a strong theoretical basis" but have not been demonstrably proven, which could help to increase the amount of "truly effective programs." Typically, government budget constraints can serve as a disincentive to pilot new programs that have not been proven. SIBs have been argued as a mechanism to fund these "innovative and promising" interventions, that can result in the "adoption of new solutions and strategies to solve problems" (Liang, Mansberger, and Spieler 2014). The extent to which SIBs have, in practice, been used to fund new and innovative programs is debatable. Maier,

Barbetta, and Godina highlight this as one of the major paradoxes of the model: the promise of SIBs to be both evidence-based and flexible (2018). Arguments that the SIB model have not lived up to this promise of flexibility will be discussed more in the next section on SIB critiques.

5. Fosters cross-sector collaboration on a pressing social issue

The SIB structure usually requires long-term contracts that involve a variety of stakeholders. As a result, they have the potential to provide a "compelling focal point" for collaboration amongst "governments, donors, private sector investors, and non-governmental social service providers that are interested in scaling social service prevention interventions for a sustained period of time" (Burand 2013). This multi-sector partnership can foster collaborative partnerships, leading to joint-working and the potential for new ideas and heightened attention to challenging social problems (Liebman 2011; Nicholls and Murdock 2012; Liang, Mansberger, and Spieler 2014; Arena et al. 2016). Maier, Barbetta, and Godina (2018) argue that the multi-stakeholder collaboration characteristic of SIB projects has the potential for improving social policies by bolstering knowledge flows across sectors.

Baliga (2013) argues that the structure of the SIB "acknowledges that the largest social ills of our day cannot be confronted by the government, the private sector, or philanthropic organizations acting alone," and brings these actors together to tackle complex challenges. Others echo this argument, suggesting that the SIB structure allows for aligned incentives between the public, private, and voluntary sectors working towards a common goal (Cohen 2011; Liebman 2011; Arena et al. 2016). Scholars and practitioners arguing that SIBs enable collaboration often remark on the beneficial importation of private sector techniques, discussed previously in this section.

Other promises common in the early days of SIBs

There are a few other promises of the SIB model worth mentioning here for more context on the model's original intentions. These are not included in the list of 'five major promises' because, while popular

arguments in the early days of SIB development, they have since lost traction. These include the SIBs potential to: realize cashable savings, shift spending from remediation to prevention, and scale up successful interventions.

- *Realize cashable savings:* When SIBs were first proposed, it was widely argued that the projects' achieved outcomes would result in savings to the public sector (Fraser et al. 2016). For example, one of the first SIBs in the US provided supportive housing to high-frequency users of the hospital, police, and emergency systems with the intention that the project would help individuals reduce their reliance on these systems and thus save government money (Urban Institute 2019). In theory, payable outcomes in SIB projects represent avoided downstream social problems, and if outcomes were achieved then the state would be paying for them out of the savings generated by not needing to fund as many expensive remedial services (Fraser et al. 2016; Social Finance 2011; Mulgan et al. 2011; Albertson et al. 2018). It is debatable whether cashable savings have been realized in practice, despite many projects achieving outcomes, so this promise has stopped being cited as an argument for SIBs.
- *Shift spending from remediation to prevention:* Related to realizing cashable savings were arguments that SIBs would shift spending from remediation to prevention. Under traditional budget constraints, governments are often incentivized to fund remedial services for more immediate needs rather than invest in preventative programs. For example, in the UK £92b is spent on health expenditure, of which only 3.7% is spent on preventative interventions (Albertson et al. 2018). A further £1.5b is spent on school truancy and exclusion, while only £111m is spent on preventative initiatives (Ibid). SIBs have been lauded by some as a potential solution to this issue, by paying for preventative programs only after potential remedial savings are realized (Liebman 2011; Loder 2011; Bafford 2012). SIBs were argued to have the "potential to shift the government funding cycle", by "creating a system centered on prevention and measurable outcomes instead of one focused on remedial interventions and inputs" (Bafford 2012). While SIB projects do frequently focus on the prevention of challenging social programs (Fraser et al. 2016), the argument that the model shifts spending away from remediation has lost traction.

- *Scale up successful interventions:* Another early SIB promise was that, by relying on private investment to test social interventions, SIBs could encourage successful models to be scaled up and new models to be experimented with (Loder 2011; Liang, Mansberger, and Spieler 2014). There are a few ways in which SIBs have been lauded as a new way to scale up evidence-based interventions. With upfront working capital from private or philanthropic funders, SIBs could bring together and scale up multiple service providers to collectively work towards achieving the same outcomes (Burand 2013). The mechanism could also provide "seed capital" to cultivate innovative programs that may have not otherwise been government funded (Baliga 2013). Additionally, the model could bring new sources of funding from private investors who have not previously invested in similar interventions or social issue areas (Disley et al. 2011). Through these means, SIBs could have provided a "new way to harness private capital to support the scaling of government contracted social service interventions" (Burand 2013). This argument has lost traction, as in practice SIBs have not been used to scale up successful interventions.

SIB critiques: undermining public service delivery

In response to all of the promises of social impact bonds laid out in the existing literature, there are a number of critiques related to how the model could undermine public service delivery, including:

1. High transaction and ongoing management costs
2. Financializes social services and weakens democratic accountability
3. Encourages perverse incentives
4. Creates a narrow scope of services, targeting symptoms instead of the system
5. Fails on key promises to foster innovation and transfer risk

As with the promises, not all of these criticisms are specific to the SIB mechanism, as some overlap with outcome-based commissioning more generally.

1. High transaction and ongoing management costs

For most commissioners, SIBs are a completely new way of contracting that often involves a multi-stakeholder partnership between government, service provider, investor, and intermediary. This process can contribute to high transaction costs, stemming from legal and procurement expenses, hiring an intermediary, and conducting any feasibility analysis (Boyle 2011; Disley et al. 2015; Demel 2012; Edmiston and Nicholls 2018). Lessons from the first SIB, launched in Peterborough, suggest that setting up robust outcome measures was "a time-consuming and analytically complex process" (Disley et al. 2011). As more projects have been launched, SIBs have developed a reputation as "notorious for their high set-up and management costs;" for example with one UK SIB reporting start-up costs of £200,000 (FitzGerald et al. 2019). Fox and Albertson (2012) raise the further issue of evaluation costs, noting that rigorous (and often costly) evaluations are key to the SIB model.

In addition to high start-up costs, scholars and practitioners have reflected on the ongoing resource intensity required throughout SIB projects, as partners are often required to frequently collect, monitor, and share data (Disley et al. 2015; Lowe et al. 2019; Ronicle and Smith 2020; Fraser et al. 2020). These performance management expectations can create pressure on both staff time and resources, particularly for provider and government partners (Fox and Albertson 2012; Tan et al. 2015; Roy, McHugh, and Sinclair 2018; Ronicle and Smith 2020). Lowe and Wilson (2017) have found that ongoing performance management responsibilities can take staff time away from supporting clients and encourage perverse behaviors related to meeting data requirements or performance targets. Pressures on data collection can be particularly draining for provider and government partners if the monitoring system focuses on processes of service delivery rather than service user outcomes directly (Fraser et al. 2020). Ronicle and Smith (2020) found that inefficient systems of "duplicative performance management" diminished an effective use of resources for the Youth Engagement Fund SIBs; while Edmiston and Nicholls (2018) highlight the administrative burden placed on service providers created by the intense management costs and requirements of SIB projects.

High start-up and ongoing management costs raise concerns about the cost-effectiveness of the SIB mechanism, as commissioners must consider whether using a SIB is worth the time and resource intensity. This is especially relevant given the arguments that SIBs are limited in scope and size, discussed more thoroughly in critique number four.

2. Financializes social services and weakens democratic accountability

There is a concern in the SIB literature that the introduction of private values to fund social interventions has dangerous implications for policy making (Fraser et al. 2016). Some scholars see SIBs as a tool that is allowing private investors to profit from populations in need of support; there are concerns that the private-sector element of the SIB encourages the financialization of vulnerable segments of the population, by allowing investors to shape policy and turn service users in to potential sources of revenue (Dowling 2017; Sinclair, McHugh, and Roy 2019; Tse and Warner 2018; Neyland 2018). Lake (2015) argues that SIBs encourage the narrative that public policy exists to support the economy at the expense of meeting social needs. Others suggest that the SIB model exploits the voluntary sector to achieve profit for investors (Fox and Albertson 2012; Dowling and Harvie 2014; Fraser et al. 2016). Warner (2013) and Rosenman (2014) add that SIBs have created a new profit-making industry, made up of private entities, who benefit at the expense of the voluntary sector and service users.

Others argue that SIBs mark a broader challenge to the ethos of the voluntary sector, and suggest that the profit motive may compromise social impact achieved (McHugh et al. 2013). It is possible that the use of SIBs could "alter the moral character" of service provision while changing the "nature of citizenship rights" (Sinclair et al. 2014). Roy, McHugh and Sinclair (2018) suggest that SIBs "transform citizens into commodities" and reduce their rights "both as service users and as a polity." Cooper, Graham, and Himick (2016), in an evaluation of a SIB providing homelessness support in London, argue that the model contributes to the securitization of homeless individuals into potential cash flow for investors. These

arguments reflect concerns that SIBs treat service users as agents to be acted upon, for the benefit of a profit-motive, rather than active participants who have a say in the services they receive.

In addition, there are fears that the privatization of social policy will have negative implications for democratic accountability (Joy and Shields 2013; Warner 2015). The loss of a direct relationship between service providers and government commissioners "erode(s) the trail of public accountability" (McHugh et al. 2013). Providers could lose independence, and service user interests may be sidelined (McHugh et al. 2013; Joy and Shields 2013; Sinclair, McHugh, and Roy 2019; Tse and Warner 2018). A loss of democratic accountability could be demonstrated if social service program decisions are dictated by a returns-motive instead of by the public interest, potentially at the expense of the service user (Joy and Shields 2013; McHugh et al. 2013; Warner 2013; Dowling and Harvie 2014; Fraser et al. 2016; Sinclair, McHugh, and Roy 2019). Morley (2019) argues for the need for more transparency and regulation with the SIB model to ensure that the preferences and voices of service users are prioritized.

3. Encourages perverse incentives

One of the most common critiques on the use of SIBs is the scope for perverse incentives, as payment is only made if particular outcomes are demonstrated (Considine, Lewis, and O'Sullivan 2011; Finn 2011; Rees, Whitworth, and Carter 2014; Lowe and Wilson 2017; Edmiston and Nicholls 2018). Edmiston and Nicholls (2018) argue that perverse incentives resulting from the "marketization of social outcomes" can compromise "service quality and integrity". Stout (2014) adds that pay for performance programs in general can "promote selfishness and opportunism" in the voluntary sector, resulting in possible "uncooperative, unethical, and illegal employee behavior." Payments based on outcomes can lead to ticking the box of the outcome being measured and disregarding quality of service delivery (McHugh et al. 2013). Even if there are no changes to actual service delivery behavior, the use of SIBs risk "encouraging service providers to manipulate how their performance is measured and reported" (Sinclair et al. 2014).

Perverse incentives could also be demonstrated through explicit gaming of the outcomes contract, for example by focusing on easier to serve clients who are seen as more likely to achieve the payable outcome (Joy and Shields 2013; Sinclair et al. 2014; Carter and Whitworth 2015; Fraser et al. 2018; Ronicle and Smith 2020). Cherry-picking, serving the easiest segments of the target population, is one such mechanism for gaming outcome payments (Fox and Albertson 2012). 'Creaming and parking' is another: taking a list of all the referrals and parking (i.e. ignoring) harder to serve clients and creaming (i.e. serving) easier to serve clients (Carter and Whitworth 2015; Roy, McHugh, and Sinclair 2018). Concerns over outcome gaming stem from long-standing claims that a focus on measurable performance in delivery of public services could lead to corrupt practices or distort provider behavior (Barberis 1998; Doig 1997; Minogue 2001). Scholars studying SIBs have found some evidence that the SIB model yielded distorted provider behavior, including by focusing staff time on clients who are easier to help (Ronicle and Smith 2020; Lowe et al. 2019; Edmiston and Nicholls 2018; Fraser et al. 2018; R. Griffiths, Thomas, and Pemberton 2016; Joy and Shields 2013).

4. Causes narrow scope of services, targeting the symptoms instead of the system

Relatedly, SIB projects can create a narrow scope of services by focusing only on interventions that can be funded by private investors, outcomes that can be measured, or target populations that can be easily reached (Fox and Albertson 2012; Dowling 2017; Tse and Warner 2018; Ronicle and Smith 2020). Dowling (2017) highlights the inherent narrow focus of SIBs, as they tend to target individuals at the symptom-level rather than "transforming the social, political and economic structures that create such social problems in the first place." Tse and Warner (2018) warn that SIBs can "silence boarder policy goals," by narrowing service provision to "low-cost programs with short-term returns, when more comprehensive approaches are needed." Only programs that are attractive to investors, and by extension unattractive to government commissioners, are likely to receive funding through the SIB mechanism (Giacomantonio 2017). SIB projects also risk supporting service users to achieve measurable outcomes without attention to the

underlying or wider factors that contribute to ongoing social issues (Lake 2015; Dowling 2017; Ronicle and Smith 2020).

The narrow focus on achieving measurable outcomes can lead to further obstacles, as outcomes are "notoriously difficult to measure" (McHugh et al. 2013) and many challenges arise when attempting to define outcome measures (Fox and Albertson 2012; Warner 2013; Liang, Mansberger, and Spieler 2014; Lowe and Wilson 2017). The simplistic model of defining measurable outcomes and then paying on them fails to account for the complexity or context of challenging social problems (McHugh et al. 2013; Tse and Warner 2018). Frequently, SIB payment mechanisms face attribution challenges – where it is difficult to determine whether or not an intervention was actually responsible for an outcome (Loder 2011; Sinclair et al. 2014; Roy, McHugh, and Sinclair 2018). Jackson (2013) argues that the innovations promised by SIBs have "deflected attention away" from the challenges with attaining ambitious outcome targets, as many factors that could influence outcome achievement are outside the control of stakeholders involved with a SIB. These challenges could lead to payment for narrow, less meaningful outcomes that do not correspond to a project's actual impact and neglect wider-spread, systemic issues (Tse and Warner 2018).

Paired with the intense time and resources required from stakeholders across sectors, SIB projects' narrow focus highlights the risk of disproportionality. Previous studies suggest that SIBs may bring additional layers of bureaucracy (L. Roberts and Cameron 2014; R. Griffiths, Thomas, and Pemberton 2016) more robust systems of program management (Fraser et al. 2018), and high transaction costs (Tan et al. 2015). These high costs, in terms of monetary resources as well as government and service provider time, spent on a relatively narrow intervention, raise questions about the equity of SIB projects. Disproportionality can increase the opportunity cost of doing SIB projects; the more time and resources are spent on a relatively small cohort, the less are available for other populations or the wider system.

5. Fails on key promises: does not create flexibility or transfer risk away from government

Many scholars have argued that the SIB model has failed to follow-through on two of its key promises: creating flexibility for innovation and transferring risk away from government agencies. In terms of flexibility at the provider-level, it is argued that by increasing the administrative burden on providers and front-line staff they actually have less scope for flexibility and less time working with the populations they seek to serve (Roy, McHugh, and Sinclair 2018). At the government-level, there are arguments that SIBs create a lack of flexibility by locking public agencies into multi-year outcomes contracts without discretion to change course (Neyland 2018). At the program level, Jackson (2013) argues that SIBs are not experimenting with new models and are only being used to supplement core government finances. Warner (2013) suggests that "only proven intervention models receive investment" because investors are not willing to take on model risks, leading to a lack of process innovation. Roy, McHugh, and Sinclair (2018) echo this sentiment, arguing that "almost all SIB-funded projects are based on well-established models" because investors motivated by a return have "little incentive to fund risky innovative policy experiments."

Scholars have argued that in practice SIBs do not actually transfer risk from government to investors (Demel 2012; McKay 2013; Fraser et al. 2016). This is due to the common practice of building in interim payments on output, and to the potential scope for cream-skimming to maximize outcome payments without actually achieving impact. Warner (2013) takes this argument one step further, arguing that SIBs do not live up to the promise of attracting private capital at all. In the UK, most of the SIB investors are social investors, rather than commercial, who have a dual mission to generate both social and financial returns (INDIGO 2021). There is more commercial financing in the US SIB projects, but even those are often underwritten with philanthropic capital (Ibid). Fox and Albertson (2012) add that investors seeking a "decent return" may find problems with the SIB model, as project sizes and thus overall changes in outcomes are likely to be small.

On providers and front-line staff

Many of the arguments around how SIBs might influence public service delivery or intervention outcomes rely on behavior and/or attitude changes among front-line staff in the service provider organization. The role of incentives, impacts on service quality, the degree of flexibility, and the shifting focus to outcomes are all components that play out to some degree on the front-line.

In analyzing the potential impact of the SIB model on front-line staff, it is helpful to draw on the wider literature of street-level bureaucrats (SLBs). In his seminal work, Michael Lipsky (1980) coined the term 'street-level bureaucrat' to refer to public service workers who directly interact with citizens and have wide discretion over services delivered, such as police officers, teachers, or social workers. His primary assertion is that the front-line decisions, routines, and coping mechanisms "effectively *become* the public policies they carry out" (Ibid). Referred to as "the founding father of the bottom up perspective" (Hill and Hupe 2009), Lipsky set the stage for attempting to understand the realities of policy as they translated into practice. His work, and the work of subsequent scholars, alludes to the vital importance of understanding how front-line discretionary behavior is shaped, and the influences of central-level policies (such as using a SIB mechanism) on front-line behavior.

So, what can the literature from street-level bureaucrats tell us about the impact of the SIB model on providers and front-line staff? Street-level bureaucrat literature suggests that front-line behavior is "is patterned by the structural conditions of the working environment," rather than "random acts of self-interested individuals" (Gilson 2015). In particular, structural conditions such as resource constraints, workload pressures, ambiguous policies, relations with clients, and bureaucratic efforts to exert control can have influences on front-line behavior (Lipsky 1980). Positive supports to front-line staff that are potentially reiterated in the SIB model include increased discretion, opportunities for collaboration, and feedback loops focused on learning by doing (Lipsky 1980; Brodkin 2012; Rowe 2012; Rothstein and Stolle 2008; Gilson 2015). Detriments to front-line staff that are potentially reiterated in the SIB model include an increased

administrative burden, the use of incentives, and complex systems of managerial oversight (Lipsky 1980; Marinetto 2011; Evans 2011; Brodkin 2012). These structural conditions are discussed more thoroughly in the paragraphs below.

Some front-line discretion is inevitable in program delivery, and high levels of discretion can help to advance important social policy goals while also allowing front-line staff to remain responsive to service user concerns (Brodkin 2012; Gilson 2015). The SIB model may afford providers and front-line staff more discretion over service delivery, through increased flexibility in interventions and processes they use. As resources may not be tied to the same regulations as public funds, and emphasis is on outcomes over processes, front-line staff may have more discretion to exercise their judgement and greater latitude in determining how to obtain results (Fox and Albertson 2012; Griffiths and Meinicke 2014; Carter et al. 2018). Providers may also have more discretion on resource allocation and prioritization, in addition to flexibility with direct service provision (Fox and Albertson 2012). In ways previously constrained, providers and front-line staff may be able to tailor services and incorporate innovative approaches in their interactions with clients (Fraser et al. 2018). Some scholars argue that the SIB model has the potential to enable service providers to "reconfigure[e] the structure of service delivery" and leave them free to design new initiatives or approaches through which to achieve outcomes (Arena et al. 2016). This could free front-line staff to personalize services according to the specific needs of services users or the particular context in which they are operating (Leventhal 2012; Jackson 2013). Some providers have reported that the SIB structure, particularly the focus on outcomes paired with the input of investors, has "motivated them to adapt" their processes of service delivery to better achieve outcomes (Ronicle and Smith 2020).

The benefit of building in time for reflection to review practice and learn from experience in order to support front-line staff is one lesson from the street-level bureaucrat literature (Lipsky 1980; Rowe 2012). This function could manifest in the SIB model through the increased performance management requirements, which are generally more intensive compared with traditional methods of service delivery (Fraser et al.

2018). In the SIB model, at least in principle, outcomes are rigorously tracked and monitored, often times with regular meetings involving the provider to troubleshoot challenges and brainstorm service changes. Some scholars have found that this performance management system can lead providers to increase their use of program data and have an aligned focus on outcomes, and in some cases may lead to improved practice (Edmiston and Nicholls 2018; Lowe and Wilson 2017). Providers have reported that structure has encouraged the use of data to review practice, and served as one of the main benefits of the SIB model (Disley et al. 2015).

Evans (2011) highlights the importance of positive relationships and collaboration between front-line workers and managers, particularly when managers provide professional guidance and support. May and Winter (2009) suggest that front-line understanding and acceptance of policy goals are one of the most important factors influencing their behavior with citizens. Relatedly, scholars studying SIBs have found that the support networks available to providers and front-line staff operating under the SIB model can be different from typical government contracts. First, the collaborative nature of the model has the potential to improve the flow of knowledge to and from the provider to other project partners, in order to improve service delivery (Maier, Barbetta, and Godina 2018). In addition, SIB contracts can last much longer and are typically locked in for a set number of years, which can provide front-line staff a higher degree of job security (Fraser et al. 2018).

The SIB model may also create structural conditions that have a negative impact on front-line behavior. Brodkin (2012) argues that interventions such as target-setting and incentives can influence the scope of front-line discretion, but cannot control it, and cautions against the use of incentives to influence front-line behavior. Many scholars studying SIBs have reflected that the outcome-based contracts could create disincentives to providers and front-line staff, because payment is only made if particular outcomes are demonstrated (Rees, Whitworth, and Carter 2014). This could create pressure to meet outcomes or outputs, which may manifest in a change in provider behavior (Fraser et al. 2018). Most explicitly, the SIB model

could encourage providers to game the contract, by focusing on easier to serve clients at the expense of cases that may be deemed as more challenging (Joy and Shields 2013; Sinclair et al. 2014; Carter and Whitworth 2015; Fraser et al. 2018; Ronicle and Smith 2020). For example, Lowe et al. (2019) suggest that for providers in one SIB project, the process of meeting payable outcomes "distorted practice" and "encouraged creaming of clients - working with those who are easiest to help."

The use of incentives could manifest in more subtle changes as well, including through a narrow scope of services provided, focusing only on achieving payable outcomes while neglecting more holistic, but difficult to measure, indicators of impact (McHugh et al. 2013; Tse and Warner 2018). The overall target population might also be narrowed, to focus only on those for whom data can be reliably collected, or those who are more easily reached (Lowe et al. 2019). For example, a SIB launched in North England to address social determinants of health only aimed to serve a subset of the population that General Practitioners thought would benefit due to provider restrictions created by the SIB structure (Lowe et al. 2019). In the Youth Engagement Fund SIBs, launched across the UK to support disadvantaged young people, evaluators found that providers did not necessarily support the most disadvantaged young people in target areas (Ronicle and Smith 2020).

Brodkin (2012) suggests that under time and resource constraints, it is sometimes very difficult for front-line staff to remain responsive to individual client needs. The SIB model may yield greater time and resource constraints, as front-line staff may feel an administrative burden by the increased performance management expectations typical of the SIB model (Disley et al. 2011; Roy, McHugh, and Sinclair 2018; Ronicle and Smith 2020). The time-intensiveness of data collection and reporting may detract from front-line staff's face time with service users (Edmiston and Nicholls 2018; Lowe and Wilson 2017). For example, Lowe et al. (2019) found that the "onerous data driven nature" of the SIB reduced the time that front line staff spent developing relationships with their clients. Furthermore, increased data management and monitoring requirements can exacerbate incentives for perverse provider behavior, as it may manifest

in pressure to meet payable outcome goals or output requirements (Fraser et al. 2018; Lowe and Wilson 2017; Arena et al. 2016)

Marinetto (2011) shows, as Lipsky (1980) hypothesizes, that managerial interventions can have negative effects on service quality and efficacy by reducing responsiveness to client need. In the SIB model, managerial interventions are increased and providers and front-line staff face additional lines of accountability - to the service user client, to their direct supervisors, to investors, and to the government partner (McHugh et al. 2013). Misalignment over goals or service delivery processes could lead to inter-organizational conflict (Warner 2015; Fraser et al. 2020; Ronicle and Smith 2020), and providers may express resistance to efforts to shift their organizational culture "to become more business-like" in response to the SIB model (Lowe et al. 2019). Additional project partners, including the presence of a private investor who may not have relevant programmatic experience, could lead to front-line frustrations with inconsistent advice on service direction, particularly as power imbalances are likely in provider relationships with investors (Morley 2019).

SIBs create structural conditions that could have an effect on providers and front-line staff. This effect could be positive by providing more front-line discretion and flexibility, increasing commitment to learning from data and creating feedback loops, or fostering support networks and collaboration with other sectors. Effects could also be negative, by encouraging and introducing perverse incentives, creating an administrative burden yielding stricter time and resource constraints, or creating more lines of accountability that might reduce front-line responsiveness to service user needs.

On service users

One of the most important yet understudied aspects of the SIB effect is on the outcomes for service users. Separating the SIB effect from the intervention effect is difficult, particularly when analyzing outcomes for

service users, and most existing studies tend to conflate the two (Fox and Morris 2019). The claims that the SIB model's effects on service delivery and front-line behavior, such as introducing market discipline to increase the focus on outcomes, giving service providers more discretion, or concentrating time and resources on a particular issue, are theoretically meant to translate to improvements in outcomes for service users. Most of the SIB projects in the UK are not analyzing the effect of SIB-funded interventions on service user outcomes, while projects in the US typically include impact evaluations to determine the SIB-funded intervention effects on service user outcomes (INDIGO 2021). No studies so far, however, have sought to explicitly determine the effect of the SIB component on outcomes for service users (for, by example, comparing intervention outcomes from a non-SIB funded project to a SIB-funded project).

The most developed evidence base for SIBs is in the UK, where there have been over 85 SIB projects launched to date and over $95m in upfront capital raised (INDIGO 2021). In a review of evaluations of UK SIB and PbR projects, Fox and Morris (2019) find 30 evaluation papers related to PbR programs and only 15 related to SIB projects (plus one related to both a PbR and SIB project). Of these, only four papers have been published in peer-reviewed journals. Published quantitative evaluations span the first launched SIB, Her Majesty's Prison (HMP) Peterborough SIB project, the Greater London Authority (GLA) Rough Sleepers SIB (Fox and Morris 2019), and the 10 Innovation Fund SIB projects. Evidence from the Peterborough SIB suggests that the intervention resulted in a 9% reduction in the reoffending rate, compared to a relevant counterfactual (Disley et al. 2011). The evaluation of the GLA Rough Sleepers SIB suggests that the intervention significantly reduced rough sleeping over a two-year period for participants, compared to a 'well-matched comparison group' (DCLG 2017). The evaluation of the Innovation Fund SIBs suggest that the projects were not successful in helping young people engage in education or employment compared to the comparison group: for participants over 16 years old, after one year 59% of the SIB group were in education, employment or training compared to 75% of the matched comparison group (Salis, Wishart, and McKay 2018). These evaluations, which point to mixed findings on the effectiveness of SIB-funded interventions, are helpful in that they seek to determine the impact of the SIB-

funded intervention on service user outcomes compared to relevant comparison groups. However, these studies focus on the intervention impact without being able to draw any conclusions on the impact that the SIB model may have had on service user outcomes.

In their review, Fox and Morris (2019) conclude that there is a paucity of evaluation on UK outcomes-based commissioning projects, and furthermore that the quality of existing evaluations is not high. The authors call for the use of randomized control trials (RCTs), in particular to separate the effects of the SIB model and the interventions commissioned (Ibid). In particular, there are limited studies that seek to determine the 'value-add' (if any) of using a SIB to fund particular interventions. The authors also comment that this finding confirms other evidence reviews that suggest that evidence has not been prioritized within the wider field of social investment (P. O'Flynn and Barnett 2017).

Second to the UK, the US has launched the next highest number of SIBs with over 25 projects to date (INDIGO 2021). Compared to the UK, the US SIB projects rely more heavily on impact evaluations to trigger payments rather than relying on administrative data (Albertson et al. 2018). As a result, there are various completed and ongoing impact evaluations of SIB projects in the US. The first ever US SIB, tackling juvenile justice in New York, used a quasi-experimental impact evaluation to determine payments. The evaluation found no significant impact of the program, and so no outcome payments were made (Vera 2015). While this is the only completed SIB in the US, there are a thirteen ongoing RCTs and two ongoing quasi-experimental evaluations that should be coming to an end in the next few years (Albertson et al. 2018). This will help to determine the impact of SIB-funded interventions on service users, but these evaluations do not test what the impact of the SIB model has been.

Existing publications are largely inconclusive on both the impact of the intervention and the effect of the SIB financing structure on outcomes for service users. There are other potential effects on services users aside from outcomes achieved, including more (or less) responsive or flexible services, enhanced (or

worsened) relationships with providers, or more (or less) discretion over the service direction. However, there is a notable lack of literature related to the effect of the SIB model on service users. Going forward, more direct empirical analysis of SIBs is needed (Tan et al. 2019; Fox and Morris 2019; Fraser et al. 2020).

Intended Contribution to the Literature

So, what exactly is the SIB effect? Social impact bonds have been around for over ten years, and, as laid out above, there is much debate around their impact on public services, provider organizations, and service users. What have we actually learned so far?

In terms of the impact of SIBs on systems of public service delivery, scholars hypothesize that the SIB model can have either improving or undermining effects. The SIB model could introduce market discipline with greater focus on performance management, create flexibility for innovation, transfer risk away from government, foster productive cross-sector collaboration, or shift the focus of public service delivery to achieving impact. Conversely, the SIB model could also put financial pressure on the system through high transaction and ongoing management costs, financialize social services and weaken democratic accountability, encourage perverse incentives, create a narrow scope of services, or fail on its key promises to foster innovation and transfer risk.

Similarly, for the impact of the model on providers; the SIB model could improve front-line discretion and flexibility, increase practices of learning from data to improve service delivery, or foster support networks and collaboration with other sectors. However, the model could also introduce perverse incentives to distort provider behavior, create an administrative burden yielding stricter time and resource constraints, or reduce front-line responsiveness to service user needs.

Finally, for service users the SIB model could have a (positive or negative) effect on intervention outcomes or the service user's experience with the funded intervention. Empirical evidence for any of these effects, however, is weak.

Table 5 below connects back to the original conceptual framework for the literature review, and offers a summary of my reflections across the different levels of the 'SIB effect'.

Table 5: Reflections on SIB literature across levels

Level	Unit of analysis	Reflections
Macro	System of public service delivery	• Polarization of claims - mechanisms of SIB model claimed to both improve and impede service delivery • Limited empirical evidence
Meso	Service provider and front-line staff	• Different structural conditions of SIB can have positive or negative impacts on front-line staff • Limited empirical evidence
Micro	Service users/intended beneficiaries	• Mixed findings of intervention effects on service users ○ No findings that isolate SIB structure • Lack of empirical evidence

Despite the many claims made around the promises and pitfalls of SIBs, and the growing number of SIB projects around the world, the empirical evidence behind the model is still limited. We do not yet have a solid understanding of how a SIB works to impact service delivery or provider behavior, and how this in turn impacts intervention outcomes for service users. More direct empirical analysis of SIBs is needed, particularly analysis that incorporates the complex role of different actors, pays great attention to context, and seeks to try and disentangle the intervention effect from the SIB effect (Tan et al. 2019; Fox and Morris 2019; Fraser et al. 2020). In the remainder of the thesis, I offer three separate papers that aim to contribute to critical gaps in our understanding on how SIBs work in practice, related to the model's impact on service delivery systems and front-line staff.

In the first paper, I make theoretical and empirical contributions to the debate around how using a SIB impacts the system of public service delivery. I first propose a conceptual framework for analyzing the impact of the SIB model on public services, stemming from the four key features of SIBs to date and based on the literature reviewed here. I then apply this framework in a case study of a SIB launched in Essex, using in-depth interviews, document review, and observation of front-line staff. With these insights, I discuss how the SIB model impacted public service delivery in Essex and conclude with suggestions for where knowledge may be transferable. I conclude with a discussion of policy implications, including the importance of safeguarding democratic accountability, developing positive relationships, and mitigating disproportionality if government agencies continue to pursue SIBs. This paper addresses key gaps in the literature by making a theoretical and empirical contributions to the question around how the SIB model impacts systems of public service delivery.

In the second paper, I investigate whether using a prescribed and replicable (i.e. high-fidelity) intervention insulates providers and front-line staff from potential negative effects of the SIB model. Based on the work of previous scholars laid out in this literature review, I lay out four categories through which the SIB model may affect front-line staff: performance management, support networks, flexibility, and incentives. SIBs often fund 'black-box' interventions, where the process of delivering a social intervention is not prescribed nor monitored and only the outcomes matter. This causes concerns amongst public servants and scholars around the potential of gaming outcomes contracts, for example by focusing on service users most likely to generate outcome payments at the expense of other, harder-to-reach service users. In this paper, I draw on my in-depth case study of the Essex SIB to analyze how using a high-fidelity intervention with the SIB model affects providers and front-line staff. I conclude with a discussion on the promises and drawbacks of using the SIB approach in various contexts and suggest avenues for further research, including the consideration of 'open-box' interventions: those which are not prescribed but are monitored and held to minimum quality standards. This paper builds on the existing body of knowledge by analyzing the effect of the SIB on providers and front-line staff in the particular context of a high-fidelity intervention.

For the third paper, along with Drs. Eleanor Carter and Mara Airoldi, we investigate whether SIB projects have been designed to accomplish two of the model's original intentions: shifting the focus of public service delivery to achieving impact and transferring risk away from the government partner. By conducting qualitative content analysis on the SIBs launched to date in the US and UK, we plot and analyze how the design of SIB projects vary across these two key dimensions, i.e. the degree to which payment is linked to impact and the nature of the working capital. Analyzing the variation of SIB projects across these components helps to determine how (if at all) the SIB model is different than other commissioning approaches (such as fee-for service or payment by results). This will be the first systematic attempt to map variation in design of SIBs against the espoused model, and allow us to draw conclusions on the model's ability to live up to its original promises. In particular, if SIBs have "stretched" too far from what makes the model a distinctive, innovative approach this could undermine its potency as a tool for service delivery improvement.

Article One - More Than Just a Clash of Cultures: Using Social Impact Bonds to Deliver Public Services

Abstract

The social impact bond (SIB) model is subject to heated debate, as it involves private investors funding social services with the potential to generate returns for achieving outcomes. Despite this debate, and the model's increasing application to policy areas including homelessness, child welfare, and employment support, it is unclear how the model impacts processes of public service delivery. In this paper, I offer a conceptual framework for how SIBs might impact public service delivery that stems from the SIB model's four key features. I then apply this framework in a case study of a SIB launched in Essex, using in-depth interviews, document review, and observation of front-line staff. I find that the SIB had mixed effects on public service delivery in Essex. The SIB model enabled rigorous monitoring of performance, availability of additional resources for the service, and outcome goal alignment across stakeholders. However, the SIB model also generated stakeholder conflict, feelings of resource burden and disproportionality by the government partner, loss of democratic accountability, and the pressuring presence of an underlying financial driver. I conclude with a discussion of policy implications, including the importance of safeguarding democratic accountability, developing positive relationships, and mitigating disproportionality if government agencies continue to pursue SIBs.

Introduction

The UK's public sector reform agenda has long been characterized by attempts to incorporate private sector ideals into the way government functions. Following economic recession in the early 1980s, New Public Management (NPM) - the idea of applying private-sector mechanisms to increase the efficiency of public services - came to the forefront of the reform agenda (Hood 1991). NPM strategies aimed to improve the

quality of public services, while simultaneously saving the government money by introducing market discipline (Pollitt and Bouckaert 2004). Nearly forty years later these ideals persist, thriving in an era of austerity where budgets are getting smaller and local governments are being pushed to look for new ways to fund vital social services (Neyland 2018). Policymakers are again turning to the private sector for solutions, recently with the introduction of the social impact bond (SIB).

For market discipline enthusiasts, SIBs are a holy grail of public-private partnership; bringing in private sector techniques to achieve public good, all without risking public money. In their most basic form, SIBs are a three-way partnership between a government commissioner, social service provider, and private investor. Generally, a government commissions services and agrees to pay on specific outcomes, a service provider delivers an intervention that is closely monitored, and private funders provide the upfront capital and are repaid only based on the achievement of outcomes. The SIB changes the process of public service delivery by reframing the focus to impact and bringing in an external investor whose money is at risk. Since 2010, over 85 SIBs have been launched in the UK to address a variety of social issues, including criminal justice, homelessness, and children's social care (INDIGO 2021).

The introduction of private-sector actors and new forms of market discipline to the public sector, concepts stemming from NPM ideals, have been used as justification for how SIBs could improve public service delivery. In their 2016 review of the early literature on SIBs, Fraser et al. identified a consistent narrative in the literature that the SIB's private sector management techniques could improve public sector shortcomings in service design, delivery, and accountability. Proponents of the model argued that because private sources of funding were tied to explicit measures of service provider performance, mechanisms of accountability and oversight would be introduced in ways not typical to public services (Burand 2013; Lagarde et al. 2013; Stoesz 2014). Greater discipline in the use of resources, through the SIB model, could help minimize inefficiencies found in typical government spending by paying only for specified outcomes (Bafford 2012; Edmiston and Nicholls 2018). Assumptions that incorporating private sector actors and

techniques would lead to improvements in public service delivery and efficiency stem firmly from NPM ideals (Osborne 2006).

In practice, however, the SIB model may not fit neatly into all components of the NPM framework. As iterated by Ferlie et al (1996), one primary goal of the NPM movement is to increase market competition in public services (the authors identify the three M's of NPM - markets, managers, and measurement). Whilst SIBs indisputably emphasize management and measurement, it is not clear what role they have played in fostering market competition. In fact, Neyland (2018) argues that SIBs serve as an anti-market device. In his work analyzing a SIB launched in Essex (the same to be analyzed in this paper), he argues that the SIB "ruled out competition and jumped straight into guaranteeing a local monopoly" by fixing in place the intervention, the providers, and the investors. While a lack of a competitive market goes against NPM ideals, this has potentially been an outcome of context (i.e. few social investors willing to invest, provider-led projects, the way local governments approach contract negotiations, etc.) rather than an intentional design of the SIB model. Initial arguments from SIB proponents actually claimed that these projects could increase competition, leading to private sector-led innovation and and "encourag[ing] new market entrants" in the world of public services (Fox and Albertson 2011).

Other scholars have argued that perhaps New Public Governance (Osborne 2006) may be a more appropriate lens with which to analyze SIBs (Dayson, Fraser, and Lowe 2020). New Public Governance (NPG) is a framework that highlights the reliance on preferred suppliers, rather than a competitive market, and emphasizes neo-corporatist values of co-designed interventions and collaborative, trust-based relationships (Osborne 2006). While it is helpful to draw on the NPG framework when analysing the SIB's emphasis on collaboration amongst multiple stakeholders in a complex system (Lowe et al. 2019), for most components of SIB design and implementation the underpinning logic is more deeply rooted in NPM (Dayson, Fraser, and Lowe 2020). Previous scholars have discussed how SIBs have grown out of a context of austerity, where public agencies are being 'pressed to do more with less' (Tan et al. 2019; also see

Neyland 2018). Others argue that the SIB model is an extension of privatization of public agencies, and increases the risk of financializing social services with a focus on cost-reduction (Dowling 2017; Tse and Warner 2018; Cooper, Graham, and Himick 2016). SIBs have imported "ideals, values and tools" from private sector management (Lowe et al. 2019), using market-based approaches to tackle challenging social problems typically seen as government responsibility (Nicholls 2010). These arguments strongly support the applicability of the NPM framework for analyzing SIBs.

Much like other initiatives with roots in NPM logic, the SIB model is steeped in debate - as one might expect when core public services are being financed by private investors. The model itself is "strategically ambiguous" (Tan et al. 2019), and may have gained popularity in part because it "can be bent" in various directions "depending on the preferences of the audience" (Maier, Barbetta, and Godina 2018). This ambiguous nature is reflected in the debate between proponents and critics of the SIB model. Proponents of the SIB argue that the model shifts the focus of service delivery to measuring and achieving impact, introduces market discipline, encourages innovation, and creates multi-sector partnerships that improve service delivery (Fraser et al. 2016). Critics of the SIB model in turn argue that this approach introduces profitization values, encourages perverse incentives, fails to improve service quality, and necessitates high transaction costs (McHugh et al. 2013; Warner 2013; Rowe and Stepheson 2016; Roy et al 2017). Maier, Barbetta, and Godina (2018) identify paradoxes of the model, reflected in these debates, including the SIB's promises to be both flexible and evidence-based, and ensure cost-savings while transferring risk.

Existing research does not paint a clear or complete picture of how SIBs impact public service delivery in practice. In a review of empirical evaluations of SIB projects (n=16), Fox and Morris (2019) highlight the need for more high-quality evaluations, particularly those that do not conflate the SIB model itself with the intervention being funded. Furthermore, Lowe et al. (2019) point out that most studies to date have focused on "high-level outcomes" of SIBs rather than "examining the detailed complex set of interactions" between stakeholders. Whilst there have been a number of high-quality evaluations and analyses of SIB projects,

the insights of which are incorporated into this paper's conceptual framework, there is room for more research that provides nuanced, context-rich detail on how the SIB model plays out in practice. Given the SIB's increasing use by government agencies around the world, we would benefit from a greater understanding of how the model impacts public service delivery.

In this paper, I make theoretical and empirical contributions to the debate around how using a SIB impacts processes of public service delivery. I first propose a conceptual framework for analyzing the impact of the SIB model, stemming from the four key features of SIBs to date. Then I apply this framework in a case study of a SIB launched in Essex, which provided multi-systemic therapy to children at risk of entering the social care system. Using a detailed case study provides the opportunity for an in-depth examination of the sequence of events and relationship dynamics in Essex, providing a nuanced and holistic picture of a complex case (Bennett and Checkel 2014; Cartwright 2017). With these insights, I discuss how the SIB model impacted public service delivery in Essex and conclude with suggestions for where knowledge may be transferable.

Conceptual Framework: How a SIB might impact public service delivery

While the precise definition of a SIB varies across stakeholders, based on a review of the literature and the existing 206 SIB projects (INDIGO 2021) I suggest four basic features characteristic of projects considered to be a SIB: multi-sector partnership, data monitoring system, payment for outcomes contract, and external investment.

Table 6: Four key features of a SIB

SIB Feature	Description
1. Multi-sector partnership	Generally, at least three stakeholders (typically government, provider, and investor) work together to launch a SIB project. There may be multiple government agencies, providers, and/or investors involved with the deal. There is often also an intermediary and an evaluator involved.
2. Data monitoring system	Systems of data collection and analysis are set up in order to track outcomes achieved in SIB projects. These systems are used to hold provider(s) accountable for meeting performance metrics and are used to determine payments made.
3. Payment for outcomes contract	A contract is signed by relevant parties that establishes a series of payable outcomes. The commissioner (or other paying agency) only pays for services delivered if and when these pre-specified outcomes are achieved.
4. External investment	A private investor provides upfront at-risk capital to the service provider to fund the project and is repaid based on the achievement of outcomes.

These four features serve as a starting point for the development of a conceptual framework that highlights how SIBs can impact processes of public service delivery. A review of the literature on promises and critiques of the SIB model, in the *Literature Review*[7] chapter and summarized below, sheds light on the primary mechanisms through which each SIB feature could improve or impede public service delivery.[8] These are summarized in Table 7 on page 65.

The ***multi-sector partnership*** feature could enable ***collaboration*** across public, private, and voluntary sector stakeholders (Burand 2013; Fraser et al. 2016; Carter et al. 2018). In practice, this could take the form of stakeholder cooperation and joint-working, an effective governance structure with representation from all parties, and an intervention that is well-integrated into the existing service array (Liebman 2011; Nicholls and Murdock 2012; Liang, Mansberger, and Spieler 2014; Arena et al. 2016). Maier, Barbetta, and

[7] For the purposes of the thesis submission, I reference an earlier chapter here. As a standalone article for publication, I plan to cite the thesis instead.
[8] The term mechanism is used as conceptualized by Pawson and Tilley (1997) in their work on realist evaluations.

Godina (2018) argue that the collaborative element of SIBs has potential for improving social policies by bolstering knowledge flows across sectors. However, a multi-sector partnership could also lead to ***conflict*** amongst the stakeholders. This might be demonstrated by relationship tensions, unclear lines of accountability or governance structure, and a lack of integration into the existing system (Warner 2013; Rosenman 2014; Giacomantonio 2017; Ronicle and Smith 2020). Misalignment over goals, expectations, and data management and service delivery processes between government, provider, and investor partners can lead to breakdowns in relationships and inter-organizational conflict (Warner 2015; Roy, McHugh, and Sinclair 2018; Fraser et al. 2020)

The ***data monitoring system*** of a SIB could enable more ***active management of contracts,*** with additional resources allocated to project monitoring to ensure the service is running smoothly and minimize inefficiencies (GPL 2017; Mulgan et al. 2011; Bafford 2012). This could enable the continuous use of data to identify challenges and implement service delivery improvements in real-time or guide local decision-making (GPL 2017; Fraser et al. 2020). For example, in their work evaluating a SIB launched in Peterborough, Disley et al. (2015) reported the use of data to review practice as one of the main benefits of using the SIB model. However, such a system could also create a ***resource drain*** as project partners are required to collect, monitor, and share data frequently (Disley et al. 2011; Finn 2011; Demel 2012; Edmiston and Nicholls 2018). Lowe and Wilson (2017) find that while increased performance management can in some cases lead to improved practice, it can also take staff time away from supporting clients and encourage perverse behaviors related to meeting data requirements or performance targets. A resource drain might also take the form of administrative burden, pressure on front-line staff, or a feeling of disproportionality in terms of the time spent on the SIB compared to other programs or services (Fox and Albertson 2012; Tan et al. 2015; Roy, McHugh, and Sinclair 2018; Ronicle and Smith 2020). Pressures on data collection can be frustrating for provider and government partners, particularly if the monitoring system focuses on processes of service delivery rather than service user outcomes (Fraser et al. 2020).

External investment could unlock ***flexible funding***, as resources may not be tied to the same regulations as public funds (Griffiths and Meinicke 2014; Fraser et al. 2016; Carter et al. 2018). Flexible funding may enable more tailored support to service users and additional provider discretion on how best to allocate resources to achieve results (Leventhal 2012; Fox and Albertson 2012). The SIB model could create a space where service providers are "free to design new initiatives" to achieve outcomes, by "exploiting synergies between different actors", "reconfiguring the structure of service delivery", or "creating opportunities of cross-fertilization between different sectors" (Arena et al. 2016). This flexibility for innovation is arguably enabled by risk transfer, with external investors shielding government and providers from the risks of project failure or poor performance (Fox and Albertson 2012; Fraser et al. 2018; FitzGerald et al. 2019; Chiapello and Knoll 2020). However, an external investment may also lead to a ***loss of democratic accountability***, or the ability for the public sector to influence the service and make program decisions (Joy and Shields 2013; McHugh et al. 2013; Warner 2013). A loss of democratic accountability may be demonstrated by a disproportionate use of resources, program decisions dictated by a returns-motive instead of by the public interest, or a lack of transparency (Joy and Shields 2013; McHugh et al. 2013; Fox and Albertson 2012; Warner 2013; Dowling and Harvie 2014; Fraser et al. 2016; Morley 2019). There are also concerns that this element of the SIB encourages the financialization of vulnerable segments of the population, by allowing investors to shape policy and turning service users into potential sources of revenue (Dowling 2017; Sinclair, McHugh, and Roy 2019; Tse and Warner 2018).

A ***payment for outcomes contract*** may have the effect of ***outcome goal alignment*** - aligning all stakeholders to drive towards improving a meaningful outcome of interest (Liebman 2011; Lagarde et al. 2013; Stoesz 2014; Edmiston and Nicholls 2018). Outcome goal alignment could be demonstrated through consistent belief in outcome goals across all parts of the system, and a greater achievement of outcomes that lead to impact (Cohen 2011; Fox and Albertson 2012; Baliga 2013). It could also be demonstrated through increased mechanisms of market-discipline, by improving accountability practices and efficient allocation of resources, as payment is made only for specified outcomes (Edmiston and Nicholls 2018).

However, the payment for outcomes contract could also potentially enable ***perverse incentives***, as payment is only made if particular outcomes are demonstrated (Considine, Lewis, and O'Sullivan 2011; Finn 2011; Rees, Whitworth, and Carter 2014; Lowe and Wilson 2017; Edmiston and Nicholls 2018). This could result from a narrow scope of services focusing only on payable outcomes at the expense of more meaningful measures, or only on certain segments of the target population (Fox and Albertson 2012; Dowling 2017; Tse and Warner 2018; Ronicle and Smith 2020). It could also result in gaming the outcomes contract, for example by focusing on easier to serve clients who are seen as more likely to achieve the payable outcome (Joy and Shields 2013; Sinclair et al. 2014; Carter and Whitworth 2015; Fraser et al. 2018; Ronicle and Smith 2020).

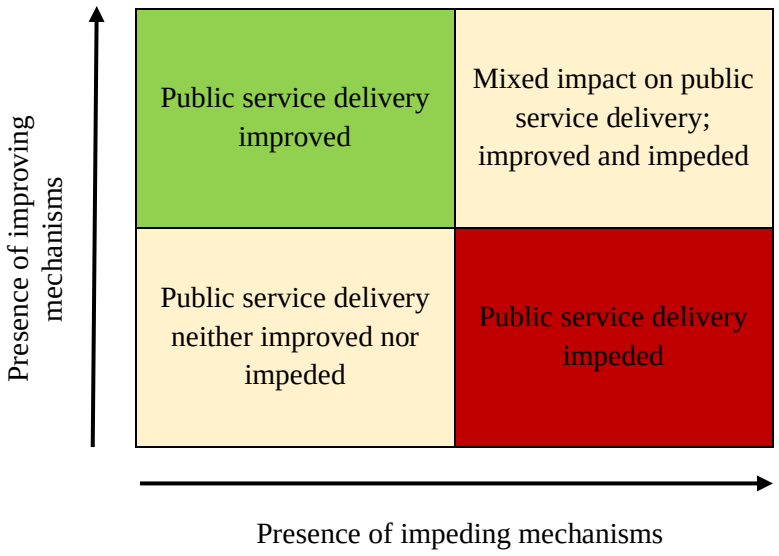
Table 7: How particular features of SIBs might impact public service delivery

SIB features	1. Multi-sector partnership	2. Data monitoring system	3. External investment	4. Payment for outcomes contract
Mechanisms to improve public service delivery	**Collaboration**	**Active contract management**	**Flexible funding**	**Outcome goal alignment**
What this may look like on the ground	Stakeholder cooperation; Effective governance structure; Intervention well integrated into existing system	Continuous use of data to identify challenges; Implementation of improvements in real-time	Tailored support to service users; Provider discretion; Additional resources available for service delivery; Risk of project failure with external investor	Outcome-driven service with buy-in at all levels; Outcomes achieved lead to impact for service users
Mechanisms to impede public service delivery	**Conflict**	**Resource drain**	**Less democratic accountability**	**Perverse incentives**
What this may look like on the ground	Stakeholder clashes; Relational tensions; Lack of service integration; Unclear lines of accountability	Administrative burden; Pressures on front-line; disproportionality	Decisions on public services dictated by non-public motive; Lack of transparency; Inefficient use of resources	Narrow scope of services; Skepticism around outcome goals; Gaming to earn outcome payments

This framework, stemming from the specific features of the SIB model, provides a starting point for analyzing the impact of using a SIB on the system of public service delivery. It draws on insights from related scholarly literatures. Contract theory would suggest the potential for moral hazard - where the investor partner may be motivated to act in its own best interest to the detriment of the government partner's best interest, thus creating space for perverse incentives (Holmstrom 1979). Theories of joint-working would suggest that relationships, goal alignment, and management structures are important elements of successful collaborations (Huxham 2003). Literature on performance management points to the potential for pressures on front line-staff when ramping up reporting requirements (Bruijn 2002). While the framework builds explicitly on the literature around SIBs, it is important to recognize that the SIB literature in turn draws on insights from various related disciplines.

The mechanisms identified in the framework to improve public service delivery (improving mechanisms) are not mutually exclusive from the mechanisms which may impede public service delivery (impeding mechanisms). In other words, it is possible for any combination to appear during the course of a SIB project. The goal for policymakers, however, is to maximize the presence of improving mechanisms while minimizing the presence of impeding mechanisms, to end up in a space where public service delivery is improved (see Figure 3 below).

Figure 3: Improving and Impeding Mechanisms



Methods: Using the case of Essex

This conceptual framework provides the basis for a case study of a SIB launched in Essex, United Kingdom. The case study draws on the key principles of process evaluation. Process evaluations are typically applied to interventions with a clear theory of change, and attempt to measure the implementation of program activities; for example whether the program was delivered as intended, in the appropriate quantities, and to the full range of the target population (Steckler and Linnan 2002). It is difficult to apply this traditional lens of process evaluation to the SIB mechanism itself, particularly as the SIB model does not have a clearly defined theory of change. The SIB model doesn't fit neatly into an implementation fidelity framework because a "SIB" is an ambiguous term (Tan et al. 2019), and it is unclear exactly how it is meant to impact service delivery. The conceptual framework, laid out in the previous section, attempts to clarify the mechanisms through which a SIB may impact service delivery. As articulated above, these mechanisms can be improving or impeding, similar in concept to barriers and facilitators of successful intervention implementation (Schneider et al. 2009). The case study that follows attempts to measure how the SIB model was implemented throughout the project, and the impact that the different features of the SIB had on service delivery. Using interviews, document review, and participant observation, the analysis was designed to uncover how the SIB model altered processes of service delivery (i.e. through which mechanisms), what the effect of those alterations were, and for whom (Pawson and Tilley 1997; Steckler and Linnan 2002). This research received ethics approval from the Blavatnik School of Government's Departmental Research Ethics Committee at the University of Oxford (see Appendix 2).

Using a single detailed case study has advantages, as it allows for the in-depth examination of the events, processes, and relationship dynamics, providing a nuanced and holistic picture of a complex case (Bennett and Checkel 2014; Cartwright 2017). SIB projects, involving complex partnerships and multi-faced policy goals, are well suited to detailed and contextualized case studies that explore questions relating to how, why, and for whom (Yin 2013; Pawson and Tilley 1997). Other scholars who have studied SIBs agree;

arguing that qualitative case studies are helpful for exploring implementation of policies (Fraser and Mays 2020) or that small 'n' designs are well suited to the evaluation of policy instruments like SIBs (Fox and Morris 2019), and that single case study designs are useful when studying new phenomena such as SIBs (Lowe et al. 2019). However, single case study designs also have a number of disadvantages. Case studies are based on the subjective interpretation of the researcher, calling in to question replicability and objectivity (Berg and Lune 2012). Furthermore, conclusions from case studies are often difficult to generalize outside of the specific context and are therefore likely not to have external validity (Gerring 2004). Depending on the selected case there may be different relationships between the variables or phenomena being studied (Berg and Lune 2012). These disadvantages are important to note in the context of the following case study on the Essex SIB.

In 2012, policymakers decided to test the SIB model in Essex to tackle a growing crisis: record numbers of children entering statutory care. Ofsted, the UK government's inspector of children's services, had recently rated Essex County Council 'inadequate' – the worst possible rating indicating widespread and serious failures within the council's children's social care system (Ofsted 2010). To tackle this challenge, the council began to implement a series of reforms, including the launch of a SIB to provide multi-systemic therapy (MST) to children at-risk of entering care ('on the edge of care').

MST is a license-based intervention where therapists work with parents to manage young person behavior, and work with the young person to engage in positive activities and improve family relationships. In the SIB project, investors provided the upfront financing to deliver MST in Essex. The council agreed to pay investors back based on the number of days the children served by MST spent out of care (compared to a historical baseline). This was the first time a SIB was used in the local government context, and the second ever SIB project in the UK. The key stakeholders of the project included:

- Essex County Council (ECC) or the 'council': the government partner responsible for children's social care in Essex, whom agreed to make payments on outcomes achieved. Key stakeholders

include the Head of Children's Services (who directs the entire array of children and family services in Essex), the Project Lead (who has a leadership role in the children's services agency and served as the primary government lead on the SIB project), the Service Manager (who has a leadership role in Essex's children social care work), the former Head of Partnership (who had a leadership role in Essex's commissioning unit and helped plan the SIB prior to its launch), the Strategy and Innovation Manager (who works in Essex's commissioning unit and worked directly on the SIB), and front-line staff serving children and families through programs similar to MST.

- The provider: a large charity operating throughout the UK, responsible for delivering the MST service in Essex. Key stakeholders include the Operational Director (who oversaw a number of the provider's projects, including MST in Essex), the Program Manager (who worked directly on the Essex project with MST supervisors, Council leadership, and the SIB Board), and MST supervisors and therapists (who delivered the service in Essex).
- The SIB Board: A special purpose vehicle (SPV) made up of investors and an advisor, responsible for upfront funding of the SIB project. The SIB Board met frequently throughout the project and played an active role in data and project management. Key stakeholders include investors (who contributed to the upfront funding of the project and remained engaged in managing the project throughout its delivery), and an advisor (who had previous experience of MST delivery in other jurisdictions).
- The intermediary: An organization specializing in SIBs, responsible for managing the project and serving as the link between all project stakeholders. Key stakeholders include the Intermediary Leads (who worked directly on managing the Essex project).

Representatives from each of these four groups (council, provider, SIB board, and intermediary) met regularly throughout the entire duration of the project to monitor project performance. In addition, the SIB board met separately (monthly and then quarterly) to receive updates from the provider and intermediary and play an active program management role. Appendix 5 provides a list of the key stakeholders from the Essex project.

In December 2018, the SIB project in Essex came to an end. Over 380 young people received MST over five and a half years, and the council paid approximately £7.1m to investors in outcome payments (the agreed upon payment cap) for a program that cost £5.8m (ECC 2012; GO Lab 2019). Approximately 19% of the young people who received MST entered care after the services ended[9] (Drew et al. 2019). This rate of children entering care compares poorly to other sites around the UK: in a randomized control trial of nine MST sites throughout the UK, Fonagy et al. (2018) report 13% of young people receiving MST had an out-of-home placement in the 18 months after services.[10] However, the performance of MST in Essex was much better than the historical baseline project partners used to determine outcome targets. At the time the baseline was created, Essex had the third highest level of children in care across all UK local authorities (GOV.UK 2011). In addition, the children in care were unsurprisingly facing serious challenges: in 2011, of all children in care, 56% were at risk of experiencing abuse or neglect, 10% were in families experiencing acute stress, and 18% were in families experiencing dysfunction (Ibid). Stakeholders assumed a rate much higher rate than 19% of children accessing MST would enter care after services ended, and as a result, the SIB far surpassed its outcome targets and investors were repaid in full plus a return.

As the Essex SIB was in its final stages of delivery at the time of this analysis, it serves as a useful case study through which to investigate the impact of the SIB model throughout the entire process of service delivery. Semi-structured in-depth interviews were conducted with all available project stakeholders (n=31) aside from service users. Service users were kept separate from the SIB element of the project, and so were interviewed in a separate study (Drew et al. 2019) that aimed to unpack their experiences with the MST service. Interviews conducted included Essex County Council leadership and employees (n=11), SIB Board members (n=3 investors, 1 advisor), intermediary leadership and staff (n=3), provider leadership and staff (n=10), and others involved with the project (n=3). The interview process began with warm contacts who

[9] At the time of analysis, 302/388 young people served by MST had at least one year of post-MST data. 56 of these children entered care after exiting MST services (19%).
[10] For further comparative context, MST UK (2020) reports that nearly 95% of young people throughout the UK and Ireland remain at home after receiving MST.

worked directly on the SIB project, and then used a snowball sampling approach (Bryman 2015) to identify additional project stakeholders to invite to interview. Interviews ended when the list of contacts generated by the snowball method was fully exhausted (interviews were conducted with 31/32 individuals recommended), and took place between July through December 2018. Interviews were conducted individually with each informant, to help facilitate an environment where respondents could openly reflect on their experience and opinions of the SIB project. Using a semi-structured approach (Bryman 2015), interviews were conducted flexibly, allowing informants to answer open-ended questions and shifting the emphasis of the conversation based on the informant's role and previous responses. The interview topic guide was designed using the conceptual framework and set out questions related to relationships between the various SIB actors, project management processes and expectations, processes for responding to project challenges and perceptions around flexible decision making, and processes and perceptions on working towards outcome goals. The guide also included a portion that invited respondents to reflect freely on how (if at all) they felt the SIB structure impacted the process of delivering MST, in order to open the conversation to ideas outside of the original framework.[11]

Interviews were supplemented by comprehensive documentary review on existing board papers and presentations, meeting minutes, contracts, evaluation papers, and project reports (n=229 documents, 2028 pages). These documents include both confidential and proprietary information that project partners were willing to share (board papers and presentations, meeting minutes, contracts), as well as documents in the public domain (reports and evaluations). The documents, particularly the board papers and meeting minutes, allowed for the creation of a comprehensive project timeline that laid out how the project was developed, from inception through completion, including evidence of project challenges and how they were dealt with, progress towards meeting outcome goals, data monitoring processes and subsequent actions taken, points of disagreement between various actors, and approaches to project governance. The

[11] See Appendix 6 for list of topic guides.

stakeholder interviews and documents were coded thematically, with help from the NVivo 10 software, using an integrated approach (Weston et al. 2001) combining both inductive and deductive approaches by beginning with the initial themes set out in the conceptual framework (deductive) but allowing space for the data to inform the theory (inductive). Themes that were explored included stakeholder collaboration and conflict, data monitoring and performance management practices, resource drain and disproportionality, flexibility in service delivery, loss of government control over service direction, stakeholder alignment to outcome goals, and perverse incentives.

To further supplement the data collected, participant observation was conducted through three days of shadowing front-line staff while the service was still being delivered. I was able to silently observe front-line teams conduct their daily activities serving families. While this data collection strategy was agreed to be used strictly for background purposes, participant observation can help to provide a richer understanding of the project and enhance and inform further data collection and analysis (Musante (DeWalt) and DeWalt 2010). Participant observation is a useful tool to help the research understand context and also critically examine any assumptions of the research (Ibid).

In addition to the disadvantages in research design articulated above, it is important to also note that the specific qualitative methods undertaken for the analysis are subject to a number of weaknesses. Interviewee responses and participant behavior can be biased by the presence of the researcher (Bryman 2015). For example, participants may have felt the need to self-censor when discussing relationship tensions between and within stakeholder groups, pressures to meet outcome goals, enrollment or referral strategies, or other topics they felt may have cast them or other stakeholders in a negative light. In addition, as interviewees were asked to reflect on past events, responses are subject to recall bias (relying on memories of events instead of events themselves) (Ibid). Interpretation of the results is also be subject to bias, as it is influenced by my perspective, past experiences, and beliefs (Berg and Lune 2012). My previous work experience providing technical assistance to government agencies pursuing SIBs, as well as my current affiliation with

the Government Outcomes Lab at the University of Oxford (which conducts research and provides technical guidance on SIB projects) influences the way I view SIBs and the assumptions I hold about the model. Acknowledging this through a lens of reflexivity suggests that my own beliefs and values have likely influenced the research process and outcomes reported in this paper (Berger 2015). Efforts were taken to mitigate potential biases, for example by conducting interviews individually, offering all respondents to be anonymized, and triangulating data from three different sources.

Despite limitations in design and methods, findings from this case study provide rich, contextualized details on how the SIB affected public service delivery of MST in Essex. These findings are important because there is a lack of clarity in both policy-making and academic fields on the impact of SIBs on processes of public service delivery. While the particular impact of the SIB model in the Essex case may not be generalizable to all other SIB projects, there are important lessons we can draw regarding the potential for a SIB to impact public service delivery.

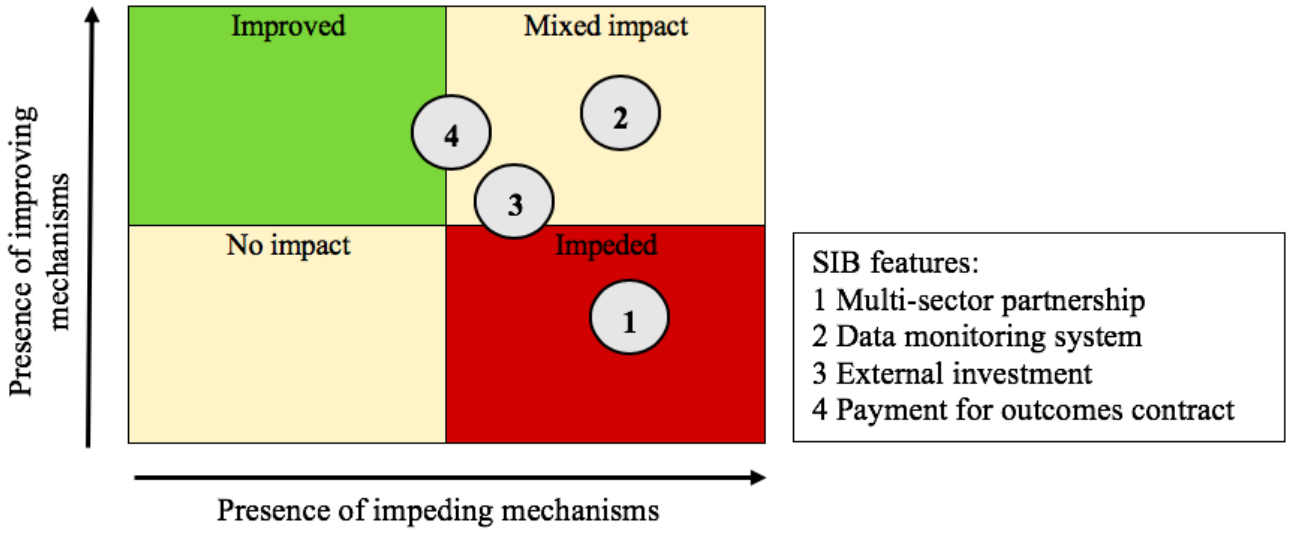
Findings: How the SIB impacted service delivery in Essex

The SIB had mixed impacts on the delivery of MST in Essex. The following section provides detail on how each of the SIB features impacted public service delivery in this case.

Table 8: Summary of findings from the Essex SIB

SIB features	Findings
1. Multi-sector partnership	**More conflict than collaboration** - Difficulty integrating SIB into the system - Challenges around governance structure and relationship tensions - Clash of cultures between private, public, and voluntary sector ways of working
2. Data monitoring system	**Active contract management at a high perceived cost** - Continuous monitoring enabled identification of challenges and trailing of solutions - Burden on stakeholders and a sense of disproportionality relative to other programs
3. External investment	**Limited flexibility and loss of democratic accountability** - Flexibility to spend additional resources but often faced pushback - Locked into contract specifications with restricted ability to course correct - Enabled new service in array of support for children at the edge of care
4. Payment for outcomes contract	**Aligned to outcome goals, but not fully protected from financial drivers** - Aligned to keeping children out of care, albeit with misalignment over scope of the target population and services provided - Presence of clear underlying financial driver putting pressure on system

Figure 4: Impact of SIB features in Essex



Multi-sector partnership: facilitating collaboration or creating conflict?

Integration of the SIB into the system

"Being able to really identify how we worked differently took just a little bit – like any service really – took a little time to embed into the wider perspective" (Interview, MST Supervisor 1).

The MST service faced challenges integrating into the existing social care system in part due to the lack of operational buy-in at the project's outset. The SIB project was developed by Essex's commissioning unit; described by Head of Children's Services as in a "parallel universe" from the social care system. There is a perception that the operational-level staff would have designed something very different had they been involved in the project's initial stages (Interview, SIB Board Advisor). The Head of ECC Youth Offending Services, who served as the council's lead on implementation of the project, stressed the importance of involving operational social care services staff alongside the contracting and procurement processes at the project's inception. Doing so would have allowed them to "adjust some of the assumptions that colleagues in commercial services might make about the way that services are set up" (Interview, ECC Project Lead). The operational leadership at the council "didn't own it in the way that they would have done if they had been involved" upfront (Interview, Investor 1) and in addition, the individuals who had structured the SIB had turned over so there was a sense that the project "was not anyone's baby" (Interview, Investor 2). In the "handing over of the baton" from commissioner to social care leadership "some of the buy-in was lost and never recovered (Interview, Intermediary Lead).

Exacerbating the difficulties with integration into the existing system was the length of time it took for the SIB to develop and start delivering services. Notional discussions of the project started in 2010, and it took until 2013 to start delivering services. This three-year lead-in time is described as "really problematic" by the council, as Essex had been facing a crisis in 2010 with too many children in care and needed immediate action (Interview, ECC Head of Children's Services). Under government intervention, the council faced "a

lot of pressure and urgency to bring about improvement within children's services" quickly (Interview, ECC Head of Partnership). As a result, the council launched a service that "overlapped in its applicability with MST," reflecting Essex's efforts to improve children's services "in the time between the [MST] service being commissioned and it actually mobilizing" (Interview, Investor 1). This newly developed internal service became one source of ongoing tension between project stakeholders and slowed the integration of the MST service in Essex.

The internal service was cause for concern amongst the SIB Board due to fear it might impact the success of the SIB. Its launch was "a bit scary because it felt like an internal competing service was being setup which would mean we would have no chance of working with anyone" (Interview, Investor 2). This concern was initially felt on the front-line as well, as the internal service was "fairly new and wanting to forge their role and…there was no clear sense of which referrals were going to fit better for us" (Interview, MST Program Manager). One MST front-line supervisor reflects that "differentiating ourselves" from the internal service and "being able to really identify how we worked differently… took a little time to embed into the wider perspective" (Interview, MST Supervisor 1). From the council's perspective, the launch of an internal service was necessary in order to significantly bring down the number of children in the care system. "We know that the investors were struggling with the idea that we had an in-house project… but there's no way [our numbers of children in care] would have come down at the rate they came down without [the internal service] delivering three times the volume that MST were delivering" (Interview, ECC Project Lead).

Governance structure and relationship network

"There was a triangle rather than a two-way dialogue. There was always a triangle bit. So therefore, when things weren't going well, you were having to negotiate three ways around rather than two ways around. Which is always slightly painful." (Interview, ECC Head of Children's Services)

In the project's initial stages, most communication was done through the intermediary. As a result, "there [was] no direct line or accountability" between the council and the service provider (Interview, Intermediary Lead). For example, complaints from the council on MST referral decisions would be taken to the intermediary, and then the intermediary would contact the MST supervisor to inquire about a specific case (Interview, MST Program Manager). This structure was "frustrating" and felt "hugely bureaucratic" to the council (Interview, ECC Service Manager) and was much different to how the service provider typically operated (Interview, Provider Operational Director). This dynamic shifted with the hiring of an MST program manager who directly liaised with the council, but the council still felt frustration with the governance structure and the lack of control over the service provider (Interview, ECC Project Lead).

Despite indirect lines of communication between the provider and the council, the most difficult relationship to manage was arguably between the council and the SIB Board. Already fragile relationships, due to a lack of operational buy-in at the outset and the launch of a similar internal service, were worsened when referrals were slower than expected at the project's start. SIB Board members felt that the council leadership were not in favor of the MST program, and were "preferencing their own service" in an "unhelpful competitive relationship" (Interview, SIB Board Advisor). Members of the SIB Board felt that the council leadership were "suspicious" of this sort of service that was backed by private investors (Interview, Investor 2). One investor commented, "I don't think Essex would have put two competing services in if one of them wasn't being delivered from outside in a SIB" (Interview, Investor 1). From the council's perspective, the slow stream of referrals at the project's outset was to be expected for any new service, and they were put off by the SIB Board's attempts to make referrals flow faster. The SIB Board "became quite fraught and tense" about the referral volume, which led to significant conversations and strained relations (Interview, ECC Project Lead). The council felt that the SIB Board's reaction reflected "complete inexperience," and led to a series of "unhelpful" suggestions that would have "detracted from the fact that this was about children who were on the very cusp of having care" (Interview, ECC Head of Children's Services).

Stakeholders eventually "got to a mutually agreeable point" (Interview, ECC Project Lead) where the council "accommodated what was happening" with the SIB (Interview, SIB Board Advisor), but relationships never fully recovered. Members of the SIB Board reflect that despite the SIB's success "in terms of the outcomes its delivered and the financial return," the council "were never very happy with the results" (Interview, Investor 3). In contrast, council leadership contend that the SIB has provided "absolutely no added value" and they would not do it again (Interview, ECC Head of Children's Services). The relationships were "quite political" and so stakeholders "hadn't necessarily gotten the value" from those partnerships that they otherwise might have been able to (Interview, ECC Strategy and Innovation Manager).

Merging public, private, and nonprofit ways of working

"Because money was changing hands and money was flowing, and the people who were involved were a mixture of people from a financial background and people from a therapy delivery and medical background, it was a much more eclectic oversight than probably there would be if it was being delivered through the state." (Interview, Investor 1)

There was some sense from the interviews that the intermediary and the SIB Board did not have a comprehensive understanding of the social care system. On-the-job learning about components such as thresholds for entering statutory care, social worker practice, and risk management took place; which is "unsurprising because [the investors] are not working in social care and child protection services" (Interview, SIB Board Advisor). Members of the SIB Board themselves noted that for future projects they should include someone with local authority children's services commissioning experience. This (real or perceived) lack of understanding sometimes led to difficulties in implementation, as the SIB Board and intermediary were the primary drivers of contract management efforts. One MST supervisor reflected on the challenges of getting across to his managers "a little bit more understanding of processes of how social care, social services, social programs work" (Interview, MST Supervisor 2). Clashes with the council

occurred as well, as ECC Head of Children's Services describes "some grim moments" where they didn't agree with suggestions by the SIB Board on how to handle certain social care challenges.

> *"I often felt that I wasn't quite sure why we were spending huge amounts of time going over and over and over stuff with people who had very little knowledge of the social care context in which this particular MST program was being landed, and I understand that there were benefits, but certainly in the early days it was hugely frustrating"* (Interview, ECC Service Manager).

This clash of cultures was most evident in the relationship between the SIB Board and the service provider. The provider, being a large charity organization, has institution-wide standards for salary rates, performance bonuses, dismissal policies, and retention schemes. "They have a certain way of working," while investors bring "a whole different value set" including increasing salaries when you want someone to stay and firing people who are performing poorly (Interview, Investor 3). Offering higher salaries to therapists, increasing the on-call allowance, and developing a retention scheme with performance bonuses were a few of the initiatives tried by the SIB Board which faced significant pushback from the service provider. According to the service provider's Operational Director, the SIB Board "felt that we were slow, process-driven and a bit risk averse" and didn't understand that as a big national organization that we were bound by rules, regulations, and procedures (Interview, Provider Operational Director). In many cases, the SIB Board "pushed and pushed and pushed until we got what we wanted" in pursuit of better outcomes for the children and families involved (Interview, Investor 3).

Data monitoring system: actively managing the project, but at what cost?

Identifying challenges and implementing solutions

"Because you've got a social impact bond with a [SIB Board], we could make very fast decisions to change things" (Interview, Investor 3)

The SIB Board identified various instances where their contract management style allowed them to identify challenges early on and brainstorm potential solutions. Low referral volume was a problem from the project's outset, addressed by the SIB Board through various marketing campaigns within the social care system in Essex to raise awareness of MST. When referrals did start to flow, the MST service struggled to maximize utilization rates, in part due to high staff turnover and infrequent MST trainings available. To mitigate these challenges, the SIB Board created a new therapist-in-waiting role in order to have a replacement therapist ready in the event of any staff turnover. The SIB Board also began to send therapists to the US for expedited MST training, as MST is originally a US program. When the project was coming to a close, the SIB Board recognized the risk of staff leaving as they began to move on to their next job. To mitigate this, the SIB Board rolled out a retention scheme that rewarded therapists for staying through the end of the contract. These instances reflect the difficulty ensuring that MST was serving enough families throughout the course of the project, which continuously "ended up being a really big challenge that required a lot of creativity to try and solve" (Interview, Investor 2).

Attempts to improve service quality were more limited. The SIB Board monitored therapist adherence measures to ensure therapist quality, and suggested routes of action when therapists were falling below expected standards. When issues with children's school attendance were flagged, the SIB Board brought in an educational expert to attend schools and try to rectify these issues. The degree to which these 'solutions' had a positive impact is unclear, but they do reflect continuous monitoring of program progress. From the perspective of the SIB Board, this active contract management style marked a stark change from public

service delivery as usual. “If we hadn’t had a board looking at things and trying to be very active on fixing problems early, things could just run and run for months and years and never get fixed” (Interview, Investor 3).

The service provider and the council had mixed reactions to this management style. From the provider perspective, the push to find solutions was helpful but the resulting actions did not always translate well on the front-line. The therapist-in-waiting role, for example, drew heavy resentment from the front-line staff because it meant a therapist was earning the same salary without doing the same amount of work. In addition, when sent to the US for training therapists “don’t get the same context” that is relevant to the UK experience (Interview, MST Program Manager). From the perspective of Essex’s Head of Children Services, many of the problem-solving techniques employed by the SIB Board are actually common practice in public service delivery and “just what you do when you run big services” (Interview, ECC Head of Children’s Services). Albeit helpful, she did not feel that these tactics were “incredibly creative” or “unique” (Ibid).

Sustained engagement, sustained burden

“It was too heavy and demanding on the system, because it was disrespectful of what people’s roles were within the system” (Interview, ECC Head of Children’s Services).

One of the things brought about by the SIB structure was sustained engagement by high-level stakeholders all aiming to support a specific cohort. Once the SIB was launched, the SIB Board and operational leadership at the council remained stable and engaged over the course of the five-year project, meeting regularly to discuss project progress and workshop issues. However, with this sustained engagement came a need for continuous resources to keep track of the project’s progress – from the council, provider, and the SIB Board. The payment mechanism, in particular, was described by many stakeholders as overly complex and burdensome. Performance management expectations were high, with an unusual degree of oversight

requirements. Provider leadership described being "heavily, heavily performance managed around the delivery of every aspect" in Essex, and commented that other service delivery projects "four times the size" would "entail half of the work." (Interview, Provider Operational Director). The provider's reporting requirements, the SIB Board's tracking and monitoring role, and the council's role in management and data resolution all required significant resources over the entire duration of the project.

As a result, the Essex MST project was monitored more closely by the council than other public services delivered. Typically, in children services in Essex, "a lot of resource goes in at the start of commissioning services, and very little goes into really monitoring and checking" once services are being delivered, but this was not the case for the SIB project (Interview, ECC Project Lead). While this extra monitoring could be considered beneficial, key members of the local authority did not see it as adding value. To the contrary, they felt extra monitoring was necessary because of concerns around the inexperience of the SIB Board in dealing with vulnerable youth. ECC's Head of Children Services felt that the amount of time and resources spent on this relatively small project was hugely disproportionate. She describes it as a "disrespectful, high-maintenance way of operating," because she had to personally shepherd a relatively small contract totaling £7 million over five years, while meanwhile running 550 schools, over 4,000 staff, and £550 million of spend (Interview, ECC Head of Children's Services).

External investment: flexible funding or locked into a rigid contract?

Availability of additional resources

"Focusing on outcomes and being paid per success frees you up to do a load of stuff that you wouldn't be able to do under a normal contract" (Interview, Investor 2)

External funding allowed the MST service to be delivered in Essex, adding another option to support children at the edge of care: "at the end of the day, [the investment] has provided an additional level of service activity in Essex" that would not have been funded otherwise (Interview, ECC Strategy and Innovation Manager). In addition, the SIB Board's response to arising issues throughout the course of the project was frequently characterized by attempts to inject more money into the system; which is often thought of as "completely impossible" in the public sector (Interview, SIB Board Advisor). This injection of additional resources reflected a different approach with more flexibility than if the project had been directly delivered by the council (Interview, Investor 1). The therapist-in-waiting, retention scheme, education expert, and sending therapists to the US for training were all initiatives made possible due to the additional resources available for the project. The SIB Board also decided to continue the service past its original wind-down date in order to reach the target of 380 children served. In addition, a discretionary fund was made available for MST therapists to use to support families by making one-off purchases aimed to alleviate sources of stress. Having resources come from external investors, rather than directly from the council, provided more "freedom and flexibility" to support families and the therapists (Interview, ECC Strategy and Innovation Manger).

Despite having extra resources available to trial improvements, the system wasn't always flexible to accommodate this influx of resources. More often than not, these attempts faced pushback from various stakeholders, and were subject to significant negotiations and delays on implementation despite the SIB

Board being "perfectly happy to pay" (Interview, Investor 2). In addition, there have not been cost-effectiveness analysis on the additional spend. It is unclear whether initiatives ultimately funded through the availability of additional resources increased positive outcomes or if those resources could have been better spent elsewhere. And ultimately, although investors provided these additional resources upfront, the Council ended up paying for them and then some in the form of outcome payments. From the perspective of the council, the project cost £7million and some felt that they could have done it in-house at a lower cost (Interview, ECC Head of Children's Services).

Inflexibility to course correct

"I think there's a disconnect between the use of a license evidence-based program and a social impact bond…You've got one which is a supposedly flexible model, a model that can react quickly if the circumstances change to one that is like a rigid dinosaur and you cannot shift it"(Interview, ECC Project Lead)

Due to the license-based nature of the MST intervention, it was difficult to flex the model to contextual needs or for the SIB Board to pilot new approaches. Any service changes that aren't backed by existing research face pushback from the license holders., as MST Inc. strive for total model fidelity. The process of trying to add in an extra therapist, for example, was described as "tortuous" and "counter-cultural to everyone involved," in part because having an additional therapist was not part of the license-based intervention (Interview, Investor 2). From the council's perspective, they had hoped that MST would be able to assist with some of their high-priority populations, such as teenagers on the autistic spectrum or those at risk of child-sexual exploitation. The model wouldn't allow for it, and was described by one member of the council as "a rigid dinosaur" that you "cannot shift" (Interview, ECC Project Lead). Council stakeholders felt that MST did not give them enough "flexibility to explore the benefits" that the SIB may have brought (Interview, ECC Strategy and Innovation Manager).

The SIB contract also made it difficult for the council to maintain decision-making power on the service direction. By the time the SIB was launched, the context had started to change in Essex and the profile of children entering care shifted, with an increase in gang-involved youth in the local authority (Interview, MST Program Manager) and problems facing the council "looked completely different" (Interview, ECC Head of Partnership). As early as 2014, the council began to feel that MST, with its inflexibility and exclusions, was not the best fit service. Normally, the council would have greater control and more flexibility around the contract, but in this case, they had "limited influences to flex the contract" compared to normal commissioning arrangements (Interview, ECC Project Lead). The perceived complexities of renegotiating the SIB contract prevented the council from changing course. "If it had been a direct local authority contract, we would have activated the break clause and then renegotiated the model of delivery…but the SIB model, just so complex, bought us in to that" (Interview, ECC Head of Children's Services). The nature of the SIB tied them down to "a particular intervention and a particular approach for a significant period of time" (Interview, ECC Head of Partnership). Despite belief that a different service would be most beneficial to children on the edge of care, public servants felt locked into MST for the duration of the SIB contract.

Payment for outcomes contract: aligned to outcomes or to generate returns?

Everyone aligned to (safely) keeping children out of care

"[The project] was about reducing the care population which, at the time it was constructed, was very, very high" (Interview, ECC Project Lead)

Stakeholders expressed widespread consensus around the project's primary outcome of care days avoided, i.e. trying to maximize the number of days children are safely at home with their families and not in the care system. Additionally, it was important to all project partners that the decision of whether or not a child

was placed into care was made independently in order to avoid instances where children were being kept out of the care system unsafely. The decision to place a child in care remained with the family's allocated social worker and maintained independence from the MST service. Secondary outcomes were also tracked, including school attendance, indicators of emotional and behavioral challenges, and offending behavior. However, "tracking these other things is quite tricky," and they were less comprehensively monitored over the course of the project (Interview, Investor 2).

Despite strong agreement on the primary goal of the contract, there was misalignment in expectations on the scope of services provided. In particular, there were disagreements over who constituted as 'edge of care,' how holistic and pervasive the MST service should have been, and the total number of children served through the SIB contract. At the project's initial stages, a debate between the council and the SIB Board arose over "how seriously bad the children and the families had to be before they got referred to MST;" and thus what specific cohort would be eligible for referrals (Interview, Investor 3). This debate continued for the first few years of project implementation. As the project began to wrap-up, the council questioned the wider impact of MST and expressed frustrations around engagements with children previously served by MST in various parts of the social care system. Stakeholders from the front-line, however, debate this perception with the view that MST is "not there make people's lives perfect. [It is] there to keep young people out of care" (Interview, MST Program Manager). Finally, at the end of the project, Head of Children Services at the council expressed disappointment in the total number of children served, saying "for me, it did what it said on the tin and nothing else" (Interview, ECC Head of Children's Services). The internal service, by comparison, was seen as able to feedback learning and have wider impact on the larger social care system (Interview, ECC Strategy and Innovation Manager). Misalignment over expectations on the scope of services persisted throughout the entire course of the contract.

An underlying financial driver

"As social investors we're thinking about: how do we deliver outcomes, and get a financial return? Those are the two things the social investor cares about. But obviously with a SIB it's easier because they're generally very aligned" (Interview, Investor 3)

The project's slow start highlighted some of the potential complexities of financing a social service through private shareholders. After the first few children served by MST went into statutory care, the SIB Board was "super nervous in terms of the commitment and the money" because they'd already "hoovered through most of the start-up money and some of the core capital" (Interview, SIB Board Advisor). During this period, there was "quite a lot of pressure from the investors" on the council to generate more referrals to the service (Interview, ECC Project Lead). The council commented that these concerns ultimately stemmed from a relationship that was not philanthropic and a project that "had to provide back to shareholders" (Interview, ECC Head of Children's Services). At this point, it felt to some council members that the "needs at the local authority in respect to certain children weren't aligning with the requirements of the investors to get suitable referrals through and generate returns on investment" (Interview, ECC Strategy and Innovation Manager).

Despite the payment-for-outcomes contract, the heavy focus on output targets (such as referral and enrollment numbers) remained throughout the course of the project, even as issues with the referral pipeline subsided. Because the SIB Board were not experts in social care, the MST model was mostly left to speak for itself (Interview, Investor 2). As a result, the components focused on by the SIB Board were largely output targets such as referral rates and enrolment volumes. This pressure on volume was felt at the front-line management level, where there was "a pressure from the SIB in terms of generating referrals and opening cases" that fell largely on the supervisors (Interview, MST Program Manager). Indeed, most of the changes initiated by the SIB Board were for the purpose of increasing volume of cases served by MST. One investor commented that much of the drive to increase performance "would have been possible if we were

just being paid on outputs" (Interview, Investor 3). This pressure on output targets, with little more than monitoring of outcome targets, persisted throughout the duration of the contract.

Despite a slow start, the project did begin to experience success above expected levels in terms of care days avoided for children served by MST. The SIB Board began to gain confidence as outcomes were met, with one investor feeling as though "this process really works" and hypothesizing that "the counterfactual was too easy for us" (Interview, Investor 3). As implied by this investor, it is possible that the payable outcomes were too easy to achieve, and this caused some concern amongst the council. The payment cap of £7 million, the maximum amount investors could be paid for outcomes, was reached before the service had finished, suggesting to some that this "was not a good deal for the local authority financially" (Interview, ECC Head of Children's Services). However, even though the payment cap had been reached, the SIB Board extended the service in order to meet the target number of cases. "We did that because we weren't in it to maximize profit, we were in it to deliver the service which was going to help the lives of the children who needed it. We consulted shareholders, and everyone was in favor of it" (Interview, Investor 1). Despite this, the ease with which outcome targets were achieved in Essex and the underlying financial driver in the SIB model caused some concern within the council.

Discussion: What can we learn from Essex

As evident from the findings, the impact of the SIB on the delivery of MST in Essex was mixed. The SIB model provided a new service for vulnerable children and families and enabled rigorous monitoring of performance. However, the SIB model also generated stakeholder conflict, feelings of resource burden and disproportionality by the government partner, and some loss of democratic accountability. These mixed findings mirror the mixed narratives from key stakeholders involved in the project, with the investors generally very proud of the project for meeting the agreed upon outcomes, and the government partner generally disappointed with the overall experience. Despite these divergent views, learning from this case

reflects more than just a clash of cultures and there are overall lessons we can draw. For example, many elements of the specific context in Essex contributed to the way that the SIB impacted service delivery. The lack of operational buy-in at the project's outset, the choice of a license-based intervention, a complex payment mechanism and governance structure, and the changing landscape of children at the edge of care are all contextual components that influenced the impact of the SIB. These learning points may not be inherent to all other SIB projects, but they do provide insight in to how the features of the SIB model can interact with the local context to produce unintended impacts on public service delivery.

Studying the case of Essex also illuminates lessons that relate to the nature of the SIB model itself and will be useful to consider in any future explorations of the model. These include:

Importance of relationships

As SIBs involve bringing together diverse partners who may have different approaches or worldviews, conflict can arise if relationships aren't sufficiently developed. In the case of Essex, relational tensions throughout the project resulted in clashes between stakeholders and difficulties integrating the service in the wider system. Project partners recognize that spending more time developing relationships upfront and ensuring operational-level staff are involved in the planning from the start would have helped to mitigate some of the conflict present throughout the project. La Torre et al. (2019) confirm this finding, arguing that when SIBs are low-collaboration partnerships they face increased risk of failure. The "complex inter-relationships" inherent in SIBs require a large amount of institutional work both to launch a SIB project and then to stabilize it (Lowe et al. 2019). It is often difficult to integrate SIB-funded projects with the wider system and array of other services offered, and communicative relationships are key (Ronicle and Smith 2020).

Risk of disproportionality

Disproportionality increases the opportunity cost of doing SIB projects; the more time and resources are spent on a relatively small cohort, the less are available for other populations or the wider system. This raises a question about the equity of SIB projects and highlights the importance of mitigating disproportionality in future iterations of the model. The complex payment mechanism and resource-intensive governance structures of the Essex SIB project contributed to stakeholder frustrations and a sense of disproportionality, as the SIB funded services for a small proportion of young people relative to the entire children's social care system. This mirrors findings from previous studies, which suggest that SIBs may bring additional layers of bureaucracy (L. Roberts and Cameron 2014; R. Griffiths, Thomas, and Pemberton 2016) more robust systems of program management (Fraser et al. 2018), and high transaction costs (Tan et al. 2015). High degrees of administrative burden, exacerbated by duplicative systems of performance management, can lead to an inefficient use of resources in a SIB project (Ronicle and Smith 2020). This issue is worsened when data collected is used only for purposes of the SIB model itself, and not to improve practice for service users (Fraser et al. 2020). Dowling (2017) highlights the inherent narrow focus of SIBs, as the projects tend to target individuals at the symptom-level rather than "transforming the social, political and economic structures that create such social problems in the first place." The narrow focus, paired with intense time, effort, and monetary requirements of stakeholders across sectors, signal the prevalent risk of disproportionality in SIB projects.

Need for flexibility, particularly to safeguard democratic accountability

SIB projects notoriously take a long time (often several years) to set up, and then set out multi-year contracts that in practice lock stakeholders in to the agreed upon service direction and outcome goals. When the local context or needs change, it is difficult for the government partner to shift direction. Over the course of the SIB project, the context in Essex's children's social care changed dramatically – with fewer children entering care, and the priority population shifting. Despite desires to shift the service

direction in response to these changes, there were two components of inflexibility that limited the council's discretion. First, there was inflexibility within the license-based intervention to target the population deemed highest priority. This largely relates to the specific context in Essex through the choice of MST, whereas in other SIB projects increased flexibility within the chosen service has been demonstrated (Ronicle and Smith 2020). Maier et al. (2018) identify the paradox between flexibility and evidence-based services as a major contradiction of the SIB model, and note that in Essex it was not dealt with in a satisfactory way. However, second there was also inflexibility within the contract arrangements for the council to switch to a new service or provide more service direction. Because the contract was fixed and the perceived renegotiation costs were high, the government agency lost their ability to switch to an intervention that would "provide the right outcomes at the right level and at the right price" (Neyland 2018). This loss of control over the service direction may have contributed to a loss of democratic accountability, which resonates with critiques of the SIB model and other NPM strategies (Warner 2013). By allowing investor or provider partners the discretion to "shape social policy according to their interests," which include a financial gain, SIBs contribute to a "process of de-democratization" that takes power away from the state (Dowling 2017).

Implications of incomplete outcomes specifications

Fitzgerald et al. (2019) identify three components to generating outcome specifications in SIB projects: alignment between payable outcomes and policy objectives, a clearly defined eligible cohort, and accurate price-setting of outcomes. Essex was successful in the first component, as the payable outcome of care days avoided was directly tied to the overall policy goal. However, the eligible cohort was not clearly defined and there is some concern over the accuracy of the benchmark used to trigger outcome payments. Numerous discussions on the cohort eligibility criteria, particularly who counts as 'edge-of-care', carried on throughout the project and were the source of ongoing tension. These concerns resonate with findings from other SIB evaluations; Lowe et al. (2019) found that a SIB launched in North England to address social determinants of health could only serve a subset of the population that General

Practitioners thought would benefit and Ronicle and Smith (2020) found that providers from the Youth Engagement Fund SIBs may not have supported the most disadvantaged young people that they could have. In addition, in Essex, outcome prices were set based on a historical benchmark during a time of very poor practice, perhaps making it too easy to achieve outcome payments in future years. Neyland (2018) notes that in Essex, the "business as usual counterfactual" that payments were based on "depended on a 'business' (sending large numbers of children into residential care) that was no longer 'usual'." Tse and Warner (2018) argue that SIBs in general can lead to overpayment by narrowing the conversation to financial metrics, "extracting funds" from the same public sectors that "SIBs were designed to support." Carefully specifying payable outcomes has important implications for the value government agencies stand to gain (or lose) from using the SIB model.

Taking these four considerations in to account can help stakeholders involved in SIB projects maximize the presence of improving mechanisms (collaboration, active contract management, flexible funding, and outcome goal alignment) while minimizing the presence of impeding mechanisms (conflict, resource drain, loss of democratic accountability, and perverse incentives) to try and drive improvements in public service delivery.

As iterated in the *Methods* section of this paper, it is important to note that this study is subject to a number of limitations in the design, data collected, and analytical methods. Relying on a single qualitative case study makes it difficult to generalize to other SIB projects, calling in to question the external validity of the findings (Gerring 2004). In other cases, there may be different ways that the SIB impacted public service delivery, or different contextual reasons for impacts that the SIB had, for different groups of stakeholders.

Conclusion

This paper offers a new framework for analyzing the impact of the SIB model on the process of public service delivery. Building on previous literature and categorizing existing research, the framework stems from the key features of the SIB model and suggests the mechanisms through which they might improve or impede public service delivery. I argue that this framework is helpful to apply because it enables thorough investigation of contradictory predictions in the literature on the likely effects of the SIB, and aids in the understanding of what features of the SIB model are likely to lead to improvements or impediments in public service delivery.

Applying this framework to the SIB project launched in Essex provided insight on how each feature of the SIB model affected processes of public service delivery. The multi-sector partnership allowed for joint working across sectors, but also lead to significant relationship tensions that greatly influenced the way that stakeholders viewed the project. The data monitoring system enabled rigorous monitoring of project performance with the aim of implementing improvements, but also led to a large resource burden and feelings of disproportionality by the local council. External investment allowed for additional resources to be injected into the system, but flexibility was limited and the local council lost control of the service direction even when they felt a different service would have been better for their constituents. Finally, the outcomes contract aligned all stakeholders to keeping children safely out of care as the ultimate project goal, but introduced an underlying financial driver with potential consequences related to pressure on generating referrals and outcome payments being too easy to achieve.

Through the creation of a new framework and its application to the Essex case, I make theoretical and empirical contributions to the debate on how using a SIB impacts processes of public service delivery.
While the empirical analysis is limited to one case, a noteworthy limitation of this work, both context-specific and SIB-specific lessons can be extracted and applied to future SIB projects. Contextual

components that played a role in how the SIB affected public service delivery in Essex include a lack of operational buy-in at the project's outset, the choice of a license-based intervention, and a complex payment mechanism and governance structure. More generalizable lessons from the case include the importance of developing relationships, the risk of disproportionality, the need for flexibility to safeguard democratic accountability, and the implications of incomplete outcomes specifications. Further research should extend this framework to incorporate other mechanisms through which the SIB could impact public service delivery. Future research could also apply such a framework to other SIB cases to analyze the impact of the model on public service delivery, and directly compare these with other, more traditional forms of government commissioning.

Article Two - Thinking Outside the Black-Box: Using a SIB With a High-Fidelity Intervention

Abstract:

Government agencies are increasingly buying social services by paying for outcomes achieved rather than services delivered; for example, paying for child welfare services based on a reduction in child maltreatment rates rather than the total number of families served. The trend of paying for outcomes is often paired with the use of 'black-box' interventions, where the process of delivering a social intervention is not prescribed nor monitored and only the outcomes matter. This pairing has caused some concerns amongst public servants and scholars around the potential of gaming outcomes contracts, for example by focusing on service users most likely to generate outcome payments at the expense of other, harder-to-reach service users. These fears are further exacerbated with the social impact bond (SIB) model, which involves private investors who provide the upfront capital and stand to gain a return from achieving outcome targets. In this paper, I investigate whether using a prescribed and replicable (i.e. high-fidelity) intervention, as opposed to the black-box approach, insulates providers and front-line staff from potential negative effects of the SIB model. Through an in-depth case study of a UK-based SIB, I find that while using a high-fidelity intervention with the SIB model mitigates incentives to game outcome payments, it also restricts the ability of the provider to innovate and deliver flexible, context-driven services. Using a high-fidelity intervention does not protect providers from pressure to meet output targets, which could potentially create other disincentives in service delivery. I conclude with a discussion on the promises and drawbacks of using the SIB approach in various contexts and suggest avenues for further research.

Introduction:

Government agencies are increasingly experimenting with outcome-based financing, purchasing social programs by paying for outcomes achieved rather than services delivered (Albertson et al. 2018). For example, governments are paying for services for the homeless based on the number of individuals who remain stably housed, paying for workforce development projects based on the number of clients in employment, or paying for education programs based on student's scores. Social Impact Bonds (SIBs) are one type of outcome-based financing that have garnered a lot of attention in recent years. In the SIB model, typically a government commissions services and pays for specific outcomes, a service provider delivers a social intervention, and private investors provide the upfront capital and are repaid based on the achievement of outcomes. Since 2010, over 200 SIB projects have been launched worldwide (INDIGO 2021).

In many ways the SIB model is an extension of New Public Management (NPM) logic, which surfaced in the early 1980's as a series of reforms for increasing the efficiency of public services through the application of private-sector mechanisms (Hood 1991). NPM strategies introduced 'market discipline' in an effort to save the government money while improving the quality of public services (Pollitt and Bouckaert 2004). Hands-on professional management, private-sector involvement, explicit performance standards, and greater discipline in use of resources are all features of the NPM movement (Hood 1991) that are reiterated in the SIB model. A results-driven focus on achieving target outcomes, without prescribing the process by which outcomes are achieved, is one NPM technique largely associated with the private-sector (Mongkol 2011) that has been repurposed for use in the public sector through the application of social impact bonds.

Some scholars argue that New Public Governance (Osborne 2006), "with its emphasis on the role of multiple stakeholders and recognition that social outcomes are emergent properties of complex systems" is a more fitting framework for SIBs (Fox and Morris 2019). Whilst this is a useful framework to consider,

particularly for analyzing the complex relationships inherent to SIB projects, elements of SIB design and implementation more closely align with the NPM logic (Dayson, Fraser, and Lowe 2020). In particular, private sector mechanisms of management and measurement are heavily emphasized in the SIB approach (Tan et al. 2015). In addition, the logic put forward to justify SIBs closely relates to NPM ideals. One of the overarching narratives for using the SIB model at its outset was that private sector management techniques, explicit outcome measures, and more efficient use of resources (leading to cost-savings) would improve public service design and delivery (Fraser et al. 2016). Many criticisms of the model also stem from an NPM lens, arguing that SIBs "financialize" social services by focusing on cost-reduction in an era of austerity and turning service users into potential sources of private sector revenue (Dowling 2017; Sinclair, McHugh, and Roy 2019; Tse and Warner 2018).

Like many of its NPM predecessors, the SIB model is the subject of heated debate, with critics claiming that the approach introduces profitization values, encourages perverse incentives, and fails to improve service quality (Warner 2013; Dowling and Harvie 2014; Lake 2015; Roy, McHugh, and Sinclair 2018). As these projects are based on paying for outcomes, they are often paired with the use of black-box approaches, where the process of delivering services is not prescribed nor monitored. The term 'black-box' is used as a metaphor to describe an approach where the process and characteristics of an intervention cannot be measured or viewed (Pearson et al. 2001). These types of interventions largely encourage focusing on outcomes over processes, and whist this may enable flexibility to better achieve outcomes, it may also open the door for less than standard service provision (Carter and Whitworth 2017).

Particularly when dealing with external service providers, and private investors in the case of SIBs, the pairing of outcome payments with a lack of process monitoring leads to concerns about creating disincentives in service delivery where outcomes are achieved at the expense of delivering meaningful, quality services. It is notoriously difficult to choose the right outcome to measure and target for complex public service projects (Smith 1995; Thiel and Leeuw 2002). If done incorrectly, "attention may be focused

on achieving targets that have little connection to the desired changes" by government (Boyne and Law 2005), meaning that services users of social programs may not necessarily be better off, and resources may be wasted. Payments based on outcomes can lead to ticking the box of the outcome being measured and disregarding quality of service delivery (McHugh et al. 2013). Furthermore, outcome payments could create incentives to game the contract, particularly if providers are not held accountable to standard provision of high-quality services to each client. Cherry-picking, serving the easiest segments of the target population, is one such mechanism for gaming outcome payments (Fox and Albertson 2012). 'Creaming and parking' is another: taking a list of all the referrals and parking (i.e. ignoring) harder to serve clients and creaming (i.e. serving) easier to serve clients (Carter and Whitworth 2015; Roy, McHugh, and Sinclair 2018). Concerns over outcome gaming stem from long-standing claims that a focus on measurable performance in delivery of public services could lead to corrupt practices or distort provider behavior (Barberis 1998; Doig 1997; Minogue 2001).

To determine whether these concerns manifest in reality, previous scholars have looked at the effect of the SIB structure on providers and front-line staff and found some evidence of outcome gaming. In their work analyzing SIBs launched in the UK and conducting in-depth interviews with project stakeholders, Fraser et.al (2018) found some evidence of provider staff 'parking' difficult cases to focus staff time on clients more likely to generate outcome payments. Other scholars echo these findings and have found that the SIB model yielded distorted provider behavior, including focusing staff time on clients who are easier to help (Ronicle and Smith 2020; Lowe et al. 2019; Edmiston and Nicholls 2018; R. Griffiths, Thomas, and Pemberton 2016; Joy and Shields 2013). Perverse behaviors, particularly 'creaming and parking', are arguably enabled by the use of black-box approaches, where providers (and investors) are not held accountable to a certain process for service delivery. This raises the question of what the effect of using a SIB model would be with a high-fidelity intervention; i.e. an intervention where providers are held to specific standards and processes for service delivery across the entire target population. Would using a

high-fidelity intervention insulate provider and front-line staff from potential perverse incentives of the SIB structure?

In this paper, the term "high-fidelity intervention" is used to refer to an intervention that is *intended* to be implemented with fidelity, including many evidence-based programs that are highly prescribed and closely monitored. Implementation fidelity is often measured across five elements: intervention adherence (intervention is being delivered as it was written), dosage or exposure (amount of intervention received by participants), quality of delivery (manner that the provider delivers an intervention), participant responsiveness (how engaged participants are), and program differentiation (identifying essential program components) (Carroll et al. 2007; Van Urk 2016). Typically, evidence-based programs are referred to as high-fidelity because it is assumed that these interventions are "implemented in full accordance with [their] published details" as they are intended to be (Carroll et al. 2007). There are longstanding debates in the implementation science literature on the benefits of program fidelity versus local adaptation (Blakely et al. 1987). Adaptation refers to the scope for program modification depending on the local context or to accommodate the needs of a particular user group (Castro, Barrera, and Martinez 2004). It is "widely recognized" that program adaptation occurs regularly to "fit the needs of different communities in different contexts" (Van Urk 2016). Allowing for the adaptation of different intervention components, particularly for complex interventions, can enable programs to be tailored to local conditions and improve program effectiveness (Hawe, Shiell, and Riley 2004). However, others argue that strict implementation fidelity is crucial to the success of evidence-based interventions across contexts (Breitenstein et al. 2010). The debate on fidelity versus adaptation in the implementation science literature in many ways mirrors the potential tradeoffs between using black-box and high-fidelity interventions with the SIB model.

In particular, one potential effect of using a high-fidelity intervention with the SIB model may be diminishing flexibility afforded to service providers. Unlocking flexible ways of working is a common promise of the SIB model, and paired with the emphasis on evidence-based practices represents one of the

model's most prevalent paradoxes (Maier, Barbetta, and Godina 2018). Positioning SIBs as a model that emphasizes evidence is argued by some as a means to depoliticize SIBs in the face of "ethical criticisms" around financializing vulnerable populations (Fraser et al. 2020). Simultaneously, arguments that SIBs enable flexible service provision and more discretion to providers adds to the model's "strategic ambiguity," yielding claims that SIBs can enable public sector innovation (Tan et al. 2019). Previous projects have dealt with this paradox of flexible yet evidence-based services in a number of ways, most notably by taking on flexible definitions of evidence-based services or acknowledging that using an evidence-based practice may reduce the flexibility of the intervention (Maier, Barbetta, and Godina 2018). The tension and potential trade-offs between evidence-based interventions and flexible interventions that can be adapted at the discretion of service providers is one that could be further unpacked in empirical analysis of SIB projects. This paper aims to contribute to this debate by investigating the impacts on providers and front-line staff under the SIB model when an evidence-based, high-fidelity intervention is used.

In this paper, I investigate the effect of pairing a high-fidelity intervention with the SIB model on the providers and front-line staff tasked with service delivery. Based on the work of previous scholars, I lay out four categories through which the SIB model may affect front-line staff: performance management, support networks, flexibility, and incentives. Through an in-depth case study of a UK-based SIB in Essex, I use an implementation science lens to determine how changing an element of the SIB context (i.e. using a high-fidelity intervention) affects stakeholders at the provider level (Bauer et al. 2015). In the case of the Essex SIB, I find using a high-fidelity intervention with the SIB model does protect from incentives to game outcome payments, but also restricts the ability of the provider to innovate and deliver flexible, context-driven services. Using a high-fidelity intervention does not mitigate against pressure felt by the provider to meet output targets, which could potentially create other disincentives in service delivery. I conclude with a discussion on the promises and drawbacks of using the SIB approach in various contexts and suggest avenues for future research.

Conceptual Framework: SIB effect on front-line staff

There are many ways that the SIB model may affect providers and front-line staff, regardless of the intervention type. Drawing on findings from the literature, articulated in the *Literature Review* chapter[12], the ways that the SIB model impacts providers and front-line staff largely fall under four key themes, summarized below, including performance management, support networks, flexibility, and incentives.

When using a SIB, the ***performance management*** requirements on providers and front-line staff are generally more intensive than in traditional methods of service delivery (Fraser et al. 2018). Because payments are based on outcomes achieved, outcomes need to be rigorously tracked and monitored throughout the entire course of SIB projects. A more robust performance management system can lead providers to increase their use of performance data and have an aligned focus on outcomes, and in some cases may lead to improved practice (Edmiston and Nicholls 2018; Lowe and Wilson 2017). Some providers have reported that this increased performance management has allowed the use of data to review practice and increase their responsiveness to project challenges, and served as one of the main benefits of the SIB model (Disley et al. 2015). However, front-line staff may feel burdened or pressured by the administrative and performance management requirements of the SIB (Disley et al. 2011; Roy, McHugh, and Sinclair 2018; Ronicle and Smith 2020). The time-intensiveness of data collection and reporting may detract from front-line staff's face time with service users (Edmiston and Nicholls 2018; Lowe and Wilson 2017); Lowe et al. (2019) found that the "onerous data driven nature" of the SIB reduced the time that front-line staff spent developing relationships with their clients. Furthermore, performance management requirements may also manifest in pressure on the provider to meet payable outcome goals and/or output requirements (such as the number of clients served, or number of cases closed) (Fraser et al. 2018; Arena

[12] For the purposes of the thesis submission, I reference an earlier chapter here. As a standalone article for publication, I plan to cite the thesis instead.

et al. 2016). Increased data management and monitoring expectations can encourage perverse behaviors by providers, related to meeting data requirements or performance targets (Lowe and Wilson 2017).

The ***support networks*** available to providers and front-line staff are different in the SIB model. Typically, SIB contracts are locked in for a number of years - offering front-line staff a degree of job security that they may not have otherwise (Fraser et al. 2018). Because the funding originates from an investor, they may also have additional and discretionary resources to use for service provision that they may not have in other methods of service delivery (Griffiths and Meinicke 2014). In addition, the collaborative nature of the model has the potential to improve the flow of knowledge to and from the provider to other project partners, in order to improve service delivery (Maier, Barbetta, and Godina 2018). On the other hand, providers and front-line staff also face additional lines of accountability - to the service user client, to their direct supervisors, to investors, and to the government partner (McHugh et al. 2013). Communication between all of the stakeholders involved in the SIB is critical, and misalignment over goals or service delivery processes could lead to inter-organizational conflict (Warner 2015; Fraser et al. 2020; Ronicle and Smith 2020). Additional project partners, including the presence of a private investor who may not have relevant programmatic experience, could lead to provider frustrations with inconsistent advice on service direction, particularly as power imbalances are likely in provider relationships with investors (Morley 2019). Providers may also express resistance to partner efforts to shift their organizational culture "to become more business-like," resulting in a culture clash between provider and investor stakeholders (Lowe et al. 2019).

When delivering services as part of SIB projects, providers and front-line staff may have increased ***flexibility*** in the processes and interventions they use, as resources may not be tied to the same regulations as public funds (Griffiths and Meinicke 2014; Carter et al. 2018). The SIB's outcome focus could enable a greater personalization of services (Leventhal 2012; Jackson 2013) and allow providers greater latitude to determine how to obtain results (Fox and Albertson 2012). Providers may experience the ability to tailor services in a way previously constrained, and incorporate innovative service approaches in their day to day

interactions with clients (Fraser et al. 2018). The SIB model could enable service providers to "reconfigure[e] the structure of service delivery" and leave them free to design new initiatives or approaches through which to achieve outcomes (Arena et al. 2016). Some providers have reported that the focus on outcomes paired with the input of investors have "motivated them to adapt" when delivering SIB-funded interventions (Ronicle and Smith 2020). In addition to flexibility in direct service provision, providers may also have more discretion on resource allocation and prioritization (Fox and Albertson 2012).

The SIB structure also brings a new relationship with ***incentives*** to providers and front-line staff. The outcomes contract may shift provider behavior to focus on improving outcomes (Battye 2015), aligning outcome goals and yielding a consistent focus on achieving those goals (Liebman 2011; Fraser et al. 2016). Instead of focusing on inputs and outputs, common practice among providers, they instead focus on "quantifiable social outcomes that dictate the rate and amount of payment" they receive (Edmiston and Nicholls 2018). Outcome-based contracts could create disincentives, however, because payment is only made if particular outcomes are demonstrated (Rees, Whitworth, and Carter 2014). Providers may be put under increased pressure to meet outcomes or outputs, which could manifest in a change in provider behavior (Fraser et al. 2018). For example, the outcome-payment nature of the SIB could lead to a narrow scope of services provided, focusing only on achieving payable outcomes while neglecting more holistic, but difficult to measure, indicators of impact (McHugh et al. 2013; Tse and Warner 2018). The SIB model could also lead to the narrowing of the target population, to focus only on those for whom data can be reliably collected, or those who are more easily reached (Lowe et al. 2019). For example, a SIB launched in North England to address social determinants of health only aimed to serve a subset of the population that General Practitioners thought would benefit due to provider restrictions set out by the SIB (Lowe et al. 2019). In the Youth Engagement Fund SIBs, launched across the UK to support disadvantaged young people, evaluators found that providers did not necessarily support the most disadvantaged young people in target areas (Ronicle and Smith 2020). The outcome-based contract could also lead to explicit 'gaming' of the contract, by focusing on easier to serve clients at the expense of cases that may be deemed as more

challenging (Joy and Shields 2013; Sinclair et al. 2014; Carter and Whitworth 2015; Fraser et al. 2018; Ronicle and Smith 2020). For example, Lowe et al. (2019) report that for providers, the process of meeting payable outcomes "distorted practice" and "encouraged creaming of clients - working with those who are easiest to help."

Table 9 below summarizes these themes, suggesting the SIB model can potentially have a mixed effect on front-line staff. The model can provide front-line staff with flexibility to individualize services and a more robust support network of additional resources, but it could also create rigid performance management expectations and introduce perverse incentives that could negatively affect service users.

Table 9: How the SIB structure might affect front-line staff

Main themes	Subthemes
Performance management	+ Increased use of performance data to respond to challenges -- Administrative burden (impacts time with service users) ? Pressure to meet outcomes and/or outputs
Support networks	+ Heightened job security + Availability of additional resources + Collaborative knowledge flows -- Culture clashes with investors -- Complex systems of managerial oversight
Flexibility	+ Ability to tailor or individualize services + Incorporation of innovative service approaches + Provider discretion over service delivery & resource allocation
Incentives	+ Outcome goal alignment -- Outcome gaming incentives -- Narrowed scope of services ? Consequences for missed targets ? Financial or other rewards

Table Key:
+: Positive effect
--: Negative effect
?: Unclear effect

In the sections that follow, I investigate the effect of the SIB model on front-line staff when the project uses a high-fidelity intervention. Are front-line staff protected from incentives to 'game' outcome payments? What happens to their flexibility to innovate, which is consistently cited as a major advantage to using the SIB model? Are they still subject to strict performance management expectations, or is this lessened

because of the highly prescribed nature of the intervention they are delivering? These questions are important because they add nuance to the debate around the use of social impact bonds by looking at the effects of the model under different contexts. It will help shed light on what, if any, effect the SIB model has on providers when using a high-fidelity intervention. This is particularly relevant, given the increasing use of SIBs (using both black-box and high-fidelity interventions) around the world.

Methods: Case study of a SIB with a high-fidelity intervention

To determine how the SIB structure affects providers and front-line staff in the context of a high-fidelity intervention, I conducted an in-depth case study of a SIB launched in Essex to fund multi-systemic therapy (MST), a high-fidelity license-based intervention that provides therapy to children at-risk of entering the care system. In 2012, policymakers decided to test the SIB model in Essex to tackle a growing crisis: record numbers of children entering statutory (out-of-home) care. Ofsted, the UK government's inspector of children's services, had recently rated Essex County Council 'inadequate' – the worst possible rating indicating widespread and serious failures within the council's children's social care system (Ofsted 2010). To tackle this challenge, the council began to implement a series of reforms, including the launch of a SIB to provide MST to children at-risk of entering care ('on the edge of care'). MST provides therapists to work with families in their homes for up to five months, supporting parents in managing young person behavior and working with the young person to promote positive activities and improve family relationships (MST UK 2018). The MST model is considered high-fidelity because it requires therapists to follow a set of specific processes when working with each family, and every week therapists discuss their work with an accredited MST employee to ensure the model's processes are being followed and standards of service provision are met. Any changes to the service delivery model are typically not allowed.

Key stakeholders of the Essex SIB project included:

- Essex County Council or the 'council': the government partner responsible for children's social care in Essex, whom agreed to make payments on outcomes achieved.
- The provider: a large charity operating throughout the UK, responsible for delivering the MST service in Essex.
    - The front-line staff: MST therapists, supervisors, and program manager who were employees of the provider, but were hired specifically to deliver MST in Essex as part of the SIB project.
- The SIB Board: A special purpose vehicle made up of investors and an advisor, responsible for upfront funding of the SIB project.
- The intermediary: An organization specializing in SIBs, responsible for managing the project and serving as the link between all project stakeholders.

In the SIB project, investors gave upfront financing to a large children's charity (the provider) to deliver MST in Essex. The investors also established a 'SIB Board,' which oversaw the project and played an active project management role throughout service delivery. The council agreed to pay investors back based on the number of days the children served by MST subsequently spent out of care (compared to a historical baseline). The provider set up two MST teams in Essex, each consisting of up to five therapists and a supervisor. Both project teams were overseen by a program manager. Therapists, supervisors, and the program manager were hired explicitly to deliver MST in Essex as part of the SIB project.

After delivering MST to over 380 young people in Essex over five and a half years, the SIB came to an end in December 2018 and MST services were stopped. Approximately 19% of the young people who received

MST entered care after the services ended[13] (Drew et al. 2019), significantly less than their projected baseline, and the SIB successfully achieved its outcome targets. Other sites around the UK have performed better than MST in Essex, for example in a randomized control trial of nine MST sites throughout the UK, Fonagy et al. (2018) report 13% of young people receiving MST had an out-of-home placement in the 18 months after services.[14] Despite this, the performance of MST in Essex was much better than what project partners projected and the SIB far surpassed its outcome targets; investors were repaid in full plus a return. The low baseline can be (at least partially) explained by the context in Essex at the time the SIB was launched; in 2011 Essex had the third highest level of children in care across all UK local authorities, and 56% of these children were deemed at risk of experiencing abuse or neglect (GOV.UK 2011).

Using Essex as a case study allows for a nuanced, detail-rich investigation of the effect of the SIB model on providers and front-line staff under the context of a high-fidelity intervention. Fox and Morris (2019) have highlighted the need for more high-quality analysis of SIB projects, and Lowe et al. (2019) note that most studies to date do not "examin[e] the detailed complex set of interactions" between various SIB stakeholders. The case study that follows doesn't seek to determine whether MST was implemented with fidelity, but rather focuses on the impact of the SIB model on the provider and front-line staff using an intervention meant to be implemented with high-fidelity. Stemming from an implementation science lens, this investigation focuses on analyzing how the SIB was implemented and its effects on the level of the provider and front-line staff, building knowledge around what works, where, and why (Bauer et al. 2015). Advantages of using a single detailed case study include thorough analysis of complex relationships and multi-level effects (Pawson and Tilley 1997). Using a case to examine events, processes, and relationship dynamics allows for a holistic picture of a complex case (Cartwright 2017), making this design well-suited for studying the implementation of SIB projects (Fraser and Mays 2020; Lowe et al. 2019).

[13] At the time of analysis, 302/388 young people served by MST had at least one year of post-MST data. 56 of these children entered care after exiting MST services (19%).
[14] For further comparative context, MST UK (2020) reports that nearly 95% of young people throughout the UK and Ireland remain at home after receiving MST.

As part of the case study, I conducted in-depth interviews with provider and front-line staff delivering MST in Essex (n = 10), including all of the MST therapists and supervisors currently working on the project at the time of this analysis. Interviews were also conducted with front-line staff delivering services similar to MST, but through a program delivered directly by the council (n=7). The interview process began with a warm contact and continued using a snowball sampling approach (Bryman 2015) until all front-line staff working on MST had been interviewed and until the list generated by the snowball method was fully exhausted for staff working for the council. Interviews took place between July through December 2018, and were conducted individually with each respondent. Front-line respondents were all made anonymous, in an effort to protect them from revealing any sensitive information and help make them comfortable to respond freely. Interviews were conducted using a semi-structured approach (Bryman 2015), allowing informants to answer open-ended questions and to respond to follow-ups. The interview topic guide was designed using the themes set out in advance, and included questions related to comparing delivering MST with previous experiences, interactions with intermediary, council, and investor partners, discretion over approaches or ways of working, and practices and perceptions around the outcome targets.[15] At the end, respondents were asked to reflect freely on their knowledge of the SIB and how (if at all) they felt the SIB structure impacted the process of delivering MST. Interviews were coded thematically, with help from the NVivo 10 software, using an integrated approach (Weston et al. 2001) combining both inductive and deductive approaches by beginning with the initial themes set out in advance (deductive) but allowing space for new themes to emerge from the data (inductive). Themes that were explored included performance management expectations, relationship dynamics and support networks, flexibility in service delivery, the use of incentives, and pressure on outcomes or output targets.

In addition to the interviews, I observed front-line staff teams through three days of shadowing while the service was being delivered. I silently observed front-line teams conduct their daily activities serving

[15] See interview topic guides in Appendix 6.

families, using the experience for background purposes and to get a firmer understanding of the project context. Participant observation can help to provide a richer understanding of the project and enhance and inform further data collection and analysis, and is a useful tool to help the researcher critically examine any research assumptions (Musante (DeWalt) and DeWalt 2010). This research received ethics approval from the Blavatnik School of Government's Departmental Research Ethics Committee at the University of Oxford (see Appendix 2).

It is important to note that case study designs have a number of limitations. Firstly, it is difficult to generalize outside of the specific case and context when drawing conclusions from single case studies (Gerring 2004; Berg and Lune 2012), and there may be different impacts on front-line staff and providers under high-fidelity interventions in other cases. Relatedly, replicability and objectivity are difficult to ensure, as case studies are based on the subjective interpretation of the researcher and may be influenced by my perspective or past experiences (Berg and Lune 2012). In particular, my past experience with SIB projects may have influenced the way that I interacted with interviewees and my thematic interpretation of the data collected. Previously, I have worked on SIB projects, providing technical assistance to government agencies rolling out or considering SIBs. Additionally, this research was written while I was at the Government Outcomes Lab at the University of Oxford, which conducts research on SIB projects and provides technical guidance to SIB projects that are active or in development. Acknowledging this past experience with a lens of reflexivity suggests that it may have influenced the research process and outcomes reported in this paper (Berger 2015).

Furthermore, interviews may be subject to response bias, as interviewee responses could have been influenced by my presence or demeanor, or by the desire to give an 'acceptable' response as judged by either myself or other project stakeholders (Bryman 2015). In particular, respondents might have felt the need to self-censor when discussing relationship tensions or perverse incentives the SIB may have encouraged. Relatedly, participant behavior could be subject to observation bias, with front-line staff

behaving differently because they know they are being observed (Ibid). Efforts were taken to mitigate potential biases, for example by conducting interviews individually, anonymizing all front-line staff respondents and offering all other respondents to be anonymized, and triangulating data from different sources.

While subject to a number of limitations, the following case study provides rich, contextualized findings on how the SIB model affected the provider and MST front-line staff delivering services in Essex between 2012 and 2018. It can help contribute to our understanding of what negative pressures, and what value-add, the SIB model may bring to providers and front-line staff. There are important lessons we can draw from this case regarding the providers' experience delivering a SIB with a high-fidelity intervention that are relevant to both academic scholars and policymakers. By using this case study to analyze the SIB's effect on front-line staff in the context of a high-fidelity intervention, it can then be compared with previous findings on how SIBs affect front-line staff in general. This adds a degree of nuance to existing analysis of SIB implementation around how the model works, for who, and why (Bauer et al. 2015).

Findings: How the Essex SIB affected front-line staff

The following section provides detail on how the SIB in Essex affected the provider and front-line staff in regard to performance management, support networks, flexibility, and incentives.

Performance Management

Typically, in SIB projects the 'payable' outcome provides direction to the provider on the primary outcome(s) of focus, and this is followed-up on regularly through performance tracking and management. In the Essex project, partners agreed on the number of days children who received MST spent out of the care system as the payable outcome (i.e. investors were paid back by the council for each day an MST-treated child spent outside of the care system). Despite this explicit and sole outcome focus in the contract,

in the delivery of MST the outcome focus was driven by the practice and the model. Therapists all reported that deciding on outcome goals was done with each family in collaboration, "systematically with everyone around the [young] person involved" (Interview, Therapist 6) and were typically more specific than keeping a child safely at home. For example, outcomes identified by the therapists included lessening aggressive behavior, improving school attendance, or developing more positive parental-child relationships. Once a family's desired outcomes were identified, therapists tracked progress towards them every week in accordance with the MST model. While keeping children safely at home with their families was the ultimate goal, in their day to day practice the outcomes that therapists worked towards were contextually driven and identified by the families they worked with. In this way, the outcome-focus was driven by the high-fidelity nature of the model, which involved working collaboratively with families, rather than by the SIB contract.

While therapists were insulated from performance management pressures to meet the payable outcome goals, there was some pressure on provider staff to meet output targets - particularly enrollment and referral numbers. Regular contract review meetings involving the SIB board and the provider supervisors focused primarily on these output targets. The main focus was "really about where we are at in terms of numbers. The referral flow, how many kids have we served, how many [cases] are we going to close, how many are we going to open, they wanted to know those kinds of numbers (Interview, Supervisor 2)." There was a "pressure from the SIB in terms of generating referrals and opening cases," leading to stakeholders keeping a "closer eye" and asking more questions on why caseloads aren't higher than if MST were more traditionally funded" (Interview, Program Manager). Along these lines, increased responsiveness to project challenges - a potential benefit of the SIB model - revolved around these outputs of interest (referrals/enrollment/utilization rates) rather than family outcomes. For example, low referral volume was a consistent challenge, particularly at the project's outset, which the SIB Board attempted to address through marketing campaigns within Essex's social care system. In response to high staff-turnover, project partners created a new therapist-in-waiting role to be available in case of a therapist vacancy, began sending therapists to the US for MST training, and created a retention scheme that rewarded therapists for staying

through to the end of the project. While these performance improvement techniques focused on processes, attempts to improve actual service quality were much more limited.

For therapists, the administrative burden of the performance management requirements was perceived as generally similar to non-SIB funded projects (i.e. they were largely protected from additional performance management expectations). However, for supervisors and the program manager, the SIB component of the project added a layer of performance management requirements with an extra "pressure on to deliver" (Interview, Program Manager). The "regularly held contracting meetings felt extraordinarily tedious at times" (Interview, Supervisor 2), in addition to "weekly calls liaising and problem solving about any particular issues" (Interview, Program Manager). Provider leadership commented that other, much larger, projects would entail "half the work" compared to Essex, and described the SIB project as "heavily, heavily performance managed around the delivery of every aspect" (Interview, Operational Director). Supervisors expressed that the SIB stakeholders were generally very supportive and collaborative in working through issues and keeping everyone on track towards the project goals, but at times it was difficult for them to communicate "understanding of processes of how social care, social services, and social programs work" to the SIB stakeholders (Interview, Supervisor 1). The additional performance management expectations potentially "made life difficult," serving as a mechanism to allay the SIB Board's anxieties around meeting outcome goals at the expense of extra pressure on the supervisor's time (Interview, Program Manager).

Support Networks

As in other SIB projects, the SIB component of the project in Essex provided front-line staff with additional resources and a sense of job security. Because the direct financing was coming from the SIB board, rather than the local authority, the front-line staff were able to receive additional funding in an effort to improve service delivery. For example, a discretionary fund was made available for MST therapists to use to support

families by making one-off purchases aimed to alleviate sources of stress.[16] Therapists expressed appreciation for this support, noting that it "makes a massive difference - even down to engaging families" (Interview, Therapist 2) and that they were enabled to "provide more than anyone else could as a service" (Interview, Therapist 5). In addition, the external funding provided a source of job security to front-line staff, as they knew they had funding for five years. "We weren't ever worried it was going to end before that" (Interview, Therapist 5).

While therapists and supervisors generally felt very well supported by the SIB Board and other project stakeholders, complex layers of managerial oversight caused some difficulties. One therapist described the system as "unnecessarily layered" (Interview, Therapist 1), because of the number and variety of stakeholders involved in moving the project forward. Supervisors described the large "number of stakeholders who want to provide service direction" as a challenging component of the project, especially when some of these stakeholders didn't come from a social work or children services background (Interview, Supervisor 2). For example, at the start of the service, rather than communicating directly with the council about referral decisions or service direction, provider staff had to communicate through the SIB intermediary. This layered structure of operational oversight was described as very different to how the service provider typically operated in its work with government (Interview, Operational Director).

The relationship between the service provider and the SIB board was also subject to instances of culture clash, with differing value sets, ways of working, and approaches to problem-solving. Many of the service delivery changes proposed by the SIB board faced pushback from the provider partner, mostly because - as a large charity - the provider organization had institution-wide standards for dismissal policies, retention schemes, performance bonuses, and salary rates. As noted by provider leadership, the SIB Board "didn't

[16] A variation of this fund is available in other MST projects that are not SIB-funded. However, local stakeholders confirmed that this nature and size of this fund would likely not have been available in Essex if MST had not been funded by a SIB.

understand that as a big national organization" the provider was "bound by rules, regulations and procedures" that limited their ability to offer higher salaries to therapists, increase on-call allowance, or provide performance bonuses (Interview, Operational Director). In addition, some of the service delivery changes implemented by the SIB Board did not translate well to therapists on the front-line. The therapist-in-waiting role was subject to heavy resentment by those working with families, as the position was offered the same salary but entailed less work. Sending therapists to the US for MST training was also questioned on the ground, as it didn't offer therapists the same grounding context that was relevant to the UK experience (Interview, Program Manager).

Multiple lines of accountability and oversight also meant that some front-line staff felt they didn't have a clear home organization to which they belonged. "One thing that's difficult is that it kind of feels like we're these in-between people who don't really belong anywhere. We're employed by [the provider], but we're funded through [the intermediary], and we kind of work for Essex council … and that cycle has been a huge difficulty for working. We don't really have the support of a wider organization" (Interview, Therapist 3). In contrast to the similar internal services provided by the council, the MST service took a long time to integrate into the wider social care system and many felt that it was never fully embedded.

Flexibility

The flexibility to do things differently, a common promise of the SIB model, was largely lost in the case of Essex due to the high-fidelity nature of the intervention. Therapists and supervisors describe feelings of being stuck to the prescribed, evidence-based intervention. One therapist described the model as "quite limiting" in its focus on just "ticking the box" (Interview, Therapist 1). Another therapist describes an idea she had for setting up an MST parenting group once a month, as a voluntary program to help parents "develop a sense of community," and how it wasn't taken up because it was not a prescribed part of the program (Interview, Therapist 2). These reflections highlight the trade-off between sticking with what is

known to work as required with a high-fidelity intervention, and innovating to adapt to particular context-specific challenges and issues as enabled by a black-box approach. Other services "follow some really great ideas that can be adapted and changed and moved," but in this case therapists and supervisors were beholden to the MST model - making it sometimes challenging (Interview, Supervisor 2).

Over the course of the project, service changes that were not backed by existing research faced pushback from MST leadership; despite the availability of extra resources to trial process improvements, an inflexible system prevented or significantly delayed these attempts. For example, early in the project the issue of high staff turnover was preventing the program from operating at capacity and serving as many families as possible. SIB stakeholders aimed to create the therapist-in-waiting role to have a replacement therapist ready in the event of staff turnover, and although eventually implemented, this idea was subject to significant delay because it is not part of the specified high-fidelity intervention. One investor described this process as "tortuous" and "counter-cultural" to everyone involved (Interview, Investor 2). Members of the local council also expressed their frustration with the required specificity of the model, describing it as "a rigid dinosaur" that they were unable to shift (Interview, Essex County Council Project Lead). The council wished to adjust the target population, hoping that MST would be able to assist with Essex's high priority populations, including teenagers on the autistic spectrum, but the model would not allow for it. In sum, the high-fidelity nature of the model limited discretion of providers, the council, and SIB Board members over service delivery and resource allocation.

Incentives

Most of the therapists knew relatively little about the specifics of the way the SIB project was financed, protecting them from financial incentives or pressure to 'game' the contract by focusing on families who may more easily generate outcome payments. One therapist describes the front-line's lack of understanding around "the nitty gritty of how the funding actually works and what takes place between the [investor

partners] and Essex Council," commenting that "it goes on behind the scenes" (Interview, Therapist 6). Another therapist commented that the way the project was funded hasn't made a difference to their daily life working with families, "apart from it wouldn't have happened otherwise" (Interview, Therapist 1). Unlike in other SIB projects, front-line staff were kept separate from discussions around achieving payable outcomes, mitigating the potential for the outcomes-based contract to serve as a disincentive to try and game payments.

Supervisors, on the other hand, were much more aware of the project's payable outcomes and the way the SIB funding worked. Despite this, the supervisors and program manager were driven instead by pressure to meet output targets (referral and enrollment numbers) through rigid performance management expectations. None of them reported on any explicit pressure to meet payable outcomes, with one supervisor commenting "I've never felt any specific pressure around those [payable outcome] targets… but I think they're targets we're aware of and it fits with the MST work anyway... we do want to keep people at home… so really the drive for those targets for me has always been more directly from MST in terms of having a successful model rather than about financial pressure" (Interview, Supervisor 1). While supervisors were aware of the payable outcome targets, they were not financially incentivized or otherwise pressured to meet them, and instead focused on them insomuch as they reflected the goals of the overall project and MST model. However, there was a "pressure from the SIB in terms of generating referrals and opening cases" that fell on the service provider supervisors (Interview, Program Manager). This pressure on service volume through output targets, applied through the heavy performance management requirements described previously, persisted throughout the delivery of services.

As in other SIB projects, partners did use incentives, primarily to improve staff retention and workplace morale. A performance bonus, seasonal celebratory meals, and a bonus for therapists who remained through the end of the project are some examples of the use of rewards and incentives. However, it is unclear whether these practices had an impact on service delivery. Outcome-goal alignment amongst front-line staff

seemed to have been achieved through the nature of the selected intervention, rather than the SIB funding mechanism. One therapist describes the use of incentives as valuable to her on a personal level, but "in regard to MST as a model," is "not entirely sure" it led to anything different than if the project had been funded by the local council (Interview, Therapist 5). Because front-line staff and supervisors are beholden to the processes laid out by the MST model, they had limited avenues to change their approach to service delivery. This limited the scope for incentives other than for process-related measures, including staff retention and adhering to the MST model.

Discussion: When to use the SIB model?

In the case of Essex, using a high-fidelity intervention with the SIB model protected providers from incentives to game outcome payments and pressure on front-line staff to meet payable outcomes, but it also restricted the ability of the provider to innovate and deliver flexible, context-driven services. Using a high-fidelity intervention did not mitigate the pressure felt by the provider to meet output targets, such as referral and enrollment numbers, or the administrative burden brought by strict performance management expectations. Using the SIB model (in this case under a high-fidelity intervention and in other contexts) led to complex systems of managerial oversight, culture clashes between the investor and provide partners, the use of financial incentives, additional resources to use for service delivery, and feelings of job security amongst front-line staff. Table 10 summarizes which of the SIB effects were mitigated and which remained in Essex under a high-fidelity intervention.

Table 10: SIB effects on front-line staff in Essex

	SIB effects that were mitigated	SIB effects that remained
Performance management	+ Increased use of performance data to respond to service quality challenges ? Pressure to meet outcomes	+ Increased use of performance data to respond to operational challenges ? Pressure to meet outputs -- Administrative burden
Support networks	+ Collaborative knowledge flows	+ Heightened job security + Availability of additional resources -- Culture clashes with investors -- Complex systems of managerial oversight
Flexibility	+ Ability to tailor or individualize services + Incorporation of innovative service approaches + Provider discretion over service delivery & resource allocation	
Incentives	-- Outcome gaming incentives	+ Outcome goal alignment -- Narrowed scope of services ? Consequences for missed targets ? Financial or other rewards

Table Key:
+: Positive effect
--: Negative effect
?: Unclear effect

The findings from the Essex case study suggest that there are both advantages and disadvantages to using a high-fidelity intervention, as opposed to a black-box approach, with the SIB model. Using a high-fidelity intervention with the SIB model can limit the provider's ability to try new approaches and flexibly spend resources, mitigating a commonly cited promise of the SIB model. Using a well-defined intervention shifts the focus back on processes, rather than outcomes, and diminishes any benefits of flexibility for service providers (Arena et al. 2016). These findings resonate with arguments from Maier, Barbetta, and Godina (2018), on the paradoxical nature of SIBs promising to be both evidence-based and flexible. In the case of Essex, the promise of flexibility to providers was largely diminished by the selection of an evidence-based, high fidelity intervention. On the other hand, a high-fidelity intervention can ensure minimum standards of service delivery are provided to each service user and help protect against explicit outcome gaming. In

Essex, unlike in other SIB projects, front-line staff were kept separate from discussions around achieving payable outcomes and instead were focused solely on adhering to the therapeutic model set out by MST. This arguably helped mitigate against incentives for provider staff to game outcome payments, for example by focusing on a subset of service users who were more easily kept out of care (i.e. achieving the payable outcome) at the expense of other users who maybe were more likely to enter care (i.e. not achieving the payable outcome). The potential for gaming is mitigated in this case because the processes for serving families enrolled in MST are standardized.

While a high-fidelity intervention can help protect against explicit outcome gaming, there are other ways the SIB model can create disincentives in service delivery that are not protected by the high-fidelity nature of the intervention. Pressure on provider staff to meet certain output targets, such as the number of referrals to the MST program or the number of families enrolled, could create a harmful disincentive in service delivery. Particularly, for example, if referrals to the SIB intervention are made regardless of the program's relative success to other local alternatives or a particular family's goodness of fit to the SIB intervention as determined by social workers. In the case of Essex, supervisors generally described the extra pressure to meet outputs as an overall value-add which kept them on their toes, but it is possible that the pressure to meet referral targets had negative implications around ensuring families were matched to the best available service to meet their particular needs. In other projects, there has been evidence of service users being enrolled in programs because they were deemed easier to help or were more easily accessible (Lowe et al. 2019; Ronicle and Smith 2020). In Essex, however, there was a different kind of pressure: to generate as many referrals as possible, particularly at the start of the program. This poses a distinctive problem, whilst not explicit outcome gaming, as this pressure could result in families being referred to MST that may have been better served elsewhere.

It is worth discussing how a high-fidelity SIB project might compare to a high-fidelity intervention that is traditionally funded without an outcomes contract. Given that one of the primary advantages of using a

high-fidelity intervention with a SIB, protecting against explicit outcome gaming, is rendered irrelevant in traditional methods of financing (i.e. without an outcomes contract), and there is still potential for the SIB to create other service delivery disincentives, it is unclear what value a SIB model adds in the case of a high-fidelity intervention. If flexibility is limited, as was the case in Essex, why use a SIB model to fund a high-fidelity intervention as opposed to using more traditional funding mechanisms? Why should the public sector pay a return, plus additional transaction costs, through a SIB for an intervention that is evidence-based and/or meant to be implemented with high-fidelity (Fraser et al. 2020)? Evidence from Essex suggests that the SIB model may have brought increased accountability and performance management expectations than would otherwise occur if MST was traditionally funded. This non-financial support for the management and delivery of services has been argued as a benefit in other SIB projects as well (Fraser et al. 2020). It is debatable, however, if this increased performance management scrutiny is worth the large transaction costs typical of SIB projects (Warner 2013; Giacomantonio 2017).

So, when does it make sense to use a SIB model, and for what type of intervention? Using a SIB with a black-box approach can be problematic, as the pairing of outcome payments with a lack of process monitoring raises concerns around the potential of gaming outcomes contracts. Using a SIB with a high-fidelity intervention, on the other hand, may not bring net benefits compared to simply funding the high-fidelity intervention using more traditional methods. However, there does not have to be a dichotomy between "black-box" interventions with potentially gameable outcome payments, or high-fidelity interventions that are highly prescribed and may have already been proven to work. So far, this logic has arguably "dictate[d] that only what is profitable will be funded" (Dowling 2017) through the SIB model, and can "narrow the scope of social investment to low-cost programs with short-term returns" (Tse and Warner 2018). Broadening the focus, Maier, Barbetta, and Godina (2018) argue that perhaps it is the collaborative element of SIBs, rather than the paradox of evidence-based flexibility or the particular funding structure, that should be maintained in order to "improve knowledge flows" about promising and proven social programs.

Further research might explore different variations of SIBs, both in funding structure and intervention type, that might seek to better take advantage of the collaborative knowledge flows SIBs might have to offer. In particular, research might seek to understand whether a variation of the SIB model is compatible with an "open-box" intervention (Ssengooba, McPake, and Palmer 2012), where interventions are not highly prescribed but are transparent and subject to minimum standards of quality. This type of open-box intervention could help to protect against explicit contract gaming, while still allowing for the flexibility promised by the SIB model. These "open-box" interventions have parallels to debates around model fidelity, in particular form versus functional fidelity (Van Urk 2016). Form-based fidelity is a stricter interpretation of fidelity and emphasizes the implementation of specific program activities, while function-based fidelity allows for adaptation of the program activities depending on the local context whilst maintaining a standardized theory of change (Van Urk 2016; Hawe, Shiell, and Riley 2004). In complex interventions, such as SIBs, standardizing the general process of the intervention rather than the specific activities could allow for the intervention to be tailored to the local context and greater service provider discretion, while still maintaining minimum standards of service quality for all intended beneficiaries (Hawe, Shiell, and Riley 2004).

As covered in the *Methods* section of this paper, the design and qualitative methods employed in this analysis are subject to a number of limitations. Conclusions from the Essex case study must be used cautiously, as it is difficult to generalize outside of case studies particularly when only analyzing a single one (Gerring 2004). Response bias and observation bias may have influenced the results, as well as my own bias in the interpretation of the data collected (Bryman 2015; Berg and Lune 2012). In other cases of SIB projects using a high-fidelity intervention, there may have been different impacts on providers and front-line staff. For example, providers in other projects may not have experienced limited flexibility or pressures to meet referral targets. While this analysis adds to the body of knowledge on SIBs by

providing context-rich detail on the complex interactions of a particular case, additional research is needed to further our understanding of how SIBs work under different contexts and for different actors.

Conclusion:

The trend of government agencies paying for social service outcomes, rather than services delivered, is often paired with the use of black-box interventions. There is some evidence (Fraser et al. 2018; Lowe et al. 2019; Ronicle and Smith 2020) and much concern that this pairing could lead to gaming outcome contracts, creating disincentives in service delivery where outcomes are achieved at the expense of delivering meaningful, quality services across the entire target population of intended recipients. These concerns echo critiques of New Public Management strategies more broadly, and in particular that a focus on measurable performance in delivery of public services could lead to corrupt practices or distort provider behavior (Barberis 1998; Doig 1997; Minogue 2001). For SIBs, these concerns are particularly salient because these projects involve private investors providing upfront capital and gaining a potential return from government agencies based on achieving outcome targets.

In this paper I have investigated whether using a high-fidelity intervention with the SIB model can insulate provider and front-line staff from potential negative effects of the SIB model, including disincentives for gaming outcome payments. Drawing on principles from implementation science, the analysis focuses on how the SIB was implemented and its effects on the level of the provider and front-line staff, building knowledge around what works, where, and why (Bauer et al. 2015). While using a high-fidelity intervention with the SIB model mitigated incentives to game outcome payments in Essex, it also restricted the ability of the provider to innovate and deliver flexible, context-driven services. Using a high-fidelity intervention did not protect providers from pressure to meet output targets, which could potentially create harmful pressures around generating referrals and enrolling families regardless of fit.

Through this analysis, I have made both conceptual and empirical contributions to the debate around using outcome-based financing to fund social services provided through government agencies. I have laid out the conceptual arguments for how the SIB model could affect provider and front-line staff, and extended this conversation to include the differences between types of intervention (i.e. black-box and high-fidelity). Using an in-depth case study, I explored whether the SIB launched in Essex, with a high-fidelity intervention, insulated front-line staff from potential negative effects of the model. While the empirical analysis is based on a single case, an important limitation of this work, the findings add to a growing body of knowledge regarding the use of social impact bonds and how they interact with processes of service delivery under various contexts.

The findings from the Essex case study, in combination with previous work, suggest that further research is needed to compare SIBs under different contexts and to non-SIB alternatives. Scholars have found that using a SIB with a black-box approach can be problematic, as the pairing of outcome payments with a lack of process monitoring raises concerns around the potential of gaming outcomes contracts. The Essex case study suggests that using a SIB with a high-fidelity intervention limits flexibility and still leaves potential for disincentives in service delivery despite mitigating against explicit outcome gaming. As such, it remains unclear when, if ever, the SIB model should be used as an alternative to more traditional methods of financing social services.

Article Three - Have We 'Stretched' Social Impact Bonds Too Far? An Empirical Analysis of SIB Design in Practice

Abstract:

When the social impact bond (SIB) model was first introduced in 2010, there were many claims about how these projects could transform the way that public services are delivered. The potential of the SIB model to shift the focus of social service delivery to achieving impact, whilst simultaneously transferring financial risk of project failure away from government, became a dominant narrative used as rationale to pursue SIB projects. Over 10 years and 200 SIB projects later, are these projects - in practice - much different from business as usual? This paper investigates how SIB projects have been designed in relation to two of the model's primary original intentions: shifting the focus of public service delivery to achieving impact and transferring risk away from the government partner to external investors. Qualitative content analysis of SIB projects launched until 2020 in the US and UK (n=114) is used to plot and analyze how the design of SIB projects vary across two key dimensions: degree to which payment is linked to impact, and the independent, at-risk nature of the working capital provided. This paper represents the first systematic attempt to map variation in design of SIBs against the espoused model, and allows us to draw conclusions on the model's ability to live up to its original promises. We find that SIB design in practice has deviated from the model's original intent in several important ways. In the UK, outcomes that projects pay out for are generally not subject to rigorous validation methods, calling in to question whether payments are meaningfully linked to impact. In both the US and the UK, upfront capital is generally independent and at-risk, but risk mitigation strategies in the repayment structures may limit the intended transfer of risk to external investors. Therefore, the SIB model has indeed "stretched" from what makes it a distinctive, innovative approach in some cases, potentially undermining its potency as a tool for service delivery improvement.

Introduction: Promises of the SIB model

OECD[17] countries have a longstanding history of private sector involvement in public services, stemming from the idea that hybrid designs in service delivery are a more effective and efficient way for the government to operate (Skelcher 2007). Hybridity in social service delivery typically refers to public partnerships with private and/or voluntary organizations, relying on shared resources and governance structures to reach a common goal (Borys and Jemison 1989). In the early 1980s, New Public Management (NPM) - a hybrid philosophy focused on applying private-sector mechanisms to increase the efficiency of public services - came to the forefront of the reform agenda (Hood 1991). NPM strategies aimed to improve the quality of public services, while simultaneously saving the government money by introducing market discipline (Pollitt and Bouckaert 2004). These NPM ideals manifested as strategies to increase market competition, rely more heavily on hands-on management, and focus on explicit measurement of performance (Ferlie et al. 1996). With these strategies came bold claims around how the public sector could be transformed for the better, operating more efficiently and effectively.

Social impact bonds (SIBs) are arguably one recent permutation of NPM ideals, with their focus on performance management and reliance on outcome-based contracting mechanisms (Warner 2013; Fraser et al. 2018). Effectively a three-way partnership between government, voluntary sector providers, and private investors, SIB projects typically involve a government agency which commissions services and pays for them based on outcomes, a service provider that delivers an intervention, and private investors who pay for the intervention upfront and are repaid based on the achievement of outcomes. At its outset, the SIB model promised to shift the focus of social service delivery from completing processes to achieving impact, transfer risk of unsuccessful programs away from government, introduce 'market discipline' and enhanced performance management techniques, and enable socially-motivated service providers to bid for

[17] OECD stands for the Organization for Economic Cooperation and Development, a group of 37 countries that are typically democratic and support free-market economies (Kenton 2020).

government contracts (Roberts 2013; Disley et al. 2015; Fraser et al. 2016). Under the SIB model, payment for services is made contingent on the achievement of measurable social outcomes. In this way, returns to investors are correlated with the achievement of impact, promising a blend of social and financial returns if projects are successful. As such, social impact bonds have been lauded as a way to incorporate private-sector ideals and strategies into public service delivery, all without risking public money. These bold promises, however, have been accompanied by many criticisms.

One of the key criticisms of SIBs is that the simplistic model of defining measurable outcomes and then paying on them fails to account for the complexity or context of challenging social problems (McHugh et al. 2013). Frequently, SIB payment mechanisms face attribution challenges, where it is difficult to determine whether or not an intervention was actually responsible for an outcome, and may lead to payment for narrow, less meaningful outcomes that do not correspond to a project's actual impact (Loder 2011; McHugh et al. 2013; Sinclair et al. 2014; Roy, McHugh, and Sinclair 2018). Barajas et al. (2014) add that the focus on these narrow outcomes could lead to the neglect of systemic issues, while distracting from "more meaningful policy reform." Others argue that the SIB model allows private investors to profit from vulnerable populations in need of support, by allowing investors to shape policy and turn service users into potential sources of revenue (Dowling 2017; Sinclair, McHugh, and Roy 2019; Tse and Warner 2018; Neyland 2018). These fears are exacerbated by arguments that the SIB model, in practice, does not actually transfer risk from government to investors (Demel 2012; McKay 2013; Fraser et al. 2016).

Despite criticisms of the model, governments around the world are experimenting with the adoption of Impact Bonds - since 2010, over 200 SIBs have been launched worldwide (INDIGO 2021). The UK has been, by number of projects, the largest adopter of the model with 89 projects launched to date and over $95m in upfront capital raised (Ibid). The central government has been a primary spear-header of the SIB model adoption, launching a series of sizable investments into SIB projects; the £80m Life Chances Fund, £30m Innovation Fund, £16m Youth Engagement Fund, and £15m Fair Chance Fund have all been

launched by the UK central government in the past ten years to spur investment in SIB projects (GO Lab 2020). The US is also a big market for SIB projects, launching 27 projects with over $193m in capital raised (INDIGO 2021). In 2014, the US government's Social Innovation Fund launched a Pay for Success program that provided over $30m in grants to develop social impact bonds across the country (CNCS 2017). In 2018, the Social Impact Partnerships to Pay for Results Act (SIPPRA) was signed into law, appropriating $100m for projects that use social impact partnerships to achieve demonstrable results (US Treasury 2018). Motivations for pursuing SIBs across these two countries vary, with the US putting forward SIBs as a way to test new, innovative programs and the UK positioning SIBs as a mechanism to fund general social services in an era of austerity (Albertson et al. 2018; Neyland 2018). Rhetoric in both countries, however, emphasized the potential of the SIB model to shift the focus of service delivery to achieving impact for beneficiaries, while protecting taxpayer money by shifting risk of project failure to external investors (Fraser et al. 2016; Albertson et al. 2018).

Whilst keeping tally of the adoption of impact bonds is relatively straightforward, assessing the degree to which the model's promises have been realized is a far more challenging question - one we turn to in this paper. Are SIBs being designed in a way that would enable the model's promises, that "investors assume the risk of a failed social intervention based on rigorous outcome evaluation" (Chiapello and Knoll 2020), to be fulfilled? Or is this an "emperor's new clothes" problem, where the SIB model is not in practice much different than status quo forms of service delivery (Heinrich 2018)? If the SIB model is designed in a way that is not much different from status quo methods of service delivery, yet still offers an opportunity for external actors to profit off of theoretical risk transfer, this could signal serious problems for policymakers as they might not obtain the promised benefits of SIBs.

Turning to literature on implementation science suggests that the SIB model could have been implemented much differently across different contexts (Hawe, Shiell, and Riley 2004). The same policy or idea can take a completely different shape depending on existing state structures, institutional settings, and social contexts

(Chiapello and Knoll 2020). In addition, the SIB model itself is "strategically ambiguous" (Tan et al. 2019), and potentially "can be bent" in various directions "depending on the preferences of the audience" (Maier, Barbetta, and Godina 2018). Subject to heated debate, SIBs can be described as a high conflict model, and policies that are both high in conflict and high in ambiguity can give way to the potential for exploitation in the way they are implemented (Matland 1995). This potential for ambiguity and exploitation highlights the importance of empirically analyzing how SIBs have been designed across contexts in relation to the model's key promises.

Much of the current discussion around the SIB model fails to acknowledge considerable variation across projects, despite the fact that SIBs "display a high level of heterogeneity" and may deviate from the typical SIB description in important ways (Chiapello and Knoll 2020). For example, a homogenizing lens is applied to SIBs by scholars who argue that the model exploits the voluntary sector to achieve profit for commercial investors, although we know SIB investors also include philanthropic foundations (Warner 2013; Fraser et al. 2016; Dowling 2017; Sinclair, McHugh, and Roy 2019). This discussion neglects the fact that some SIBs may not embody all of the characteristics for which they are criticized. On the other hand, SIBs also may not embody all of the characteristics for which they are praised, such as outcome contracts that pay for impact or repayment structures that shift risk away from the government partner (Demel 2012; McKay 2013; Fraser et al. 2016).

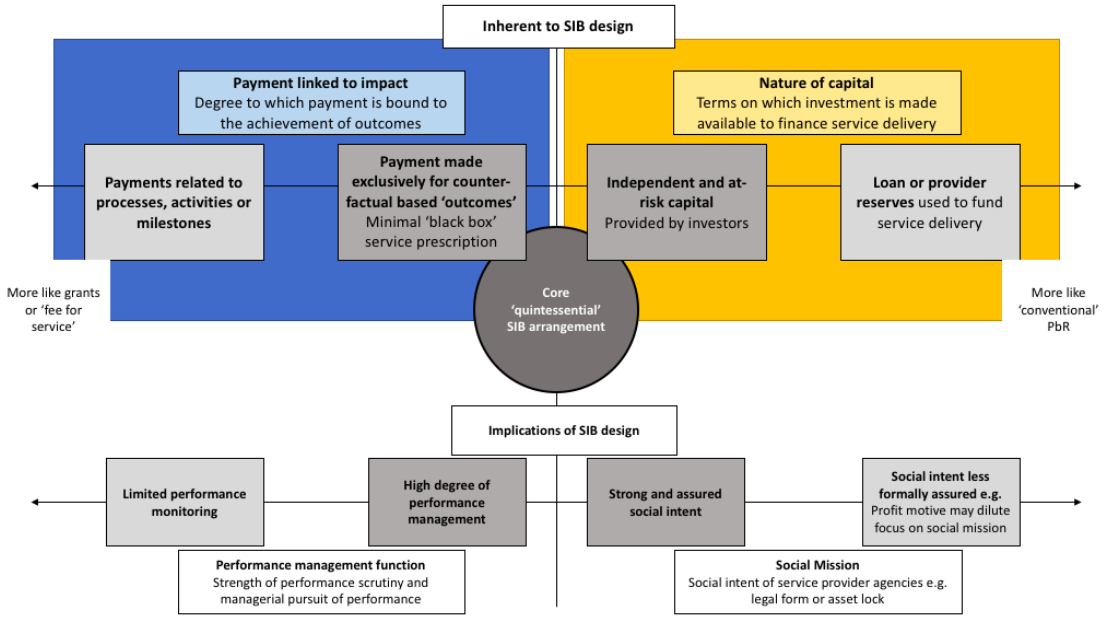
The paper that follows maps variation in the design of SIBs against the espoused model, and allows us to draw conclusions on the model's ability to live up to its original promises. We do this by investigating variation in SIB design across two major dimensions: degree to which payment is linked to impact, and the independent, at-risk nature of the upfront capital provided. First, we operationalize the two dimensions in order to create a coding frame that allows for project categorization and comparison. We then apply this coding frame to SIB projects launched from 2010-2020 in the US and the UK (n=114). After this analysis we discuss the main findings, notably that SIBs have stretched from the original intentions in important

ways: 1) a less robust outcomes framework typical of many UK SIB projects calls in to question whether these projects actually shift the focus to impact, and 2) various risk mitigation strategies, in both the US and the UK, suggest that the nature of the capital is such that risk may not be fully transferred to external investors. Finally, we conclude with implications for policymakers, and suggest avenues for further research.

Conceptual Framework: A coding frame to analyze variation in SIB design

This paper builds on a framework offered by Carter (2020), which lays out four components promised by the SIB model: payments are made for impact, up-front capital is independent (of government/service provider organizations) and at-risk, there is a high degree of performance management, and service providers have a strong and assured social intent (see Figure 5 below). Carter suggests that in practice, SIBs may have "stretched" far from these quintessential components, and hypothesizes that where SIBs fall on the scale are "likely to have very real effects on the functioning of these SIB projects and their ability to live up to the promises allied to this commissioning approach" (2020).

Figure 5: Dimensions of variation in SIBs (adapted from Carter 2020)



Interrogating Carter's original framework reveals a distinction between the dimensions that are inherent to SIB design, and dimensions that reflect implications of SIB design. The top two dimensions of the framework: 'payments linked to impact' and 'independent, at-risk capital', are the defining features of a SIB. Together, these dimensions are necessary and sufficient for a project to be characterized as a social impact bond (INDIGO 2020). The bottom two dimensions, degree of performance management and provider's social intent, are not defining features of a SIB, although they are often articulated as desired implications of the SIB design. These two dimensions are neither necessary nor sufficient for a project to be considered a SIB, but rather are implied benefits of the SIB model. For these reasons, in this paper we focus on the top two framework dimensions: payment linked to impact and nature of the capital. By doing so, we build on Carter's framework, moving beyond theory to operationalize and empirically explore variation in SIB design.

Payment made for impact

When contracting out a service, provider payments can be linked to a number of different elements: activities (what the provider does, i.e. outreach to potential service users who might benefit from a skills/training program,) outputs (what a provider delivers to service users, i.e. number of service users who completed a skills/training program), outcomes (results of service users; i.e. the number of service users employed after the program, or impact (results of service users compared to a relevant counterfactual, i.e. the difference in the number of service users employed after the program and the number of service users who would have been employed otherwise) (Gertler et al. 2016; Ebrahim 2019). In typical fee for service contracts, service providers are paid for a mix of activities and outputs, i.e. they are paid to offer a certain service, sometimes with the expectation to reach a certain number of clients, offer a certain number of courses, maintain a certain number of beds, etc. A key argument in favor of the SIB model is that switching the payment method from activities/outputs to outcomes/impact can shift provider behavior and improve social outcomes (Battye 2015; Edmiston and Nicholls 2018). By incentivizing outcome goal alignment

across participating parties, service providers are more likely to be focused on the outcomes they are trying to achieve (Fraser et al. 2016). This outcome focus, rather than a process focus, is meant to enable service providers with the flexibility to innovate and lead to greater personalization of services, thus improving service delivery (Leventhal 2012; Fox and Albertson 2012; Fraser et al. 2016). Tying SIB payments to impact, based on robust impact evaluations, helps to ensure attribution and give confidence that, through SIB projects, government agencies are paying for meaningful improvements in the lives of intended beneficiaries (Liebman 2011; Albertson et al. 2018; Chiapello and Knoll 2020).

In practice, SIB projects are likely making payments for elements across the whole spectrum: activities, outputs, outcomes, and impact. The operationalization and data analysis pioneered in this paper will show to what degree there is deviation across the spectrum of payment types for projects launched in the UK and US. Fischer and Richter (2017) argue that SIBs "tend[s] to elect a narrow category of short-term outcomes that can be monetized readily, not a full set of potential positive benefits whose social value accrues over the lifetime of program participants." Other scholars argue that SIB projects can create a narrow scope of services by focusing on outcomes that can be measured for target populations that can be easily reached (Dowling 2017; Tse and Warner 2018; Ronicle and Smith 2020). The degree to which payments are linked to impact has implications for two key SIB promises: 1) that in these projects, the focus is shifted from delivering processes to achieving tangible results for intended beneficiaries, and 2) that risk is transferred away from the public because the government agency only pays for what works.

To operationalize the degree to which 'payments are linked to impact,' it is necessary to look at two separate, but interrelated components: the payment model and validation method. The payment model is defined as the percentage of contract payments that are made based on outcomes. Outcomes are defined as results for service users achieved once the beneficiary population uses the tangible goods/service/intervention provided by the project (Gertler et al. 2016). This component alone, however, is not enough to determine whether payments are linked to impact, or in other words, if those outcomes can

be attributed to the SIB-funded intervention. The validation method sheds light on how outcomes are measured. In particular, the validation method is defined as the approach taken to assessing whether specified outcomes have been achieved. To assess the validation method, we apply the standards typical of conventional quantitative impact evaluation, as this logic of rigorous counterfactual evaluation aligns with original promises of the SIB model (Liebman 2011; Cohen 2011; Corrigan 2011; Mulgan et al. 2011; Fox and Albertson 2012; Albertson et al. 2018; Fraser et al. 2018; Chiapello and Knoll 2020).

Table 11 below offers an operationalization of both the payment model and validation method components. The payment model is scored based on the percentage of the total contract payment that is linked to outcomes, from 0% to 100%. Validation method is scored based on the rigor of the approach taken to assure outcome achievement, from 'no validation method' to the use of a randomized control trial (RCT). See Appendix 8 for term definitions used in the table below.

Table 11: Payment for Impact Operationalization

Score	0	1	2	3	4
Payment Model	0% of payment on outcomes	Less than half payment on outcomes (more than 0%)	Half or more payment on outcomes (but less than 100%)	100% of payment on outcomes	N/A
Validation Method[18]	None	Simple ex-post verification	Pre/post analysis or historical comparison	Quasi-experimental	Randomized Control Trial (RCT)

It is crucial to analyze these two components in parallel, as one without the other does not hold much meaning. For example, a contract with 100% payment on outcomes but no validation method would not provide confidence that these outcomes were actually achieved. On the other side, having the strongest

[18] If more than one validation method was used, the project was coded according to the more rigorous validation method type.

validation method possible, such as an RCT, is not relevant if all project payments are made for outputs or activities. Table 12 below offers a composite score for the payments linked to impact dimension, combining the payment model and validation method scores.

Table 12: Payment Linked to Impact Composite Score

		Payment model score			
		0	**1**	**2**	**3**
Validation method score	**0**	0	0	0	0
	1	0	1	1	2
	2	0	1	2	3
	3	0	2	3	4
	4	0	3	4	4

Independent/at-risk capital

The presence of an outside investor providing the upfront capital is a key component of what set SIBs apart from more mainstream payment by results arrangements (Cohen 2011). In principle, under the SIB model the investor takes on the risk, transferring it away from both the government and provider partners (Fox and Albertson 2011; Mulgan et al. 2011; Edmiston and Nicholls 2018; Fraser et al. 2018). As such, the SIB model represents "a contractual arrangement intended to transfer the financial risks of complex and innovative services away from government commissioners and direct service providers" (FitzGerald et al. 2019). By contrast, in fee-for-service arrangements the government agency funds upfront services and holds the risk of programs not working. In traditional payment by results projects, providers typically fund upfront

services with their own reserves or a bank loan, thus it is providers to the outcomes contract who hold the risk of the program not working.

In addition to risk-transfer, one of the key logics underpinning arguments for the SIB model is the introduction of private sector strategies (via their at-risk capital) to reform public service delivery (Fraser et al. 2016; Clist and Dercon 2014). Original SIB proponents argued that placing private sector capital at risk is the element that brings service delivery innovation; when private sector holds the risk of project failure, it can "promote more market testing" and "encourage new market entrants" in the world of public services (Fox and Albertson 2011). As a result, the degree to which upfront capital comes from a private investor and is at-risk has implications for two key SIB promises: 1) the risk of project failure is transferred away from the government and provider partners, and 2) these projects will yield innovative private sector strategies (driven by a blended profit/social motive) that make public services more effective and efficient.

To operationalize the degree to which upfront capital is independent and at-risk, we again offer two distinct but related components for analysis: the primary investor type and the repayment structure. The primary investor type is defined as the category of organizations that contribute the most money to upfront project costs; this could be the government, provider, philanthropy, social investors, or commercial investors. The repayment structure is defined as the risk-sharing split between the investors of upfront capital and the service provider. While there are many different kinds of project risk, for the purpose of this paper risk is defined as performance risk: i.e. the possibility of losing the upfront investment if project outcomes are not met. In typical payment-by-results arrangements, service providers hold all of the performance risk as they are paid based on outcomes and any loans must be paid back regardless of performance. In the textbook SIB model, an external investor is (in theory) meant to shift the risk away from the service provider by paying for upfront service delivery and getting paid back based on performance by the government outcome payor - the investor holds all of the performance risk. Project partners can also design the repayment structure so that performance risk is split between the provider and investor. In a case where the provider

primarily holds the performance risk, the provider may be required to repay the principal to investors regardless of performance, but returns to investors may be paid based on performance. In a case where investors primarily hold the performance risk, investors may be paid based on outcomes with providers paid for services, but perhaps can withhold payment to providers based on not meeting certain key performance indicators (KPIs). Risk can be shared between the provider and investor partners if both hold the outcomes contract or both are repaid based on outcomes. See Appendix 8 for term definitions used for the table below.

Table 13: Nature of the Capital Operationalization

Score	0	1	2	3	4
Primary investor type[19]	Government	Provider	Philanthropy	Social investors	Commercial investors
Repayment structure[20]	Provider holds all performance risk	Provider primarily holds performance risk	Provider and investor share performance risk	Investor primarily holds performance risk	Investor holds all performance risk

Together these two components reflect where the upfront capital comes from for a SIB project, and the degree to which it is 'at-risk'. It is important to analyze these two components in parallel, as individually they do not provide a complete picture. For example, even if the upfront capital comes from commercial investors, the repayment structure can affect the degree to which capital is at-risk. Investors may attach conditions to the upfront investment in order to lessen the risk, such as making service payments to

[19] If two separate organization types contribute equal amounts to the SIB project, the "primary" title goes to the type that stands to lose their money first in case of project failure. If it is unclear how much capital different groups contributed, projects were scored the highest of the contributing types (for example, if both philanthropy [score 2] and social investors [score 3] provided upfront investment but with no detail on what quantities, the project would score a 3 along primary investor type). In Appendix 9, a sensitivity analysis demonstrates how the results shift based on this assumption.

[20] In many cases the specific details of the repayment structure between investors and provider partners are treated as commercially sensitive and are unavailable. However, for many projects, there are indications of whether providers shared in performance risk, by either holding the outcomes contract or contributing some of the upfront capital (or both). In cases where the repayment structure was unable to be verified, projects that had evidence of provider risk share scored a 2 and projects that did not have evidence of provider risk share scored a 3. In Appendix 9, a sensitivity analysis demonstrates similar patterns for the results of only projects that have been fully verified.

providers contingent on meeting certain KPIs. Table 14 below offers a composite score for the nature of the capital, combining the primary investor and repayment structure scores.

Table 14: Nature of the Capital Composite Score

		Primary investor type score				
		0	**1**	**2**	**3**	**4**
Repayment structure score	**0**	0	0	0	0	0
	1	0	1	1	2	2
	2	0	1	2	3	3
	3	0	2	3	4	4
	4	0	3	4	4	4

Methods: Qualitative content analysis of UK and US SIBs

To investigate the variation in design of SIB projects across these two dimensions (payment linked to impact and nature of the capital) we have employed qualitative content analysis. Qualitative content analysis enables a process of coding qualitative data through systematic categorization in order to identify patterns (Hsieh and Shannon 2005). The data-driven and iterative process of qualitative content analysis allows us to investigate complex SIB projects by honing in on selected aspects that relate to the overall research question of variation in SIB design (Schreier 2014).

At the heart of qualitative content analysis is the creation of a coding frame (Schreier 2014). To create an appropriate coding frame to analyze design variation in SIB projects, we began by breaking down the two dimensions of interest - payment linked to impact and nature of the capital - in order to generate categories that might best capture these dimensions in a way that was operationalizable. We then implemented a pilot

phase to test the coding frame on a subset of the material, to ensure that the coding frame captured the elements of SIB project design that we aimed to analyze. The resulting coding frame, laid out previously in the conceptual framework, offers an operationalizable conceptualization of SIB design features that can be used to determine variation across SIB projects (and variation from the model's original promises).

We applied this coding frame to all SIBs launched in the US and UK from 2010-2020 (n=114). This list of projects was pulled from the International Network for Data on Impact and Government Outcomes (INDIGO) impact bond database, giving us access to the most comprehensive and current information on social impact bond projects (INDIGO 2021). SIB projects from the US and UK represent over 60% of all SIBs worldwide, providing a rich set of data points for analysis. The analysis also incorporates the year of SIB launch, to investigate longitudinal trends in whether SIBs have 'stretched' over time within and across the two countries. The inclusion of both countries enables a comparative lens to see how SIBs have been designed across the two countries over time. The countries share characteristics that make comparison easier, including a history of neo-liberal reforms to public service delivery and encouraging market solutions to social welfare problems (Esping-Andersen 1990). However, the US and UK also have key differences that might explain variation in trends, for example the US typically has more local government control, while the UK has shown more central political support for SIBs (Albertson et al. 2018). While this paper's analysis is not representative of all SIBs, it analyzes the evolution of SIB design in the two countries that have been leading the way in exploration of this model.

In order to populate the coding frame for SIB projects included in our sample (n=114), we conducted documentary analysis of publicly available documents, including evaluation reports, press releases and project announcements, project contracts, and government website pages. For some of the SIBs in the sample, public materials were insufficiently detailed to verify particular project details. To help address this, we relied on an advisory group, consisting of SIB investors, and intermediary, provider, and government staff (n=5) to verify project-level information. This triangulation process helped to provide

more robust project-level information whilst also allowing us an opportunity to verify data collected from documentary analysis.

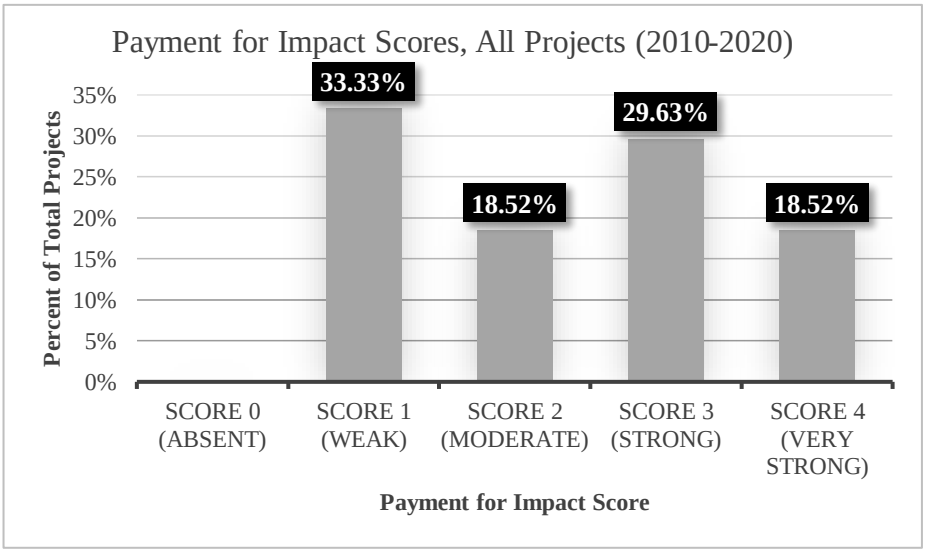
The methods described above are subject to a number of limitations. First, the construction of a new coding frame risks a degree of subjectivity, as findings and interpretation of the data are reliant on the construction of the underlying coding frame. We have attempted to mitigate this by offering the coding frame up for critique and validation from a group of experts in the SIB and public management fields (n=11).[21] Second, the methods undertaken to investigate the research question are narrow in nature and may fail to illuminate the full context of the SIB model's evolution. For example, this paper focuses only on two dimensions of SIB design, and only includes projects that have been launched in the US and UK. This narrow focus may have implications on the findings, and leaves room for future research to investigate broader implications of varying SIB design across the world.

Findings: SIB design in practice

Payment for Impact

Of the 114 projects included in our dataset for analysis, sufficient data was available to code 108 impact bonds along the 'Payment for Impact' dimension (n=81 UK projects, 27 US projects). Of these, 20 projects (19%) scored a 4 for

Figure 6: Payment for Impact Scores, All Projects

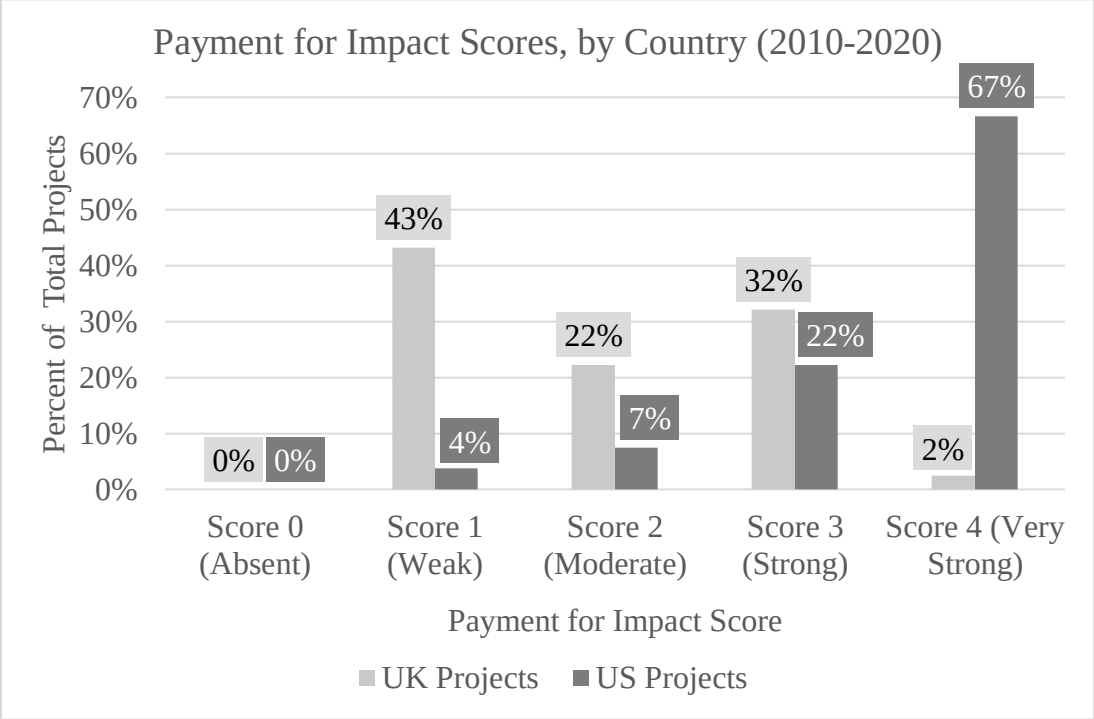


[21] The authors worked through the coding frame step-by-step with a group of SIB academics and practitioners (n=11), incorporating their critiques and suggestions and securing validation of the coding frame construction.

payments made for impact, indicating a combination of majority of payments made for outcomes and a rigorous validation method (as dictated in Table 12 on page 133). To score a 4, these projects had a combination of at least 50% payments made on outcomes and an RCT, or 100% payment on outcomes and a validation method at least as rigorous as a quasi-experimental evaluation. 32 projects scored a 3 (30%) along this dimension, and 20 projects scored a 2 (19%). Projects that scored a 3 either had at 100% payment on outcomes with a pre/post verification method or at least 50% payment for outcomes combined with quasi-experimental validation method. Projects that scored a 2 either had 100% payment for outcomes combined with ex-post verification, or 50% or more payment on outcomes combined with a pre/post verification method. 36 projects scored a 1 (33%), signaling projects with more than half payment on outcomes (but less than 100%) with a simple ex-post verification method. None of the projects scored a 0 for the 'Payment for Impact' dimension, indicating that all projects analyzed were at least partially structured so that payments were bound to the achievement of outcomes. This makes sense intuitively, and validates the population sample approach, as only SIB projects were considered in this analysis, and tying payment to outcomes has been one of the most consistent defining features of SIB projects.

Comparing scores across US and UK SIB projects reveals an interesting finding: US projects are on average more robustly and explicitly tied to impact than UK projects. Most US projects (67%, n=18) scored a 4 on the 'Payment for Impact' dimension, compared with only 2% (n=2) of UK projects. By contrast, 43% (n=35) of UK SIB projects scored a 1 on the 'Payment for Impact' dimension, compared with only 4% (n=1) of US projects. Differences between 'Payment for Impact' scores of the two countries are demonstrated in the Figure 7.

Figure 7: Payment for Impact Scores, by Country



Breaking down the 'Payment for Impact' dimension into its subcomponents reveals that the differences between the US and UK scores are largely due to the more rigorous validation methods typical of many US projects. In Figure 8 below, it is clear that the payment models of projects from both countries are relatively similar: all projects had at least 50% payment on outcomes (score 3), while in the US it is more likely that projects have 100% of payments on outcomes (score 4). More apparent differences between the two countries can be seen in Figure 9 on validation method scores - with the US consistently using more robust methods for outcome assessment compared to the UK. Of the 27 US projects analyzed along this dimension, sixteen (59%) used RCTs and three (11%) used quasi-experimental evaluations to determine outcome payments. In the UK, comparatively, out of 81 projects analyzed, no projects used RCTs (0%) and only two (2%) used quasi-experimental evaluations to determine outcome payments. Most of the UK projects relied instead on simple ex-post verification (54%) or pre/post comparisons (43%) as the outcome validation method.

Figure 8: Payment Model Scores, by Country

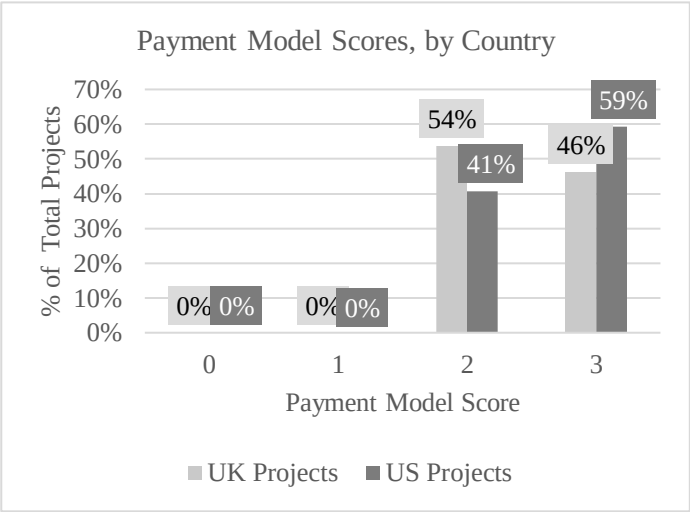
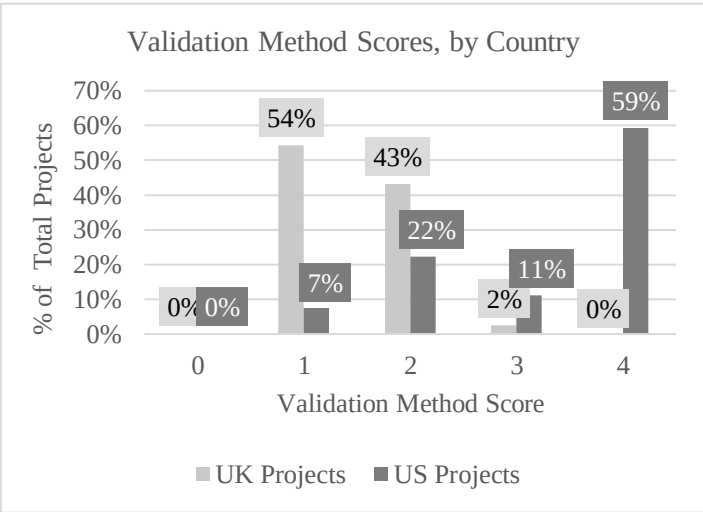
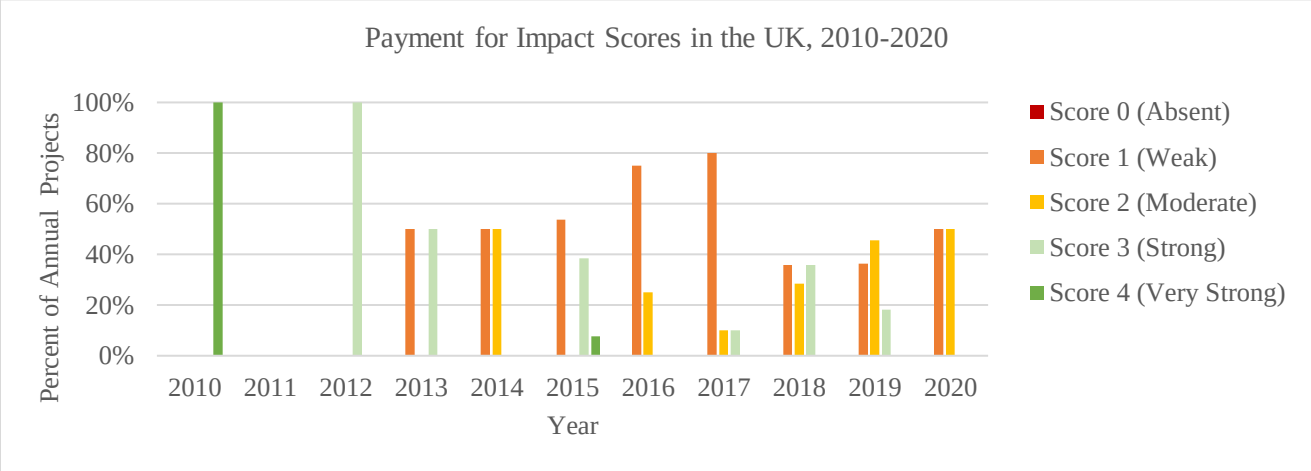


Figure 9: Validation Method Scores, by Country



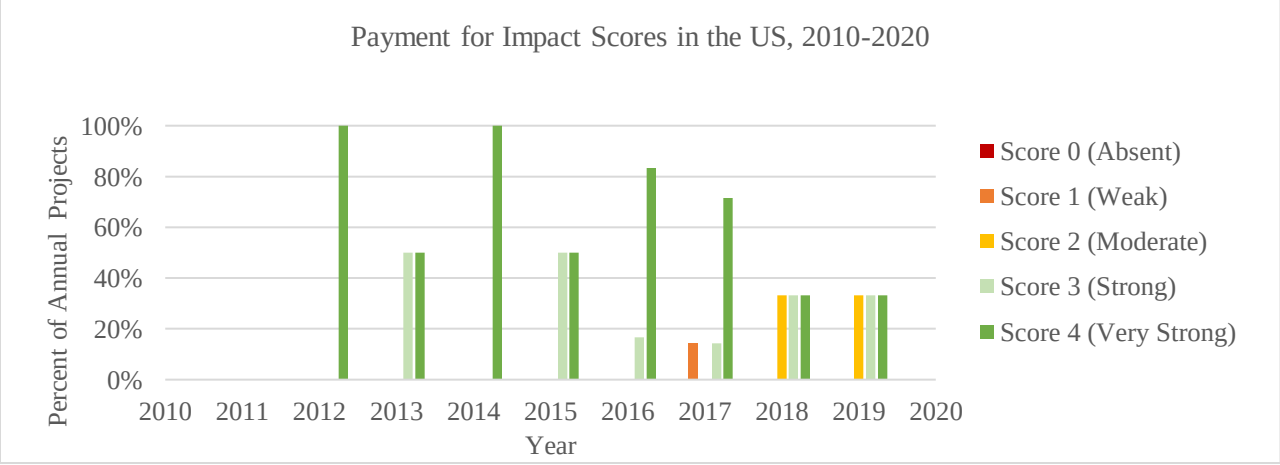
In the UK, majority of the SIB projects had all or nearly all of their payments tied to outcomes, but the process for validating those outcomes was not particularly robust (thus contributing to 'Payment for Impact' scores of 1 and 2 rather than 3 and 4). A robust validation method is important to provide confidence that outcomes were a) actually achieved, and b) can be attributed to the intervention funded by the SIB project. We use this hierarchy of evidence in the validation method to mirror the SIB promises of only paying for what works based on rigorous counterfactual evaluation (Liebman 2011; Cohen 2011; Corrigan 2011; Mulgan et al. 2011; Fox and Albertson 2012; Albertson et al. 2018; Fraser et al. 2018; Chiapello and Knoll 2020). In 2010, the first ever SIB project launched in the UK (and in the world) used a quasi-experimental evaluation to make 100% of project payments based on outcomes. As demonstrated in Figure 10, this standard of payments linked to impact largely went unmatched in subsequent projects. Increasingly, payments were made based on pre/post comparisons or simple ex-post verification, providing far less certainty that SIB projects were paying for outcomes that could be attributed to the SIB-funded intervention.

Figure 10: Payment for Impact Scores, UK



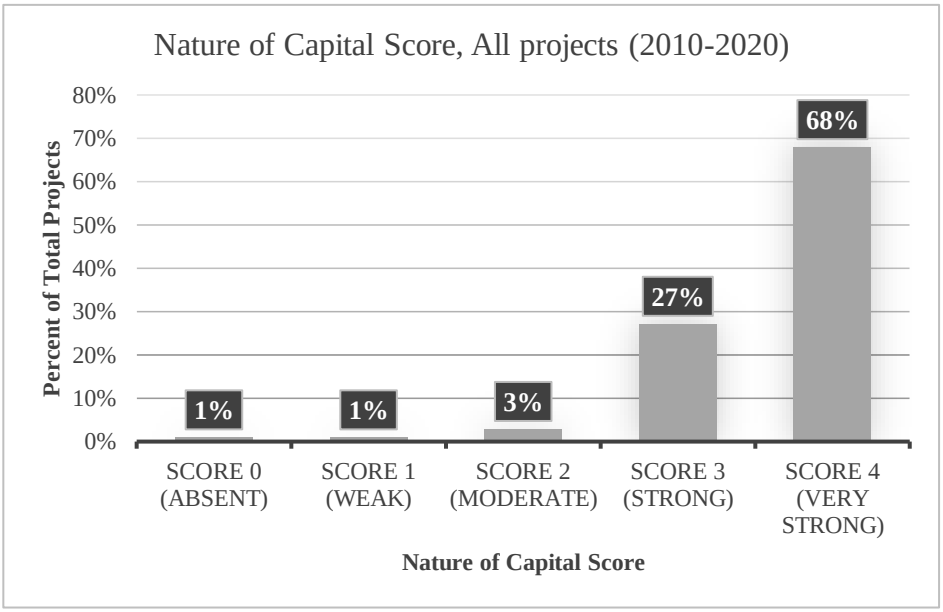
In the US, the majority of SIB projects have all or nearly all of their payments tied to outcomes paired with a rigorous validation method to ensure attribution. In 2012, the US launched their first SIB project with 100% payment based on outcomes using a quasi-experimental evaluation to determine payments. Over time, the commitment to rigorous payment for impact links has remained pretty consistent (see Figure 11), with more than half of the analyzed projects relying on RCTs to trigger outcome payments. However, in the last few years there has been some experimentation with SIB projects that have less robust validation methods (leading to lower payment for impact scores).

Figure 11: Payment for Impact Scores, US



Nature of the capital

Figure 12: Nature of Capital Scores, All Projects

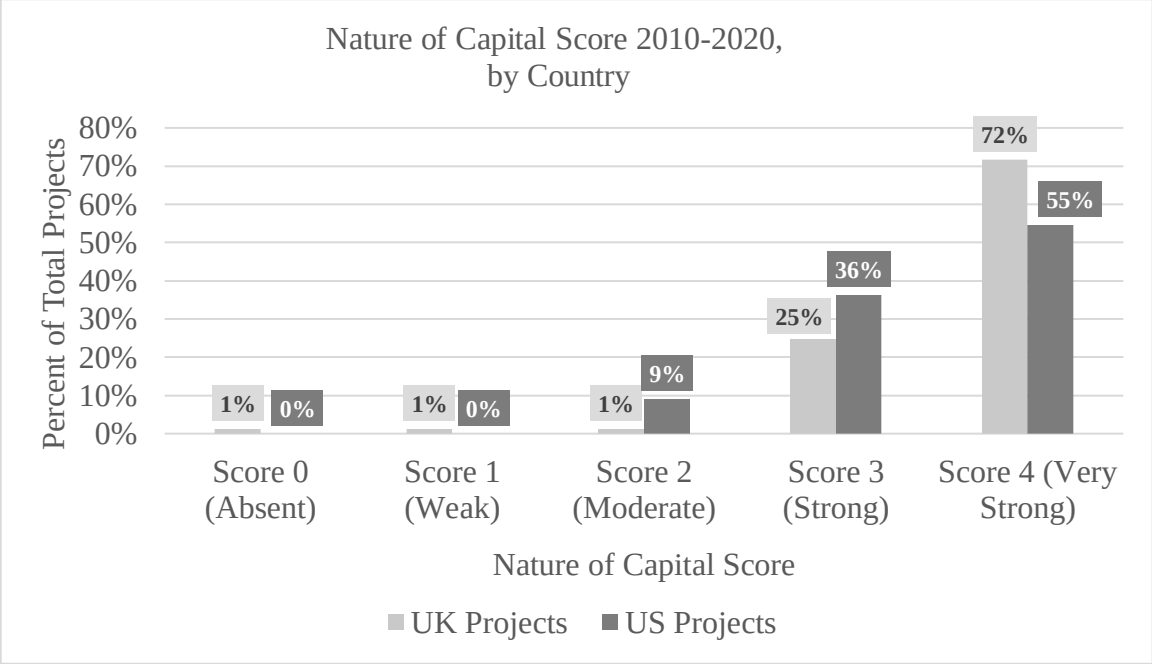


Of the 114 projects included in our dataset for analysis, sufficient data was available to code 103 projects along the 'Nature of Capital' dimension (n=81 UK projects, 22 US projects). Of these projects, 70 projects (68%) scored a 4 for the 'Nature of Capital" dimension, indicating that either a) the primary providers of upfront capital are social or commercial investors and there is limited risk to the service provider in the repayment structure, or b) if the primary provider of upfront capital is philanthropy, there is no risk to the service provider in the repayment structure. 28 projects scored a 3 (27%) along this dimension, 3 projects scored a 2 (3%), 1 project scored a 1 (1%), and 1 project scored a 0 (1%). Projects that scored a 3 either had primarily social or commercial investors of upfront capital and split risk with the provider, or a different primary investor type and limited risk sharing with the provider in the repayment structure. Projects that scored a 2 had at least some risk-sharing with the provider in the repayment structure, combined with a variety of investor types. The project that scored a 1 was primarily financed by provider organizations, and the project's service provider shared in the performance risk. The project that scored a zero was financed by the government partner, meaning the government was the sole provider of upfront capital. In this case, the upfront capital was not independent of the government partner, who held all of the performance-based financial risk.

Comparing scores across US and UK SIB projects reveals similar score patterns: in both countries, majority of the projects have scored a 4 on the 'Nature of Capital' dimension (see Figure 13). SIB projects in the UK

are more likely to score a 4 on this dimension (72%) compared to projects in the US (55%). Conversely, proportionally more US projects scored a 3 (36%) on this dimension than UK projects (25%). However, trends across the two countries are similar, with nearly all projects in both countries scoring either a 3 or 4 along the 'Nature of the Capital' dimension.

Figure 13: Nature of Capital Scores, by Country



Similar scores for the 'Nature of the Capital' across the US and the UK result from different underlying reasons, as we can see when we breakdown the dimension into its subcomponents. In the UK, in over 90% of the projects the primary investor is classified as a social investor (see Figure 14). In the US, however, primary investors are split between philanthropy (52%) and commercial investors (44%). Turning to the repayment structure scores (Figure 15), we see that in the UK it is more likely that investors share risk with provider partners. Majority of the UK projects (73%) involved partial provider risk (compared to 50% of projects in the US), while 26% of projects split risk between both investors and provider (compared to 14% of projects in the US). In the US, 36% of projects completely protected providers from performance risk, compared to just 1% of projects in the UK.

Figure 14: Primary Investor Scores, by Country

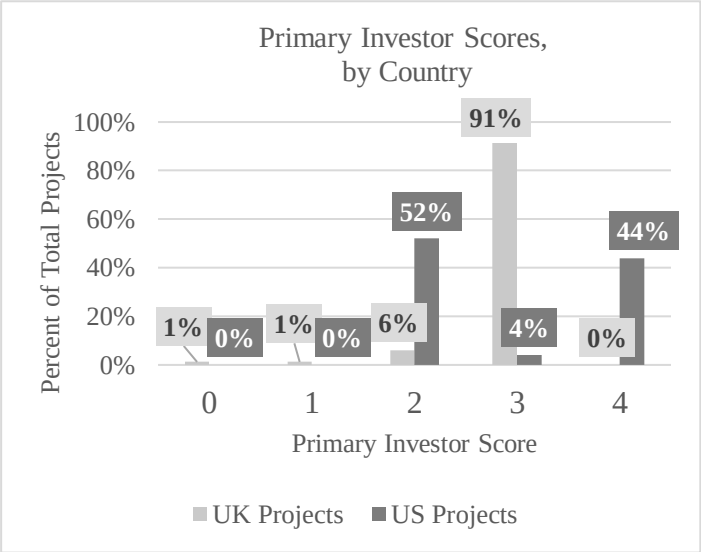
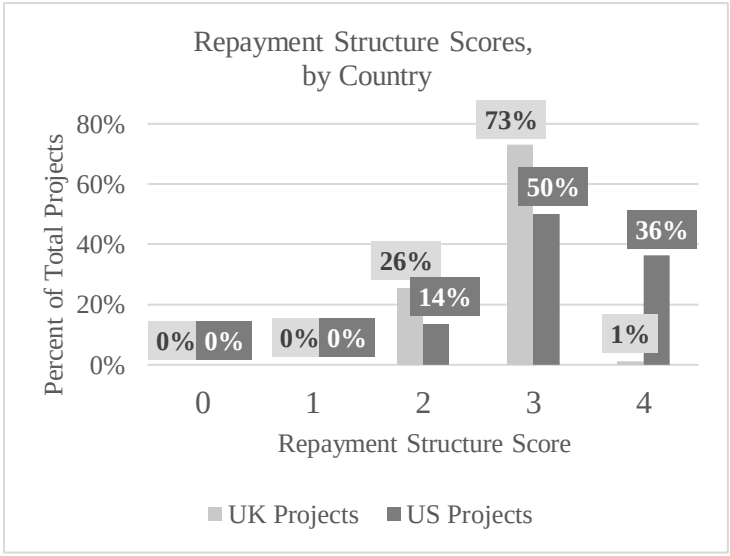
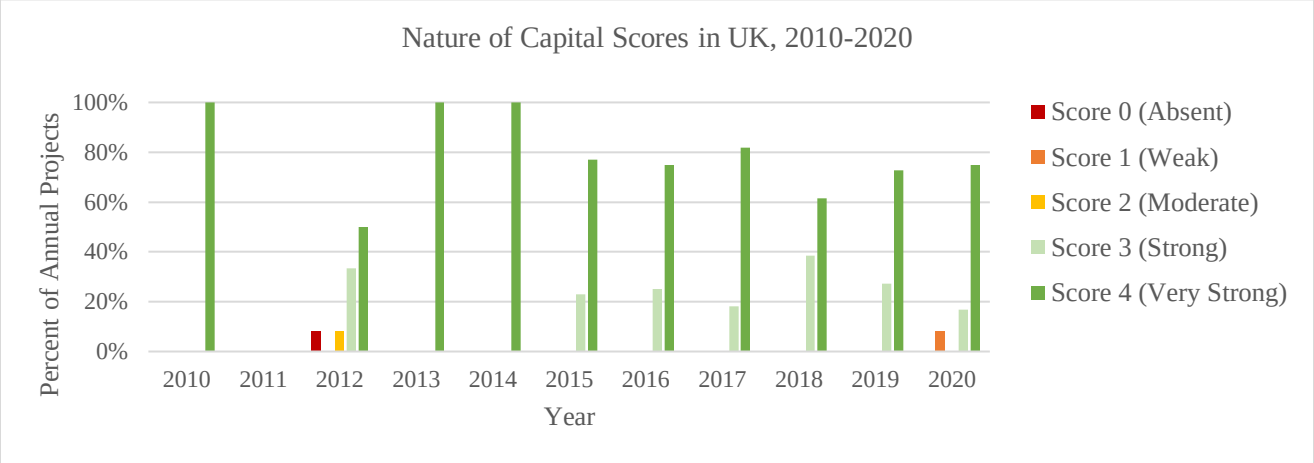


Figure 15: Repayment Structure Scores, by Country



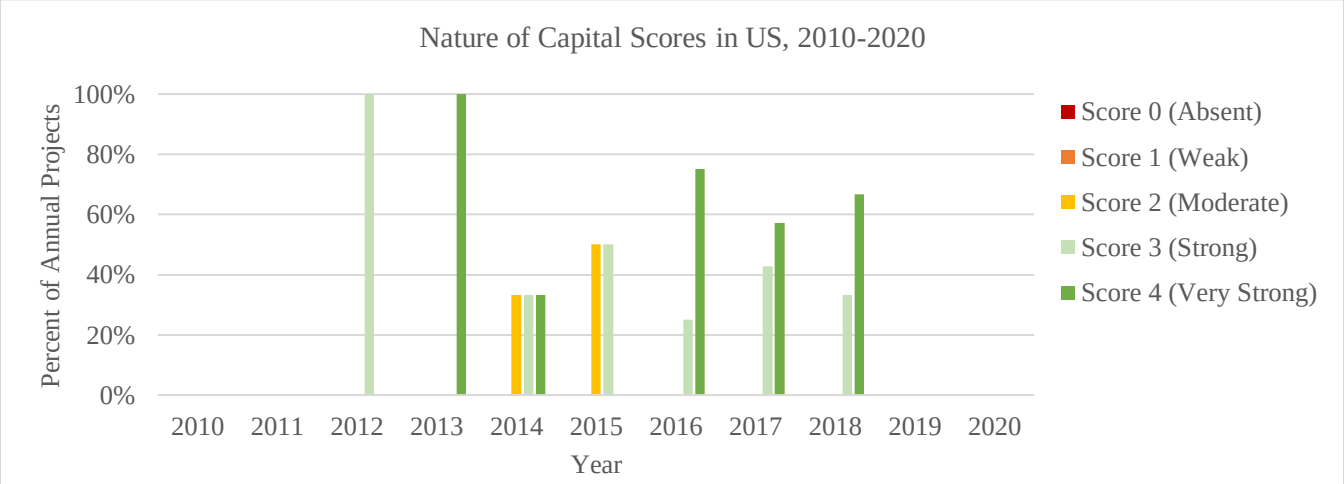
Majority of the SIB projects in the UK had social investors as the primary source of upfront capital, combined with at least some risk-sharing with providers in the repayment structure (in most cases investors could withhold payments based on service providers not meeting certain key performance indicators, but otherwise providers were paid based on services delivered rather than outcomes). In these cases, most of the financial risk of poor performance sits with social investors, yielding the nature of the capital independent and at-risk (score 4). In other cases, risk-sharing with the service provider was more explicit, and both the investor and service provider were paid based on performance. Projects structured this way received a score of 3 on the 'Nature of Capital' dimension. Over time, this structure has remained fairly consistently used and in recent years, most UK projects have scored a 4 on the 'Nature of Capital' dimension (see Figure 16).

Figure 16: Nature of Capital Scores, UK



In the US, most SIB projects incorporated partial or no risk-sharing with service providers in the repayment structure. A few projects, those that scored a 2, did have performance risk shared between investors and the service provider (in the form of deferred fees), but in most projects service providers are paid regardless of performance. The primary investors in US projects were split between philanthropy and commercial investors, and in nearly all cases philanthropy was involved. Typically, philanthropy served as a junior lender or guarantor, but in our analysis was only designated the 'primary' investor when philanthropic organizations contributed more than commercial investors. In these cases, nearly all financial risk of poor performance was held by philanthropy, contributing to lower 'Nature of Capital' scores. Over time, the 'Nature of Capital' dimension for US projects has remained generally consistent other than a relatively stronger philanthropic presence in the years 2014-2015 (see Figure 17). Commercial investors backed by philanthropic organizations have continued to be the primary investors in US SIB projects, with limited risk sharing with the service provider.

Figure 17: Nature of Capital Scores, US



Discussion: Have we stretched SIBs too far?

Based on this empirical analysis, have SIBs stretched too far from their original intentions? Are they much different from status quo contracting arrangements? Looking at how closely payments are tied to impact, a quintessential promise of the SIB model, we see two different stories across the UK and the US. In the UK, project payments are largely made for outcomes instead of outputs or activities, but these outcomes are typically not subject to rigorous validation methods. This calls into question whether or not SIB projects are making payments for outcomes that can be attributed to SIB-funded interventions . And as a result, it is hard to say for certain whether payments are truly being made on the basis of a project's impact. For example, in the 10 Innovation Fund SIBs, investors were repaid for the payable outcomes despite a subsequent evaluation demonstrating that the projects were not successful in helping young people engage in education or employment compared to a control group (Salis, Wishart, and McKay 2018). In the US, on the other hand, SIB projects typically incorporate robust evaluations that serve to validate outcomes before payments are made. This suggests that in the US it is more likely that the SIB model has been successful in shifting the status quo of service delivery to incentivize and pay for project impact. On average, SIB projects in the US are much larger than those in the UK (in dollars spent and beneficiaries served) (INDIGO 2021),

which may make it easier to fund and run experimental or quasi-experimental evaluations alongside SIB projects.

Analyzing the nature of the capital used to fund SIB projects also highlights two different approaches across the two countries. In the UK, typically social investors have served as the primary provider of upfront capital, but often rely on a risk-sharing agreement with the provider that is baked into the repayment structure. In the US, there is typically less risk-sharing with the provider, but commercial investors are often backed by philanthropic dollars as a way to mitigate risk. In both contexts, projects scored highly on the nature of the capital, signaling that upfront capital is generally both independent and at-risk. However, investors in SIB projects have relied on different risk mitigation strategies that indicates these arrangements fall short of a full transfer of risk away from the government and provider partners. In fee-for-service arrangements the government holds performance risk, and in traditional payment by results arrangements providers have held the performance risk. By not meaningfully transferring risk away from these two parties, the proposed value-add of the SIB model is potentially diminished. To fully unpack this, it is helpful to look at the two SIB dimensions (payment linked to impact and nature of the capital) in tandem.

If payments aren't made for impact, is the upfront project capital fully at-risk? While forms of operational risk remain, risk of underperformance is mitigated if the payment schedule is heavily weighted to activities, outputs, or outcomes that aren't attributed to the intervention. In these cases, when payment is not based on impact, the government agency is effectively taking on the risk that the service does not deliver the ultimate desired social outcomes. If investors are getting paid back without real attempts at attribution, then perhaps upfront capital isn't really at as much risk as the SIB model promised. Gustaffson-Wright, Massey, and Osborne (2020) found that in 48 out of 50 completed impact bonds (96%), outcomes were achieved and investors were repaid. The authors note an ongoing issue with attribution, which has likely contributed to the high success rate of projects paying out for outcomes achieved. In a study of 13 SIBs launched in five

countries[22], including the UK but not the US, Hevenstone and von Bergen (2020) found that contractual targets (i.e. payable outcomes) were never linked to evaluations, and rather payable outcomes were decided on through political negotiations that had a "strong potential to inflate government costs" and result in "achievable targets" selected. Painter (2020) reports similar findings in a study of US and UK SIBs, arguing that SIB investors are nearly always repaid by government and frequently with a higher rate of than typical government borrowing (Dixon and Rosenbach 2020).

In this sense, having a heavy weighting of payment for outcomes combined with robust validation methods can make projects riskier for investors, because it ensures that payable outcomes are those that can be attributed to the SIB-funded intervention. In the UK, an outcomes framework that does not rely on robust validation methods de-risks SIB projects to some extent, providing social investors more incentives to fund SIB projects. This less robust outcomes framework, combined with risk-sharing repayment structures with the providers, are two primary ways that SIB projects in the UK have mitigated the transfer of risk to private investors. In the US, a strong outcomes framework has potentially made SIB projects riskier, by ensuring payments are made for project impact. In this context, US projects have relied on other means to de-risk SIB investments, primarily relying on philanthropic organizations to a) serve as the primary investor of capital, b) provide junior loans that safeguard large portions of commercial investment by being the first to lose, or c) serve as guarantors in the case of underperformance to pay back commercial investments. It is important to note that these philanthropic dollars are not always structured as grants; in many cases the philanthropies seek to recover their initial investment and in others are also paid a performance return. Further research may seek to investigate the different kinds of philanthropic investments characteristic of SIB projects, to further understand the role of philanthropy in the US (and other's) SIB market.

[22] Countries represented in the study include the UK, the Netherlands, Switzerland, Germany, and Austria.

There are various reasons why different trajectories in SIB design may have occurred across the two countries. The narrative of using the SIB model to test new ideas and figure out what works was stronger in the US, contributing to the design of stronger outcome frameworks alongside SIB projects; using rigorous validation methods was seen as a tool to build the evidence base of innovative interventions funded by the SIB model (Albertson et al. 2018). It is also possible that tax advantages for philanthropy combined with a more robust landscape of philanthropic foundations in the US contributed to greater participation by these actors in the SIB model. These philanthropic funders have played a role in shaping SIB design in the US, and may have had a preference for funding interventions attached to rigorous evaluations. In the UK, on the other hand, the central government has created many 'rate card' type outcome funds, which make payments for a set list of outcomes that are not attached to rigorous evaluation methods. Perhaps the reduction in transaction costs achieved via this rate card model has been prioritized in the UK over the original promise that SIBs would incentivize payment for impact. Additionally, shorter-term outcome payment arrangements can enable the recycling of funds (using outcome payments to fund project costs), rather than requiring the full project cost to be invested upfront and put at-risk.

Differences in the range of investor partners may also play a role in the different approaches between the US and the UK; Evans (2020) reported high investor diversity in the US SIB projects and low investor diversity in the UK projects. Our findings echo this claim: in the UK, a few specialist social investment fund managers have been involved in nearly all of the country's SIB projects. These few social investors have been major players in the UK SIB market, with the potential to influence policy and shape project design across the country. One such investor, Big Society Capital, self-described as the "leading financial institution dedicated to social impact investment in the UK" (BSC 2021) was launched by the UK central government in 2012 (Cabinet Office 2012) to transform and champion the development of a social investment market in the UK (Harvie 2019). The fundamental and continuous role of the state in the creation of the UK's social investment market (Harvie 2019) may call in to question what degree this investment is really independent from government.

The data suggests that the SIB model has stretched far from its original intentions in a few key ways. In the UK, payable outcomes are generally not subject to rigorous validation methods, calling in to question whether payments are meaningfully linked to impact. Early proponents of the model argued that by "allocating resources to where they will achieve the most impact," the use of SIBs could help to achieve impact with greater efficiency than other forms of service provision (Fox and Albertson 2012). Ronicle and Smith (2020) found that one of "the strongest effects" of the SIB model was that "the SIB increased the focus on achieving outcomes," and enabled service providers to adapt in efforts to better achieve those outcomes. So, it is problematic then, if outcomes represent tick-the-box targets rather than impact that meaningfully improves the lives of beneficiaries (McHugh et al. 2013). Linking payments to outcomes that do not translate to impact might "unravel the core logic of the SIB mechanism" (Carter 2020).

In both the US and the UK, risk mitigation strategies incorporated into the nature of the upfront capital may limit the intended transfer of risk to external investors. An outside investor, bearing the risk that the project does not perform, is perhaps one of the most distinguishing features of the SIB model (Fraser et al. 2016). If performance risk sits with the government or the service provider, it renders the model similar to status-quo commissioning approaches such as fee-for-service or payment by results. If performance risk sits with philanthropy, the arrangement may resemble a grant or program-related investment and raises questions around why upfront commercial investment, and the associated returns, are needed in the first place. Perhaps there is not enough risk-appetite for external investors (commercial or social) to take on the purported risk transfer of the SIB model without employing some of the various risk mitigation strategies we have seen thus far. If risk-transfer is limited and uncertainty remains around whether payments are linked to impact, it is worth questioning what value this hybrid arrangement brings to the government partner.

Conclusion:

This paper represents the first systematic attempt to map variation in the design of SIB projects against the espoused model, and allows us to draw conclusions on the model's ability to live up to its original promises. We contribute to the ongoing debate on how SIBs impact public service delivery by offering theoretical and empirical contributions: operationalizing two major dimensions of SIB design and then plotting out the range of different SIB configurations tested in the US and UK to see how SIBs have been designed in practice. We have found that SIBs have indeed "stretched" from what makes it a distinctive, innovative approach in some cases, potentially undermining its potency as a tool for service delivery improvement.

It is important to note that this analysis is subject to limitations. The methods undertaken to investigate variation in SIB design are narrow in nature and may fail to illuminate the full context of the SIB model's evolution. For example, this paper focuses only on two dimensions of SIB design, and only includes projects that have been launched in the US and UK. Other countries may have experimented with the SIB model in different ways, perhaps designing projects differently than those in the US or UK. In addition, limited data, particularly regarding the nature of the capital used to fund SIB projects, led to the use of assumptions and estimates in the data analysis. These assumptions have been tested with sensitivity analyses in Appendix 9.

We have found that SIB projects in the UK, whilst tied to outcomes, are not rigorously linked to impact, particularly in comparison to US SIB projects. We have also found that upfront capital used in SIB projects from both the US and the UK typically appears to be independent and at-risk, but various risk mitigation strategies may limit risk transfer of poor project performance on to private investors. Future research might seek to investigate the social intent of service provider organizations participating in SIB projects. If many of the UK service providers are charities, then perhaps the differences between the US

(with a heavier reliance on philanthropic risk-taking) and the UK SIB markets aren't as pronounced as they may appear. Similarly, it would be interesting to further explore the nature of philanthropic investments, most prevalent in the US, to determine how these funding arrangements are structured and what function philanthropy serves in SIB projects. Future research might also explore how SIB projects have fulfilled other potential promises, including service innovation, market discipline, or meaningful cross-sector collaboration. Are one of these, perhaps, what the SIB model brings to the table?

Conclusion

This thesis has contributed to public management debates on how to best organize the delivery of public services by focusing on the social impact bond model. The SIB model, with roots in new public management ideals, has been positioned as an innovative, hybrid model of government commissioning involving a three-way partnership between government, private investors, and service providers. Through three separate articles, following the requirements of an article-based dissertation, this thesis has made significant theoretical and empirical contributions to questions around the impact of the SIB model on various levels of service delivery.

This conclusion chapter revisits the thesis' aims and objectives, and then summarizes findings from the three articles relative to those aims. The strengths and limitations of the thesis are then discussed, followed by an exploration of the thesis' significance to the field. The chapter closes with a discussion of policy implications and future opportunities for research.

Research Aims

The aim of this thesis was to explore the impact of the SIB model at various levels of public service delivery, contributing theoretical and empirical insights to the literature on how hybrid models of public service delivery work to support vulnerable populations. The first two articles examine a case study of a social impact bond launched in Essex, United Kingdom to support families of children at-risk of entering the care system (i.e. being removed from their parental homes). In the first article, I examine the effect of the SIB at a systems level, seeking to better understand the impact of the SIB arrangement - which brings together a multi-sector partnership, data monitoring infrastructure, outcomes contract, and external investment - on the system of public service delivery. In the second article, I examine the effect of the SIB model at an organizational level, seeking to better understand the impact of the arrangement on the service provider and front-line staff. The third article, co-authored alongside Drs. Eleanor Carter and Mara Airoldi, relied on

qualitative content analysis of all SIBs launched in the UK and the US to date (n=114). In this article, we investigate how the SIB model has evolved over time, seeking to better understand whether it follows through on some of its key promises for altering the structure of service delivery.

Following from these aims, the particular objectives of this thesis were:

1. To better understand how the different features of the SIB arrangement - a multi-sector partnership, data monitoring system, payment for outcomes contract, and external investment - combine to affect systems of public service delivery
2. To better understand whether and how the SIB model implemented with a high-fidelity intervention affects the behavior and perceptions of front-line staff in regard to support networks, performance management, flexibility, and incentives
3. To better understand the implications of variation in the design of SIB projects across two key dimensions: payment made for impact and nature of the upfront capital

Following on from these objectives, returning to the table below provides an overview of the themes explored with the high-level findings across the thesis' three articles.

Table 15: Overview of themes explored and high-level findings across three articles

SIB Features	Multi-sector partnership	Data monitoring system	External investment	Outcomes contract
Article	**Themes explored and high-level findings**			
Article 1: SIB effect on system	**Stakeholder relationships**: Culture clashes and tensions over governance structure	**Performance management & resource allocation:** Sense of disproportionality relative to other programs in the wider system	**Flexibility & democratic accountability**: Gov't loss of control of service direction and limited flexibility	**Outcome alignment & incentives**: Aligned to outcome goals, but financial driver put pressure on the system
Article 2: SIB effect on front-line staff	**Support networks**: Issues with complex systems of managerial oversight	**Performance Management**: Output-driven focus; Administrative burden at level of supervisors	**Flexibility**: Provider discretion heavily mitigated by high-fidelity intervention	**Incentives**: High-fidelity intervention limits outcome gaming incentives, but not pressure to meet output targets
Article 3: Variation in SIB design			**Nature of upfront capital (risk transfer):** Capital typically independent and at-risk, but subject to risk mitigation strategies	**Payment linked to impact**: Robustly linked to impact in US, not so much in UK

Summary of Findings

The section below summarizes findings from the three articles, including the SIB effect on the system of public service delivery, the SIB effect on providers and front-line staff, and the evolution of SIB design over time.

SIB effect on the system of public service delivery

In the first paper, I offered a conceptual framework for how SIBs might impact public service delivery that stems from the SIB model's four key features. This conceptual framework, demonstrated in Table 16 below (and also as Table 7 on page 65), is an original contribution based on a review of the existing literature. It

provides a basis for which to analyze the impact of the SIB model at a systems-level, suggesting that the four key features of a SIB (multi-sector partnership, data-monitoring system, external investment, and payment for outcomes contract) can each enable mechanisms that improve or impede public service delivery. A multi-sector partnership can enable collaboration, leading to stakeholder cooperation and meaningful linkages across organizations, but it can also enable conflict, leading to stakeholder clashes, lack of service integration, or relational tensions. The data-monitoring system might enable active contract management, leading to the continuous use of data to identify challenges and implement improvements in a way not typical to public services, but it may also create a resource drain with high levels of administrative burden and pressures on front-line staff. The external investment might enable flexible funding, yielding tailored support to services users or additional resources available for service delivery, but it might also lead to less democratic accountability, characterized by an inefficient use of resources or lack of transparency. The payment for outcomes contract may enable outcome goal alignment, leading to more outcomes achieved that lead to meaningful impact for beneficiaries, but it could also introduce perverse incentives through a narrow scope of services offered or explicit gaming of outcome payments.

Table 16: How particular features of SIBs might impact public service delivery

SIB features	1. Multi-sector partnership	2. Data monitoring system	3. External investment	4. Payment for outcomes contract
Mechanisms to improve public service delivery	**Collaboration**	**Active contract management**	**Flexible funding**	**Outcome goal alignment**
What this may look like on the ground	Stakeholder cooperation; Effective governance structure; Intervention well integrated into existing system	Continuous use of data to identify challenges; Implementation of improvements in real-time	Tailored support to service users; Provider discretion; Additional resources available for service delivery; Risk of project failure with external investor	Outcome-driven service with buy-in at all levels; Outcomes achieved lead to impact for service users
Mechanisms to impede public service delivery	**Conflict**	**Resource drain**	**Less democratic accountability**	**Perverse incentives**
What this may look like on the ground	Stakeholder clashes; Relational tensions; Lack of service integration; Unclear lines of accountability	Administrative burden; Pressures on front-line; Disproportionality	Decisions on public services dictated by non-public motive; Lack of transparency; Inefficient use of resources	Narrow scope of services; Skepticism around outcome goals; Gaming to earn outcome payments

Applying this framework to the SIB project launched in Essex provided insight on how each feature of the SIB model affected processes of public service delivery. I found that the SIB had mixed effects on public service delivery in Essex. The SIB model enabled rigorous monitoring of performance, availability of additional resources for the service, and outcome goal alignment across stakeholders. However, the SIB model also generated stakeholder conflict, feelings of resource burden and disproportionality by the government partner, loss of democratic accountability, and the pressuring presence of an underlying financial driver. In sum:

- The multi-sector partnership allowed for joint working across sectors, but also lead to significant relationship tensions that greatly influenced the way that stakeholders viewed the project.

- The data monitoring system enabled rigorous monitoring of project performance with the aim of implementing improvements, but also led to a large resource burden and feelings of disproportionality by the local council.
- External investment allowed for additional resources to be injected into the system, but flexibility was limited and the local council lost control of the service direction even when they felt a different service would have been better for their constituents.
- The outcomes contract aligned all stakeholders to keeping children safely out of care as the ultimate project goal, but introduced an underlying financial driver with potential consequences related to pressure on generating referrals and outcome payments being too easy to achieve.

A summary of these findings is demonstrated in Table 17 below (and also in Table 8 on page 74).

Table 17: Summary of Findings from the Essex SIB

SIB features	Findings
1. Multi-sector partnership	**More conflict than collaboration** - Difficulty integrating SIB into the system - Challenges around governance structure and relationship tensions - Clash of cultures between private, public, and voluntary sector ways of working
2. Data monitoring system	**Active contract management at a high perceived cost** - Continuous monitoring enabled identification of challenges and trailing of solutions - Burden on stakeholders and a sense of disproportionality relative to other programs
3. External investment	**Limited flexibility and loss of democratic accountability** - Flexibility to spend additional resources but often faced pushback - Locked into contract specifications with restricted ability to course correct - Enabled new service in array of support for children at the edge of care
4. Payment for outcomes contract	**Aligned to outcome goals, but not fully protected from financial drivers** - Aligned to keeping children out of care, albeit with misalignment over scope of the target population and services provided - Presence of clear underlying financial driver putting pressure on system

A variety of contextual components played a role in how the SIB affected public service delivery in Essex: including a lack of operational buy-in at the project's outset, the choice of a license-based intervention, and

a complex payment mechanism and governance structure. Contextual factors will present differently for each SIB project, and will likely influence the effect that the SIB model has on public service delivery.

SIB effect on front-line staff

Policymakers and scholars alike have expressed concerns around the impact of the SIB model on service providers and front-line staff. Relying on 'black-box' interventions, where the process of delivering a social intervention is not prescribed nor monitored, and paying based only on outcomes raises concerns around the potential for outcome gaming. For example, this might mean focusing on service users most likely to generate outcome payments at the expense of other, harder-to-reach service users. Based on existing evidence and a review of the literature, I offered an original thematic summary of how the SIB model might affect provider organizations and front-line staff (see Table 18 below, also Table 9 on page 104). In brief, I lay out four categories through which the SIB model may affect front-line staff: performance management, support networks, flexibility, and incentives. A SIB's performance management requirements are generally more intensive than traditional service delivery, and may detract from face time with service users. Front-line staff may have more robust support networks with a SIB, but may also face additional lines of accountability (including to an investor partner). In a SIB, front-line staff may have increased flexibility in their ability to tailor services or incorporate innovative approaches to meet outcomes. Finally, the SIB structure brings a new relationship with incentives; front-line staff may be incentivized to 'game' the contract by focusing on clients likely to achieve payable outcomes.

Table 18: How the SIB structure may affect front-line staff, according to existing evidence

Main themes	Subthemes
Performance management	+ Increased use of performance data to respond to challenges -- Administrative burden (impacts time with service users) ? Pressure to meet outcomes and/or outputs
Support networks	+ Heightened job security + Availability of additional resources + Collaborative knowledge flows -- Culture clashes with investors -- Complex systems of managerial oversight
Flexibility	+ Ability to tailor or individualize services + Incorporation of innovative service approaches + Provider discretion over service delivery & resource allocation
Incentives	+ Outcome goal alignment -- Outcome gaming incentives -- Narrowed scope of services ? Consequences for missed targets ? Financial or other rewards

Table Key:
+: Positive effect
--: Negative effect
?: Unclear effect

In the second paper, using the Essex case study, I investigated whether using a prescribed and replicable (i.e. high-fidelity) intervention, as opposed to the black-box approach, insulates providers and front-line staff from potential negative effects of the SIB model. I found that while using a high-fidelity intervention with the SIB model mitigated incentives to game outcome payments in Essex, it also restricted the ability of the provider to innovate and deliver flexible, context-driven services. Using a high-fidelity intervention did not protect providers from pressure to meet output targets, which could potentially create harmful pressures around generating referrals and enrolling families regardless of fit. A summary of the effects on front-line staff in Essex is depicted below in Table 19 (and also in Table 10 on page 118).

Table 19: Effects on front-line staff of high-fidelity SIB intervention in Essex

	SIB effects that were mitigated	SIB effects that remained
Performance management	+ Increased use of performance data to respond to service quality challenges ? Pressure to meet outcomes	+ Increased use of performance data to respond to operational challenges ? Pressure to meet outputs -- Administrative burden
Support networks	+ Collaborative knowledge flows	+ Heightened job security + Availability of additional resources -- Culture clashes with investors -- Complex systems of managerial oversight
Flexibility	+ Ability to tailor or individualize services + Incorporation of innovative service approaches + Provider discretion over service delivery & resource allocation	
Incentives	-- Outcome gaming incentives	+ Outcome goal alignment -- Narrowed scope of services ? Consequences for missed targets ? Financial or other rewards

Table Key:
+: Positive effect
--: Negative effect
?: Unclear effect

The findings from the Essex case study provide insight on how the effects of SIB model might vary under different policy contexts, contributing to the existing body of literature that has sought to unpack the SIB effect on front-line staff. Using a high-fidelity intervention with the SIB model comes with both advantages and disadvantages compared with black-box approaches. The provider's flexibility is limited, mitigating a commonly cited promise of the SIB model and diminishing potential benefits of the model for front-line staff (Arena et al. 2016). On the other hand, a high-fidelity intervention can ensure minimum standards of service delivery for beneficiaries and help protect against explicit outcome gaming. In Essex, the potential for gaming was weakened because the processes for serving families enrolled in MST were standardized.

Unfortunately, there are ways other than explicit outcome gaming that the SIB model can create disincentives in service delivery, which are not protected by the high-fidelity nature of the intervention.

Provider organizations and front-line staff may experience pressure to meet output targets, such as the number of referrals or enrollments to the SIB service. This pressure can create harmful disincentives, particularly if it means referrals are made regardless of a service user's goodness of fit to the intervention. There has been evidence of services users enrolled in programs because they were more easily accessible or deemed easier to help (Lowe et al. 2019; Ronicle and Smith 2020). In Essex, provider staff reported pressure to generate as many referrals as possible, particularly at the start of the program. While not outcome gaming, this pressure on service delivery is important to reflect on as it could result in service users being referred to a SIB intervention when they may have been better served elsewhere.

Evolution of SIB design over time

In the final paper, alongside Drs. Eleanor Carter and Mara Airoldi, we investigated variation in SIB design across two major dimensions: the degree to which payment is linked to impact, and the independent, at-risk nature of the working capital provided. Offering a theoretical contribution to this debate, we operationalized the two dimensions, creating a coding frame that allowed for project comparison. Payment linked to impact was operationalized as a combination of the payment model, the percentage of the contract payment that is linked to outcomes, and the validation method, the rigor of the approach taken to assure outcome achievement. See Table 20 below (and Table 11 on page 132).

Table 20: Payment for Impact Operationalization

Score	0	1	2	3	4
Payment Model	0% of payment on outcomes	Less than half payment on outcomes (more than 0%)	Half or more payment on outcomes (but less than 100%)	100% of payment on outcomes	N/A
Validation Method	None	Simple ex-post verification	Pre/post analysis or historical comparison	Quasi-experimental	Randomized Control Trial (RCT)

The nature of the capital was operationalized as a combination of the primary investor of upfront capital, the category of investors who contributed the most to upfront project costs, and the repayment structure, defined as the risk-sharing split between the investor and service provider. See Table 21 below (and Table 13 on page 135).

Table 21: Nature of the Capital Operationalization

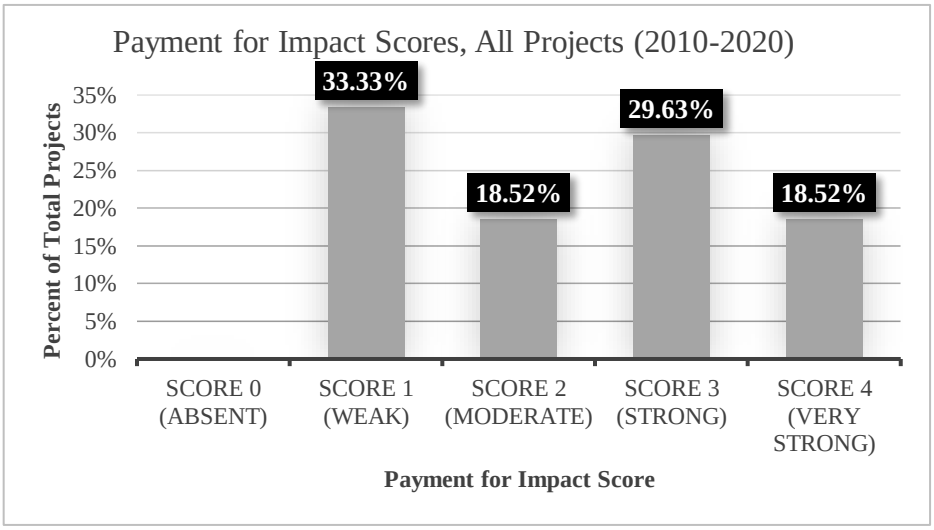
Score	0	1	2	3	4
Primary investor type	Government	Provider	Philanthropy	Social investors	Commercial investors
Repayment structure	Provider holds all performance risk	Provider primarily holds performance risk	Provider and investor share performance risk	Investor primarily holds performance risk	Investor holds all performance risk

For both dimensions, payment linked to impact and nature of the capital, we created composite scores to analyze the dual underlying components in parallel. These composite scores allowed us to rate each project one score (0-4) for both the payment linked to impact dimension and the nature of the capital dimension.

Applying this coding frame to all US and UK projects launched from 2010-2020 (n=114), we found that SIB design in practice has evolved from the original promises in a few key ways: 1) in the UK, payable outcomes are generally not subject to rigorous validation methods, calling in to question whether payments are meaningfully linked to impact, and 2) in both the US and the UK, upfront capital is generally independent and at-risk, but risk mitigation strategies in the repayment structures may limit the intended transfer of risk to external investors.

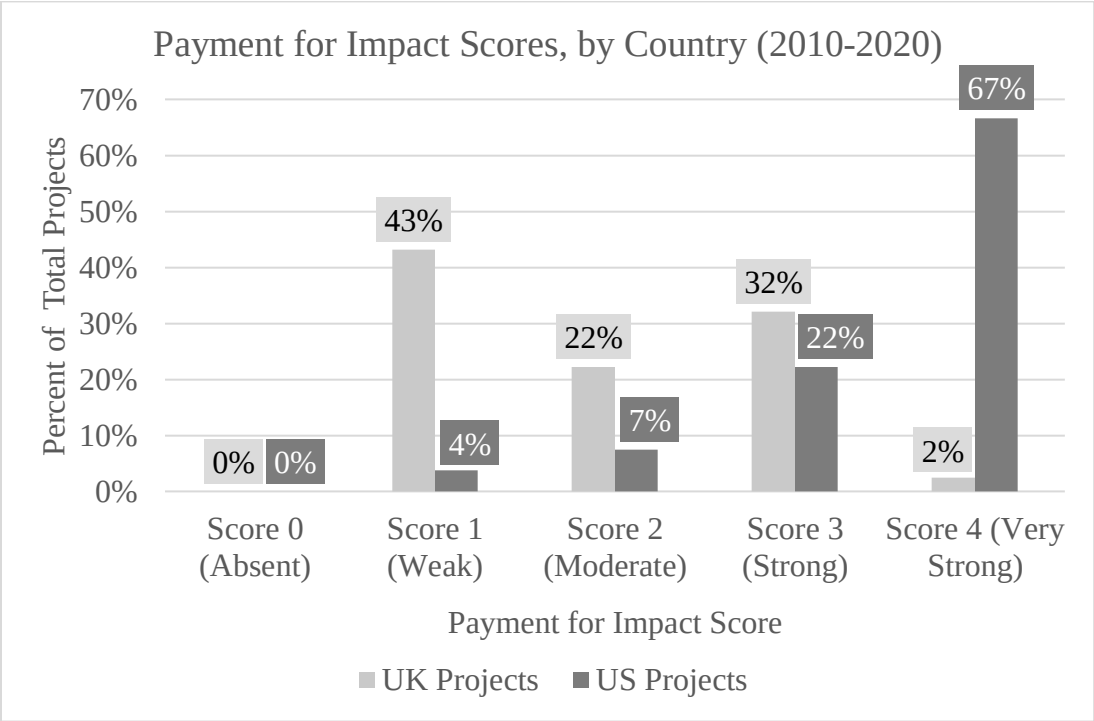
Sufficient data was available to code 108 projects along the 'Payment for Impact' dimension (n=81 UK projects, 27 US projects). Of these, 20 projects (19%) scored a 4, 32 projects scored a 3 (30%), 20 projects scored a 2 (19%), and 36 projects scored a 1 (33%). Interrogating this data reveals that all projects had at least 50% or more of contract payments tied to outcomes, but with a variety of validation method types. Projects that scored a 1 used a simple ex-post verification method, while projects that scored a 4 used either an RCT or quasi-experimental evaluation.

Figure 18: Payment for Impact Scores, All Projects



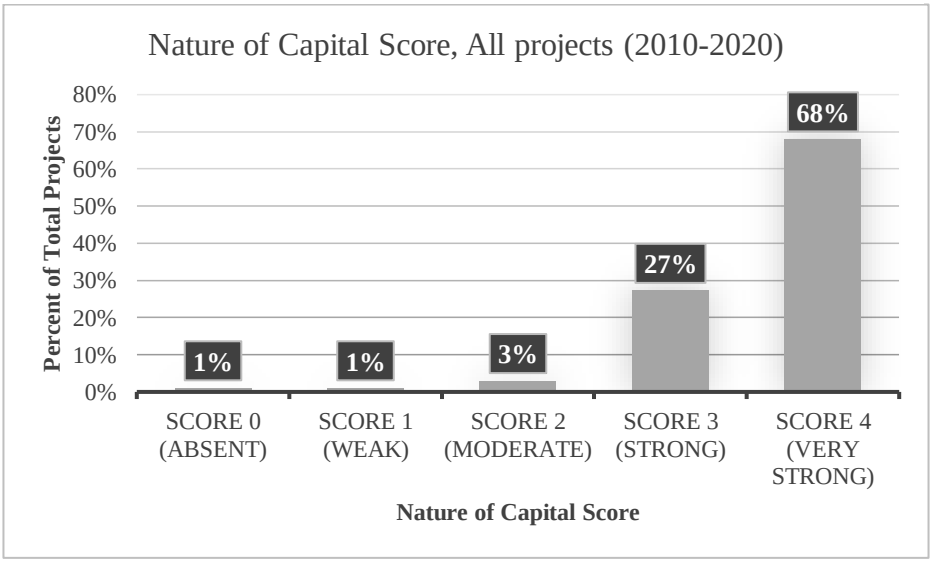
Comparing US and UK SIB projects reveals that, on average, US projects more robustly and explicitly tied to impact. Most US projects (67%, n=18) scored a 4 on the 'Payment for Impact' dimension, compared with only 2% (n=2) of UK projects. This finding is largely due to the more rigorous validation methods typical of many US projects. Figure 19 below shows the percentage of projects across each payment for impact score for both UK and US projects (see also Figure 7 on page 140).

Figure 19: Payment for Impact Scores, by Country



103 projects were coded along the 'Nature of Capital' dimension (n=81 UK projects, 22 US projects). Of these, 70 projects (68%) scored a 4, 28 projects scored a 3 (27%), three projects scored a 3 (3%), one project scored a 1 (1%), and one project scored a 0 (1%). The most common score, a 4, indicates that in a majority of the projects either a) the primary providers of upfront capital are social or commercial investors and there is limited risk to the service provider in the repayment structure, or b) if the

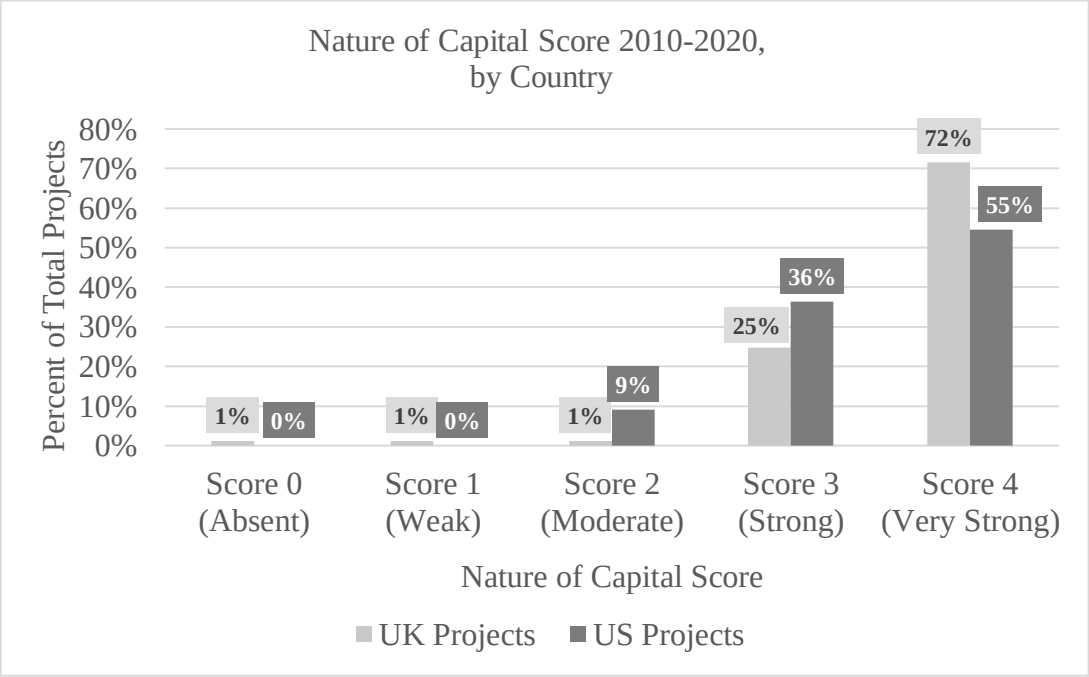
Figure 20: Nature of Capital Scores, All Projects



primary provider of upfront capital is philanthropy, there is no risk to the service provider in the repayment structure.

When US and UK SIB projects are compared, similar patterns appear but for different underlying reasons. In the UK, in over 90% of the projects the primary investor is classified as a social investor. In the US, however, primary investors are split between philanthropy (52%) and commercial investors (44%). However, in the UK it is more likely that investors share risk with provider partners. Majority of the UK projects (73%) involved partial provider risk share (compared to 50% of projects in the US); in the US, 36% of projects completely protected providers from performance risk, compared to just 1% of projects in the UK. Both of these country strategies yield nature of capital scores that signal in most SIB projects upfront capital is independent and at-risk. Figure 21 below shows the percentage of projects across each nature of the capital score for both UK and US projects (see also Figure 13 on page 144).

Figure 21: Nature of Capital Scores, by Country



Through this analysis we found that in the US it is more likely that the SIB model has been successful in shifting the status quo of service delivery to incentivize and pay for project impact. This is because in typical US SIB projects, robust evaluations validate outcomes before payments are actually made. In the UK, on the other hand, outcome payments are typically not subject to rigorous validation methods. Projects in this context might be paying for outcome metrics that don't meaningfully impact service users, or outcomes that would have happened anyway. Without a rigorous validation method, it is hard to say these projects are paying for project impact.

We have also found, in both country contexts, that investors in SIB projects have relied on different risk mitigation strategies, falling short of a full risk transfer away from government and provider partners. In the UK, social investors have served as the primary provider of upfront capital, but often rely on a risk-sharing agreement with the provider that is baked into the repayment structure. In the US, there is typically less risk-sharing with the provider, but commercial investors are often backed by philanthropic dollars as a way to mitigate risk. In both instances, investors of upfront capital, which were meant to bring new forms of market discipline and performance management, are at least partially shielded from the risk of project failure.

Strengths and Limitations

The methods used in this thesis included an in-depth case study of a SIB project launched in the UK, as well as qualitative content analysis of all SIB projects launched from 2010-2020 in the US and the UK. These qualitative methods are the most appropriate for the research design as they enabled the investigation of complex interactions characteristic of the SIB model, allowing for data gathering and analysis to occur simultaneously in a naturalistic setting (Harris et al. 2009). Through these qualitative methods I was able to gather rich data using an integrated approach, beginning with theory derived from observations of the SIB model (deductive), while allowing space for emerging empirical evidence to

inform theory development as data was collected (inductive) (Bryman 2015). Overall, these qualitative analysis methods allowed a mix of both depth (through the case study) and breadth (through the content analysis) to be prioritized in the research, in order to provide a more complete picture of how the SIB model has impacted various levels of service delivery.

The case study allowed for a rich, holistic, and context-driven exploration of how the SIB model impacted the system of public service delivery in Essex, where it was used to fund an intervention for vulnerable children and families. A key strength of the case study is that it allows for an in-depth examination of relationship dynamics and processes, providing a more holistic representation of complex phenomena compared with alternative methods (Bennett and Checkel 2014; Cartwright 2017). Relying on a case study allowed me to explore questions related to how, why, and for whom, the SIB model, has had an impact (Yin 2013; Pawson and Tilley 1997). In this vein, the case study was structed with guidance from a process evaluation framework, attempting to trace how the SIB model was implemented throughout the project, and the impact that the different features of the SIB had on service delivery. While typically a process evaluation seeks to determine how successfully an intervention was implemented using a clear theory of change (Steckler and Linnan 2002), the SIB model lacks a clearly defined theory of change. To address this, I laid out a conceptual framework that attempts to clarify the mechanisms through which a SIB may impact service delivery. These mechanisms can be improving or impeding, similar in concept to barriers and facilitators of successful intervention implementation (Schneider et al. 2009). Using interviews, document review, and participant observation, the case study was designed to uncover how the SIB model altered processes of service delivery (i.e. through which mechanisms), what the effect of those alterations were, and for whom, drawing on key principles of process evaluation (Pawson and Tilley 1997; Steckler and Linnan 2002). This framework is a key strength of the research design, providing a model that could be used in future research analyzing the SIB effect on systems of public service delivery.

A strength of this case study was the degree of project access for which I was granted to conduct my research. By developing warm contacts through a snowball method (Bryman 2015), I was able to interview a wide range of project stakeholders, including government leaders and agency staff, service providers and front-line staff, SIB investors, advisors, and intermediary staff. This allowed for a diversity of perspectives and a more complete analysis of how the SIB was impacting the system of children's social care in Essex. These insights were further complemented by an opportunity to shadow front-line staff during their service provision, which provided me with a richer understanding of the mechanics of the SIB-funded intervention, as well as how it was delivered within the wider system of children's services. In addition, I was given access to confidential documents, including investor meeting notes, contracts, and financial models, which, while specifics were not to be included for publication, allowed for greater insight into the processes, rationale, events, and relationship dynamics characteristic of the SIB project. Triangulating these methods, semi-structured interviews, participant observation, and document review, allowed for an enhanced understanding of the context and providers greater confidence in the research findings.

Conducting this case study research while the SIB was still being rolled out was another major strength of the SIB design. This allowed me to conduct participant observation by shadowing front-line staff, as discussed above, and also made it easier to connect with the full set of SIB stakeholders for interview. I conducted semi-structured interviews with 31 project stakeholders, which enabled a rich, descriptive account of individual's experiences with the SIB model and how they saw it impacting their own roles, as well as their insights on how the model was impacting the wider system. In line with the semi-structured approach (Bryman 2015), interviews were conducted flexibly and open-endedly; the conversation's emphasis would often shift based on participant's role or previous responses. Interviews were conducted individually with each informant, another strength of the research design, in order to help facilitate an environment where respondents could openly reflect on their experience and opinions of the SIB project.

The qualitative content analysis method, applied to 114 SIB projects, represents the first systematic attempt to map variation in design of SIBs against the espoused model. As evident through the case study design, SIBs are complex phenomena with an unending number of components, processes, relationship dynamics, and structures to consider. The qualitative content analysis approach allows us to reduce this material, impossible to analyze for all 114 projects within the constraints of this thesis, and focus on the aspects that most closely relate to the overall research question (Schreier 2014). Major strengths of this approach are that it is data-driven, highly systematic, and partly iterative - allowing for both theoretical and empirical contributions to be made. By creating a coding frame, the heart of this qualitative method (Schreier 2014), we have operationalized two key components of SIB design that have implications for the SIB effect on public services: payments linked to impact and the at-risk, independent nature of the upfront capital. This coding frame, a key strength of this work, provides an original operational framework from which to compare SIB design across time and geographies.

Furthermore, incorporating SIB projects from both the US and the UK in the empirical qualitative content analysis enabled a comparative lens to see how SIBs have been designed across the two countries over time. SIB projects from these two countries represent over 60% of all SIBs worldwide, providing a rich set of data points across time (2010-present) for analysis. This analysis, covering a majority of SIB projects launched to date, provides a degree of breadth to the findings (which complements the depth of the SIB case study).

Both the case study and qualitative content analysis designs are subject to limitations that are important to note. Qualitative research can be limited by concerns with subjectivity or a lack of generalizability. Case studies, in particular, raise concerns of objectivity and replicability, as they are based on the subjective interpretation of the researcher (Berg and Lune 2012). My current affiliation with the Government Outcomes Lab at the University of Oxford, which conducts research and provides guidance on SIB

projects), as well as my previous work experience providing technical assistance to government agencies pursued SIBs, has contributed to my knowledge base and influenced the assumptions I hold about the model. Acknowledging these experiences through a lens of reflexivity suggests that my own beliefs and values have likely influenced the research process and outcomes reported in this thesis (Berger 2015). Efforts have been taken throughout the process of this thesis to mitigate potential bias, including offering up the methodologies, analysis, findings and discussion to validation and feedback from other experts in the SIB field at multiple points over the course of the research.

The qualitative methods conducted as part of the case study, including interviews, participant observation, and documentary review, are also subject to a number of weaknesses. Interviewee responses and participant behavior can be biased by the presence of the researcher (Bryman 2015), and participants may have felt the need to self-censor when discussing sensitive aspects of the SIB project, perhaps including relationship tensions between and within stakeholder groups, pressures to meet outcome goals, enrollment or referral strategies, or other topics they felt may have cast them or other stakeholders in a negative light. Conclusions drawn from single case studies are often difficult to generalize outside of the specific context, as there may be different relationships present depending on the case (Gerring 2004; Berg and Lune 2012). Many efforts were taken to mitigate these limitations, including conducting interviews individually and offering all respondents to be anonymized, triangulating data from multiple sources, and validating processes and findings with external experts in the field.

The qualitative content analysis conducted as part of this thesis is also subject to limitations. Like many methods of qualitative analysis, the construction of a new coding frame relies on the interpretation of the researcher(s) and therefore risks a degree of subjectivity. The creation of the coding frame was an iterative process, during which the co-authors and I presented it for critique and validation from field experts, in an effort to mitigate concerns with subjectivity. The content analysis is also, by design, narrow in nature as it focuses on only two dimensions of SIB design. While these two dimensions of SIB

design are arguably the most relevant considering the research question, the analysis may fail to illuminate the full context of the SIB model's evolution. Finally, the analysis includes only projects that have been launched in the US and the UK. The story of how the SIB has evolved and been applied in these two countries may differ substantively from other countries where the SIB model is growing in application.

Contribution of the Thesis

This thesis contributes to the debate around social impact bonds, a model governments have been increasingly using to fund social services that aim to support vulnerable populations. Hybridity in social service delivery, characterized by public partnerships with private and third sector organizations, have long been put forward as a more effective approach to delivering public services (Skelcher 2007). In the 1980s, New Public Management (NPM), a hybrid approach focused on incorporating private sector elements to improve the efficiency of public services, gained traction across the world (Hood 1991). In many ways, the SIB model can be viewed as an extension of the NPM approach, and questions remain regarding what the 'SIB effect' is across various levels of service delivery.

Examining the 'SIB effect' relates to a number of wider public management questions in the academic literature, including collaborative hybrid-governance structures, front-line staff discretion, and the use of financial incentives in traditionally public social services. Contributions are also relevant to practitioners considering or using the SIB model, particularly pragmatic findings around SIB design and implementation that can be used to provide more effective (and equitable) social support. Over the course of this thesis, I have offered three articles that aim to address various gaps in our current understanding of how SIBs work in practice and the impact they have on service delivery systems. As such, I have made significant theoretical and empirical contributions to the literature by:

1. **Providing an original conceptual framework from which to analyze the effect of SIBs on systems of public service delivery**

   Article one of this thesis offers a new framework for analyzing the impact of the SIB model, which lacks a clear theory of change, on the process of public service delivery. Building on extant literature covered in the *Literature Review* chapter, the framework stems from the four key features of the SIB model and suggests potential mechanisms through which these features might improve or impede public service delivery. I argue that this framework is a substantive contribution to the field because it enables thorough investigation of contradictory predictions in the literature on the likely effects of the SIB model, and aids in our understanding of what features of the SIB model are likely to lead to improvements or impediments in public service delivery.

2. **Laying out the conceptual arguments for how the SIB model could affect provider and front-line staff, and extending this conversation to include differences between types of interventions**

   Article two of this thesis categorizes the conceptual arguments for how the SIB model could affect providers and front-line staff, and extends this conversation to include the differences between black-box and high-fidelity intervention types. Many of the arguments around how SIBs might influence public service delivery or intervention outcomes rely on behavior and/or attitude changes among front-line staff in the service provider organization. I argue that hypothesizing (and then empirically testing) how these effects on front-line staff might differ under a high-fidelity intervention is a significant contribution to the literature that provides insight on how the effects of the SIB model might vary under different policy contexts. This discussion adds to the growing body of knowledge that has sought to unpack the effect of the SIB model on front-line staff.

3. **Furthering the field's empirical understanding of how the SIB model impacts public service delivery and front-line staff through an in-depth case study, surfacing a unique and original discussion of the risks of disproportionality**

   Through the case study of the Essex SIB, examined in articles one and two, I have provided a rich, context-driven, and holistic exploration of how the SIB model impacted both the system of public service delivery and front-line staff. Conducting this in-depth examination of the relationship dynamics, events, and processes of the Essex case has yielded meaningful empirical findings related to the use of the SIB model to fund public services. These findings add to a growing body of knowledge on various aspects of the 'SIB effect,' related to financial incentives, performance targets, democratic accountability, multi-sector collaboration, and provider flexibility. The case study has also surfaced the risk of disproportionality in SIB projects, highlighting a key challenge that has yet to be thoroughly addressed in the literature.

4. **Constructing an original operationalized framework from which to compare the design of SIB projects across time and geographies**

   In article three, the co-authors and I provide an original operational framework to compare SIB project design across time and geographies. We constructed a coding frame through which to operationalize arguably the two most consequential components of SIB design: payments made for impact and independent, at-risk capital. We argue that the coding frame is a substantive theoretical contribution to the literature as it allows for project categorization and comparison, in order to draw conclusions on the evolution of the SIB model over time. It can help scholars pinpoint if and how SIB projects, in practice, are designed in a way that is meaningfully different from alternative methods of service delivery, enabling discussions around whether SIB projects are living up to their original promises.

5. **Offering the first systematic investigation of SIB design over time, empirically comparing the SIB model in practice to its original promises**

   Article three of this thesis represents the first systematic attempt to map variation in the design of SIB projects against the espoused model, as the co-authors and I empirically analyze the range of different SIB configurations tested in the US and UK (n=114) to explore how SIBs have been designed in practice. The potential for ambiguity and exploitation in the SIB model highlights the importance of empirically analyzing how SIBs have been designed across contexts in relation to the model's key promises. The findings from this empirical analysis represent a significant contribution to the field, demonstrating that SIBs have indeed "stretched" from what makes the model a distinctive, innovative approach in some cases, potentially undermining its potency as a tool for service delivery improvement.

Policy Implications and Further Research

This thesis has implications that are useful for practical application by policymakers. Below I offer implications for the SIB model as a policy option (i.e. for policymakers deciding between commissioning approaches) and implications for SIB design (i.e. for policymakers who have decided to use the SIB model).

Implications for the SIB model as a policy option

The SIB model is one of many tools that governments could use for commissioning and delivering social services, and it is helpful to reflect on why and under what contexts (if any) a policymaker should pick this tool. What value-add does the SIB bring to the various partners (government, provider, and service user) that are traditionally involved in service delivery? Under what contextual circumstances might it be a good idea for government agencies to pursue the SIB model? This thesis has found that the SIB model can bring cross-sector attention and resources to support a specific cohort and work with them towards particular outcomes. This outcome-aligned approach, paired with more intensive performance monitoring brought via

forms of market discipline, has the potential to shift the focus of service delivery away from activities and towards outputs and outcomes. In addition, the SIB model can be used to add an entirely new service to the service array, particularly if there are budgetary or political limitations that prevent new services from being added outside of the SIB model.

Conversely, reflecting on the drawbacks of the SIB model allows us to ask, 'does using the SIB model have negative consequences for government, provider, or service user partners'? This thesis has found that the SIB model might contribute to a sense of disproportionality from the government partner, as the model has traditionally required much time and resources for relatively small cohorts. Relatedly, SIB projects can be an expensive way of commissioning services, with high transaction costs as well as high ongoing management costs (Edmiston and Nicholls 2018). Government partners may experience a loss of control over service direction, as was reported in Essex, and provider partners might experience an increased administrative burden or added pressures to meet SIB-related targets.

Maier, Barbetta, and Godina (2018) argue that the collaborative element of SIBs has the potential to "improve knowledge flows" about promising and proven social programs. The multi-sector partnership can foster collaborative relationships, leading to joint-working and the potential for new ideas to improve service delivery (Nicholls and Murdock 2012; Liang, Mansberger, and Spieler 2014; Arena et al. 2016)
The Essex case, however, demonstrated that SIB projects are not always characterized by collaborative partnerships. Instead, in Essex, there were difficulties integrating the SIB project into the wider system, challenges with the governance structure, relationship tensions, and complex layers of managerial oversight. Many instances of culture clash between public, private, and voluntary sector ways of working occurred throughout the course of the project, limiting the potential for collaborative relationships to improve service delivery.

The potential drawbacks of the SIB model are particularly relevant if they remain despite projects being designed in a way that is not meaningfully different from status-quo service delivery. This thesis has found that the SIB model has stretched from its original intentions in a few critical ways, potentially rendering the model more similar to alternative forms of commissioning service delivery than originally thought. For example, in the UK, payable outcomes are generally not subject to rigorous validation methods, calling in to question whether payments are meaningfully linked to impact. Early supporters of the model argued that SIBs could help to achieve impact with greater efficiency than other forms of service provision by "allocating resources to where they will achieve the most impact," (Fox and Albertson 2012). More recently, Ronicle and Smith (2020) found that an increased focus on achieving outcomes was one of "the strongest effects" of the SIB model, enabling service providers to adapt in efforts to better achieve those outcomes. If these outcomes aren't linked to impact, or do not meaningfully improve the lives of beneficiaries, this might "unravel the core logic of the SIB mechanism" (Carter 2020).

Policymakers may instead be attracted to the SIB model for the potential to shift risk of project failure away from government and service providers on to private investors. External investors who bear the risk that the SIB project does not perform is perhaps one of the most distinguishing features of the SIB model (Fraser et al. 2016). This thesis has found that in both the US and the UK, upfront investment is typically independent and at-risk. However, certain risk mitigation strategies may limit the intended transfer of risk to external investors. If performance risk is not transferred to external investors, this could further diminish the distinguishing factors between the SIB model and other forms of commissioning. For example, if performance risk sits with the government, it renders the model similar to fee-for-service, if performance risk sits with the service provider it's similar to payment by results, and if performance risk sits with philanthropy it could resemble a grant. All of these configurations raise questions around why upfront commercial or social investment, and the associated returns, would be needed. If there is limited risk-transfer, paired with concerns around payment linked to impact, policymakers may question what value the model will bring to the government partner.

Findings from this thesis can also shed some light on the decision to use the SIB model in the context of a high-fidelity intervention. In the case of Essex, the flexibility promised by the SIB model was limited, due in large part to the use of a high-fidelity intervention. This resonates with arguments from Maier, Barbetta, and Godina (2018), who point out the paradoxical nature of the SIB model that promises to be both evidence-based and flexible. While flexibility was limited, evidence from Essex suggests that the SIB model may have brought increased accountability and performance management than would have happened otherwise. Other SIB projects have reported this type of non-financial support for the management and delivery of services as a benefit of the SIB model as well (Fraser et al. 2020). It is up to the policymaker, then, to determine if the increased performance management structure is worth the large transaction costs and other potential drawbacks of the SIB model (Warner 2013; Giacomantonio 2017).

Implications for SIB design

Over the past ten years, more than 200 SIB projects have been launched around the world. As policymakers continue to experiment with the model, lessons from previous projects can help inform the design and structure of future SIBs to maximize the model's potential to improve service delivery while minimizing the model's potential to impede service delivery. This thesis has uncovered lessons from the Essex case study that may be useful to policymakers in the SIB design process, including the importance of developing relationships, the risk of disproportionality, the need for flexibility to safeguard democratic accountability, and the implications of incomplete outcomes specifications.

Developing strong, collaborative stakeholder relationships can help to increase the chances that the SIB model improves public service delivery, rather than impedes it. As SIBs involve bringing together diverse partners who may have different approaches or worldviews, conflict can arise if relationships aren't sufficiently developed. In the case of Essex, relational tensions throughout the project resulted in clashes between stakeholders and difficulties integrating the service in the wider system. The "complex inter-

relationships" inherent in SIBs require a large amount of institutional work (Lowe et al. 2019); it is often difficult to integrate SIB-funded projects with the wider system and array of other services offered, and communicative relationships are key (Ronicle and Smith 2020).

SIB projects may also be subject to risks of disproportionality; as more time and resources are spent on a relatively small cohort, less are available for other vulnerable populations. Government leadership in Essex expressed frustration because the SIB funded services for a small proportion of young people, yet was very demanding of the overall social care system's time and resources. This raises a question about the equity of SIB projects in general, and highlights the importance of mitigating disproportionality in future iterations of the model. This finding also mirrors conclusions from previous studies, which suggest that SIBs may bring additional layers of bureaucracy (Roberts and Cameron 2014; Griffiths, Thomas, and Pemberton 2016) and high degrees of administrative burden that, exacerbated by duplicative systems of performance management, can lead to an inefficient use of resources (Ronicle and Smith 2020). Dowling (2017) emphasizes the inherent narrow focus of SIBs, as the projects tend to target individuals at the symptom-level rather than at a structural level. The narrow focus of SIB projects, paired with intense time, effort, and monetary requirements of stakeholders across sectors, signal a predominant risk of disproportionality.

When designing a SIB project, it is important to build in flexibility in order to safeguard democratic accountability. In Essex for example, over the course of the SIB project, the context of children's social care changed dramatically – with fewer children entering care, and the priority population shifting. Despite desires to shift the service direction in response to these changes, contract inflexibility meant that government agency lost their ability to switch to an intervention that would "provide the right outcomes at the right level and at the right price" (Neyland 2018). This loss of control over the service direction may have contributed to a loss of democratic accountability, which resonates with critiques of the SIB model and other NPM strategies (Warner 2013). By allowing investor or provider partners the discretion to "shape social policy according to their interests," which include a financial gain, SIBs contribute to a "process of

de-democratization" that takes power away from the state (Dowling 2017). This consideration, retaining discretion and control over the service discretion, is relevant to policymakers designing SIB projects.

Finally, the Essex case also demonstrated a lesson in the implications of incomplete outcome specifications in the SIB model. In Essex, the eligible cohort was not clearly defined and there is some concern over the accuracy of the benchmark used to trigger outcome payments. Outcome prices were set based on a historical benchmark during a time of very poor practice, perhaps making it too easy to achieve outcome payments in future years. Neyland (2018) notes that in Essex, the "business as usual counterfactual" that payments were based "depended on a 'business' (sending large numbers of children into residential care) that was no longer 'usual'." Tse and Warner (2018) argue that SIBs in general can lead to overpayment by narrowing the conversation to financial metrics, "extracting funds" from the same public sectors that "SIBs were designed to support." Carefully specifying payable outcomes has important implications for the value government agencies stand to gain (or lose) from using the SIB model.

Future research

Perhaps the most pervasive gap in the existing empirical evidence on the SIB model is its effect on service users. Extending the frameworks offered in this thesis for how the SIB might affect the system of public service delivery and front-line staff to understand how these alterations reach and impact the intended beneficiaries is one potential avenue forward. For example, perhaps the multi-sector partnership characteristic of the SIB model allows front-line staff to be better supported, and they in turn are able to give the service user a more comprehensive, positive experience of the intervention. Or, on the other hand, the payment for outcomes contract could lead to a narrowing of service provision that leaves service users feeling less holistically supported compared to receiving the same intervention without SIB funding. Understanding how the SIB impacts the experiences and outcomes of service users is an important piece of

the puzzle that is missing. Further work should be done to interrogate the differences (if any) to the service user experience when the SIB model is used as the commissioning approach.

Further research might also compare the impact of the SIB model on front-line staff under different commissioning contexts and to non-SIB alternatives. Previous scholars have found that using a SIB model with a black-box approach can be problematic in regard to outcome gaming (for example, focusing on portions of the target population that might more easily achieve outcomes). This thesis, using the Essex case study, suggests that using a SIB with a high-fidelity intervention can also raise concerns, notably limiting flexibility and leaving potential for output target-related disincentives in service delivery. Further work might explore and compare the impact of various commissioning approaches (fee-for-service, payment for results, and social impact bonds) on front-line discretion and incentives.

The role of the external investors, one of the critical defining features of the SIB model, should also be explored in further research. This thesis has found that many SIBs have employed various risk mitigation strategies that limit the transfer of risk on to the investors involved in SIB projects. Further interrogating what purpose the investor serves, and what value this function might add, would contribute to policymaker decisions around when (and if) to use the SIB model. Lessons from the Essex case study in this thesis suggest that the external investors brought a more intensive performance management structure, atypical of what the government agency and service normally use to track and improve project performance. This may look different in other SIB projects, depending on the nature of the investor. It would be helpful for the field to further interrogate the value-add of the investor partner, in order for policymakers to make informed decisions about the trade-offs (i.e. high transaction costs, loss of service direction, etc.) required when involving such a partner.

Finally, future research should further explore variation in SIB design, seeking to pinpoint what value-add or value-detract various configurations of the SIB model bring. This thesis provided an empirical analysis

of SIB design across two dimensions, payments linked to impact and nature of the capital, focused on SIBs launched in the US and the UK. As more SIB projects are being launched around the world, particularly in non-Western countries, expanding this narrow focus to evaluate how SIBs have been designed in other countries, and for what purpose, might provide clarity on future uses of the model.

References

Albertson, Kevin, Chris Fox, Chris O'Leary, Gary Painter, Kimberly Bailey, and Jessica Labarbera. 2018. Payment by Results and Social Impact Bonds: Outcome-Based Payment Systems in the UK and US. 1st ed. Bristol University Press. https://doi.org/10.2307/j.ctt21h4zbp.

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. https://doi.org/10.1080/01900692.2015.1057852.

Armstrong, Anona. 1998. "A Comparative Analysis: New Public Management — The Way Ahead?" Australian Journal of Public Administration 57 (2): 12–24. https://doi.org/10.1111/j.1467-8500.1998.tb01379.x.

Bafford, Beth. 2012. "The Feasibility and Future of Social Impact Bonds in the United States." Stanford Journal of Public Policy 3 (1):3-19.

Baliga, Shifali. 2013. "Shaping the Success of Social Impact Bonds in the United States: Lessons Learned from the Privatization of U.S. Prisons." Duke Law Journal 63: 437–79.

Balk, Walter L. 1996. Managerial Reform and Professional Empowerment in the Public Service. 0th Edition. Westport, Conn: Praeger.

Barajas, Adriana, Liliana Barajas, Kiyadh Burt, Toyan Haper, Patrick Johnson, Erika Larsen, Jacobo Licona, et al. 2014. "Social Impact Bonds: A New Tool for Social Financing." Princeton, New Jersey: Princeton University. https://spia.princeton.edu/sites/default/files/content/Social%20Impact%20Bonds%202014%20Final%20Report.pdf.

Barberis, Peter. 1998. "The New Public Management And A New Accountability." Public Administration 76 (3): 451–70. https://doi.org/10.1111/1467-9299.00111.

Battye, Fraser. 2015. "Payment by Results in the UK: Progress to Date and Future Directions for Evaluation." Evaluation 21 (2): 189–203. https://doi.org/10.1177/1356389015577464.

Bauer, Mark S., Laura Damschroder, Hildi Hagedorn, Jeffrey Smith, and Amy M. Kilbourne. 2015. "An Introduction to Implementation Science for the Non-Specialist." BMC Psychology 3 (1). https://doi.org/10.1186/s40359-015-0089-9.

Bennett, Andrew, and Jeffrey Checkel. 2014. Process Tracing: From Metaphor To Analytic Tool. Cambridge: Cambridge University Press.

Berg, Bruce, and Howard Lune. 2012. Qualitative Research Methods for the Social Sciences. 8th ed. Pearson.

Berger, Roni. 2015. "Now I See It, Now I Don't: Researcher's Position and Reflexivity in Qualitative Research." Qualitative Research 15 (2): 219–34. https://doi.org/10.1177/1468794112468475.

Blakely, Craig H., Jeffrey P. Mayer, Rand G. Gottschalk, Neal Schmitt, William S. Davidson, David B. Roitman, and James G. Emshoff. 1987. "The Fidelity-Adaptation Debate: Implications for the Implementation of Public Sector Social Programs." American Journal of Community Psychology 15 (3): 253–68. https://doi.org/10.1007/BF00922697.

Bode, Ingo, and Taco Brandsen. 2014. “State–Third Sector Partnerships: A Short Overview of Key Issues in the Debate.” Public Management Review 16 (8): 1055–66. https://doi.org/10.1080/14719037.2014.957344.

Borys, Bryan, and David B. Jemison. 1989. “Hybrid Arrangements as Strategic Alliances: Theoretical Issues in Organizational Combinations.” The Academy of Management Review 14 (2): 234–49. https://doi.org/10.2307/258418.

Bovaird, Tony, and Rhodri Davies. 2011. “Outcome-Based Service Commissioning and Delivery: Does It Make a Difference?” In Research in Public Policy Analysis and Management, 21:93–114. https://doi.org/10.1108/S0732-1317(2011)0000021011.

Boyle, David. 2011. “The Pitfalls and Perils of Payment by Results.” Local Economy 26 (8): 627–34. https://doi.org/10.1177/0269094211422203.

Boyne, George A., and Jennifer Law. 2005. “Setting Public Service Outcome Targets: Lessons from Local Public Service Agreements.” Public Money & Management 25 (4): 253–60. https://doi.org/10.1080/09540962.2005.10600128.

Breitenstein, Susan M., Deborah Gross, Christine A. Garvey, Carri Hill, Louis Fogg, and Barbara Resnick. 2010. “Implementation Fidelity in Community-Based Interventions.” Research in Nursing & Health 33 (2): 164–73. https://doi.org/10.1002/nur.20373.

Brodkin, Evelyn Z. 2012. “Reflections on Street-Level Bureaucracy: Past, Present, and Future.” Public Administration Review 72 (6): 940–49. https://doi.org/10.1111/j.1540-6210.2012.02657.x.

Bruijn, J. A. de. 2002. Managing Performance in the Public Sector. Psychology Press.

Bryman, Alan. 2015. Social Research Methods. Fifth Edition. Oxford: Oxford University Press.

BSC. 2021. “Big Society Capital.” Big Society Capital. 2021. https://bigsocietycapital.com/.

Burand, Deborah. 2013. “Globalizing Social Finance: How Social Impact Bonds and Social Impact Performance Guarantees Can Scale Development.” N. Y. U. J. L. & Bus 9 (January): 447–502.

Butler, David, Dan Bloom, and Timothy Rudd. 2013. “Using Social Impact Bonds to Spur Innovation, Knowledge Building, and Accountability.” Community Development Investment Review, 6.

Cabinet Office. 2012. “Launch of Big Society Capital – the World’s First Ever Social Investment Market Builder.” GOV.UK. April 4, 2012. https://www.gov.uk/government/news/launch-of-big-society-capital-the-world-s-first-ever-social-investment-market-builder.

Carroll, Christopher, Malcolm Patterson, Stephen Wood, Andrew Booth, Jo Rick, and Shashi Balain. 2007. “A Conceptual Framework for Implementation Fidelity.” Implementation Science 2 (1): 40. https://doi.org/10.1186/1748-5908-2-40.

Carter, Eleanor. 2020. “Debate: Would a Social Impact Bond by Any Other Name Smell as Sweet? Stretching the Model and Why It Might Matter.” Public Money & Management 40 (3): 183–85. https://doi.org/10.1080/09540962.2020.1714288.

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.” Government Outcomes Lab, University of Oxford. https://golab.bsg.ox.ac.uk/knowledge-bank/resources/evidence-report/.

Carter, Eleanor, and Adam Whitworth. 2015. “Creaming and Parking in Quasi-Marketised Welfare-to-Work Schemes: Designed Out Of or Designed In to the UK Work Programme?” Journal of Social Policy 44 (2): 277–96. https://doi.org/10.1017/S0047279414000841.

Carter, Eleanor, and Adam Whitworth. 2017. “Work Activation Regimes and Well-Being of Unemployed People: Rhetoric, Risk and Reality of Quasi-Marketization in the UK Work Programme.” Social Policy & Administration 51 (5): 796–816. https://doi.org/10.1111/spol.12206.

Cartwright, Nancy. 2017. “Single Case Causes: What Is Evidence and Why.” In: Chao HK., Reiss J. (eds) Philosophy of Science in Practice. Synthese Library (Studies in Epistemology, Logic, Methodology, and Philosophy of Science), vol 379. Springer, Cham. https://doi.org/10.1007/978-3-319-45532-7_2

Castro, Felipe González, Manuel Barrera, and Charles R. Martinez. 2004. “The Cultural Adaptation of Prevention Interventions: Resolving Tensions Between Fidelity and Fit.” Prevention Science 5 (1): 41–45. https://doi.org/10.1023/B:PREV.0000013980.12412.cd.

Chiapello, Eve, and Lisa Knoll. 2020. “The Welfare Conventions Approach: A Comparative Perspective on Social Impact Bonds.” Journal of Comparative Policy Analysis: Research and Practice 22 (2): 100–115. https://doi.org/10.1080/13876988.2019.1695965.

Clist, Paul, and Stefan Dercon. 2014. “12 Principles for Payment By Results (PbR) In International Development,” GOV.UK. https://assets.publishing.service.gov.uk/media/57a089d2e5274a27b20002a5/clist-dercon-PbR.pdf.

CNCS. 2017. “Federal Social Innovation Fund Celebrates $1 Billion Milestone.” Corporation for National and Community Service, AmeriCorps. February 24, 2017. https://www.nationalservice.gov/newsroom/press-releases/2017/federal-social-innovation-fund-celebrates-1-billion-milestone.

Cohen, Sir Ronald. 2011. “Harnessing Social Entrepreneurship and Investment to Bridge the Social Divide.” Paper given at the EU conference on the social economy, Brussels, Belgium, 18 November, http://www.ronaldcohen.org/sites/default/files/3/Harnessing%20entrepreneurship%20EU%20Nov%202011%20FINAL.pdf

Considine, Mark, Jenny M. Lewis, and Siobhan O’Sullivan. 2011. “Quasi-Markets and Service Delivery Flexibility Following a Decade of Employment Assistance Reform in Australia.” Journal of Social Policy 40 (04): 811–33. https://doi.org/10.1017/S0047279411000213.

Cooper, Christine, Cameron Graham, and Darlene Himick. 2016. “Social Impact Bonds: The Securitization of the Homeless.” Accounting, Organizations and Society 55 (November): 63–82. https://doi.org/10.1016/j.aos.2016.10.003.

Corrigan, P. 2011. “Social Impact Bonds. A New Way to Invest in Better Healthcare.” London: Social Finance Ltd.

Dayson, Chris, Alec Fraser, and Toby Lowe. 2020. “A Comparative Analysis of Social Impact Bond and Conventional Financing Approaches to Health Service Commissioning in England: The Case of Social Prescribing.” Journal of Comparative Policy Analysis: Research and Practice 22 (2): 153–69. https://doi.org/10.1080/13876988.2019.1643614.

DCLG. 2017. “The Impact Evaluation of the London Homelessness Social Impact Bond,” London: Department for Communities and Local Government.

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/658881/SIB_Impact_evaluation_report.pdf

Demel, Ana. 2012. “Second Thoughts on Social Impact Bonds Articles and Comments Arising from the 2012 Fall Conference: The Law & Finance of Social Enterprise.” New York University Journal of Law and Business 9: 503–10.

Disley, Emma, Chris Giacomantonio, Kristy Kruithof, and Megan Sim. 2015. The Payment by Results Social Impact Bond Pilot at HMP Peterborough: Final Process Evaluation Report. RAND Corporation. https://doi.org/10.7249/RR1212.

Disley, Emma, Jennifer Rubin, Emily Scraggs, Nina Burrowes, and Deirdre May Culley. 2011. “Lessons Learned from the Planning and Early Implementation of the Social Impact Bond at HMP Peterborough,” Cambridge: RAND.

Dixon, Ruth, and Franziska Rosenbach. 2020. “Is Risk-Transfer Largely Illusory?” The Government Outcomes Lab. Is Risk-Transfer Largely Illusory? (blog). October 23, 2020. /community/blogs/risk-transfer-largely-illusory/.

Doig, A. 1997. “Truth-Telling and Power.” Parliamentary Affairs 50 (1): 143–65. https://doi.org/10.1093/oxfordjournals.pa.a028710.

Dowling, Emma. 2017. “In the Wake of Austerity: Social Impact Bonds and the Financialisation of the Welfare State in Britain.” New Political Economy 22 (3): 294–310. https://doi.org/10.1080/13563467.2017.1232709.

Dowling, Emma, and David Harvie. 2014. “Harnessing the Social: State, Crisis and (Big) Society.” Sociology 48 (5): 869–86. https://doi.org/10.1177/0038038514539060.

Drew, Helen, Lisa Holmes, Valerie Dunn, and Neil Harrison. 2019. “Evaluation of the Multisystemic Therapy Service in Essex.” Oxford: Rees Center.

Ebrahim, Alnoor. 2019. Measuring Social Change: Performance and Accountability in a Complex World. 1st edition. Stanford, California: Stanford Business Books.

ECC. 2012. “Award of Contract for Multi-Systemic Therapy via Social Investment.” Cabinet, Essex County Council. https://cmis.essexcc.gov.uk/essexcmis5/DocumentSearch.aspx.

ECC (Essex County Council) Head of Children’s Services. In-person Interview, November 2018.

ECC (Essex County Council) Head of Partnership. Phone Interview, August 2018.

ECC (Essex County Council) Project Lead. In-person Interview, August 2018.

ECC (Essex County Council) Service Manager. Phone Interview, August 2018.

ECC (Essex County Council) Strategy and Innovation Manager. Phone Interview, December 2018.

Edmiston, Daniel, and Alex Nicholls. 2018. “Social Impact Bonds: The Role of Private Capital in Outcome-Based Commissioning.” Journal of Social Policy 47 (01): 57–76. https://doi.org/10.1017/S0047279417000125.

Esping-Andersen, Gøsta. 1990. The Three Worlds of Welfare Capitalism. Princeton, New Jersey: Princeton University Press.

Evans, Helen. 2020. “Variation in the Emergence and Growth of Social Impact Bond Markets in Advanced Capitalist Countries: A Comparative Study.” Presented at the Social Outcomes Conference 2020, University of Oxford, Government Outcomes Lab.

Evans, Tony. 2011. “Professionals, Managers and Discretion: Critiquing Street-Level Bureaucracy.” The British Journal of Social Work 41 (2): 368–86. https://doi.org/10.1093/bjsw/bcq074.

Ferlie, Ewan, Lynn Ashburner, Andrew Pettigrew, and Louise Fitzgerald. 1996. The New Public Management in Action. Oxford University Press.

Finn, D. 2011. “The Design of the Work Programme in International Context.” National Audit Office.

Fischer, Robert L., and Francisca García-Cobián Richter. 2017. “SROI in the Pay for Success Context: Are They at Odds?” Evaluation and Program Planning 64 (October): 105–9. https://doi.org/10.1016/j.evalprogplan.2016.11.012.

FitzGerald, Clare, Eleanor Carter, Ruth Dixon, and Mara Airoldi. 2019. “Walking the Contractual Tightrope: A Transaction Cost Economics Perspective on Social Impact Bonds.” Public Money & Management 39 (7): 458–67. https://doi.org/10.1080/09540962.2019.1583889.

Fonagy, Peter, Stephen Butler, David Cottrell, Stephen Scott, Stephen Pilling, Ivan Eisler, Peter Fuggle, et al. 2018. “Multisystemic Therapy versus Management as Usual in the Treatment of Adolescent Antisocial Behaviour (START): A Pragmatic, Randomised Controlled, Superiority Trial.” The Lancet Psychiatry 5 (2): 119–33. https://doi.org/10.1016/S2215-0366(18)30001-4.

Fosnot, Catherine Twomey. 2013. Constructivism: Theory, Perspectives, and Practice, Second Edition. Teachers College Press.

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 & Criminal Justice 11 (5): 395–413. https://doi.org/10.1177/1748895811415580.

Fox, Chris, and Kevin Albertson. 2012. “Is Payment by Results the Most Efficient Way to Address the Challenges Faced by the Criminal Justice Sector?” Probation Journal 59 (4): 355–73. https://doi.org/10.1177/0264550512458473.

Fox, Chris, and Stephen Morris. 2019. “Evaluating Outcome-Based Payment Programmes: Challenges for Evidence-Based Policy.” Journal of Economic Policy Reform: 1–17. https://doi.org/10.1080/17487870.2019.1575217.

Fraser, Alec, and Nicholas Mays. 2020. “Case Studies.” In Qualitative Research in Health Care, 4th ed. Wiley. https://www.wiley.com/en-us/Qualitative+Research+in+Health+Care%2C+4th+Edition-p-9781119410836.

Fraser, Alec, Stefanie Tan, Annette Boaz, and Nicholas Mays. 2020. “Backing What Works? Social Impact Bonds and Evidence-Informed Policy and Practice,” Public Money and Management 40 (3): 195-204. https://doi.org/10.1080/09540962.2020.1714303

Fraser, Alec, Stefanie Tan, Kristy Kruithof, Megan Sim, Emma Disley, Chris Giacomantonio, Mylene Lagarde, Jennifer Rubin, and Nicholas Mays. 2018. “Evaluation of the Social Impact Bond Trailblazers in Health and Social Care: Final Report.” https://www.rand.org/pubs/external_publications/EP67664.html.

Fraser, Alec, Stefanie Tan, Mylene Lagarde, and Nicholas Mays. 2016. "Narratives of Promise, Narratives of Caution: A Review of the Literature on Social Impact Bonds." Social Policy & Administration 52 (1): 4–28. https://doi.org/10.1111/spol.12260.

Gerring, John. 2004. "What Is a Case Study and What Is It Good For?" American Political Science Review 98 (2): 341–54. https://doi.org/10.1017/S0003055404001182.

Gertler, Paul J, Sebastian Martinez, Patrick Premand, Laura B Rawlings, and Christel M J Vermeersch. 2016. "Impact Evaluation in Practice, Second Edition." International Bank for Reconstruction and Development: World Bank Group. https://openknowledge.worldbank.org/bitstream/handle/10986/25030/9781464807794.pdf?sequence=2&isAllowed=y.

Giacomantonio, Chris. 2017. "Grant-Maximizing but Not Money-Making: A Simple Decision-Tree Analysis for Social Impact Bonds." Journal of Social Entrepreneurship 8 (1): 47–66. https://doi.org/10.1080/19420676.2016.1271348.

Gilson, Lucy. 2015. "Lipsky's Street Level Bureaucracy." In Oxford Handbook of the Classics of Public Policy, 24. Oxford: Oxford University Press.

GO Lab. 2019. "Case Study: Essex County Council MST." The Government Outcomes Lab. 2019. https://golab.bsg.ox.ac.uk/knowledge-bank/case-studies/essex-county-council-mst/.

GO Lab. 2020a. "Outcomes-Based Contracting/Results-Based Financing." The Government Outcomes Lab. 2020. https://golab.bsg.ox.ac.uk/the-basics/innovative-partnerships/.

GO Lab. 2020b. "UK Government Outcomes Funds for SIBs." The Government Outcomes Lab. 2020. https://golab.bsg.ox.ac.uk/basics/outcomes-funds/.

GOV.UK. 2011. "Children Looked after in England Including Adoption: 2011 to 2012." GOV.UK. 2011. https://www.gov.uk/government/statistics/children-looked-after-by-local-authorities-in-england-including-adoption.

GPL. 2017. "Active Contract Management: How Governments Can Collaborate More Effectively with Social Service Providers to Achieve Better Results." Cambridge, MA: Harvard Kennedy School Government Performance Lab. https://govlab.hks.harvard.edu/files/govlabs/files/active_contract_management_brief.pdf.

Gray, Mia, and Anna Barford. 2018. "The Depths of the Cuts: The Uneven Geography of Local Government Austerity." Cambridge Journal of Regions, Economy and Society 11 (3): 541–63. https://doi.org/10.1093/cjres/rsy019.

Griffiths, Ann, and Christian Meinicke. 2014. "Introduction to Social Impact Bonds and Early Intervention." Early Intervention Foundation. https://www.eif.org.uk/resource/social-impact-bonds-and-early-intervention.

Griffiths, Rita, Andy Thomas, and Alison Pemberton. 2016. Qualitative Evaluation of the DWP Innovation Fund: Final Report. https://nls.ldls.org.uk/welcome.html?ark:/81055/vdc_100035988353.0x000001.

Gruening, G. 2001. "Origin and Theoretical Basis of New Public Management." International Public Management Journal 4 (1): 1–25. https://doi.org/10.1016/S1096-7494(01)00041-1.

Guba, Egon G., and Yvonna S. Lincoln. 1994. “Competing Paradigms in Qualitative Research.” In Handbook of Qualitative Research, 105–17. Thousand Oaks, CA, US: Sage Publications, Inc.

Gustafsson-Wright, Emily, Sarah Osborne, and Meg Massey. 2020. “Are Impact Bonds Delivering Outcomes and Paying Out Returns?” Washington, D.C.: The Brookings Institute. https://www.brookings.edu/research/are-impact-bonds-delivering-outcomes-and-paying-out-returns

Harris, Jeffrey, Philip Gleason, Patricia Sheean, Carol Boushey, Judith Beto, and Barbara Bruemmer. 2009. “An Introduction to Qualitative Research for Food and Nutrition Professionals.” Journal of the American Dietetic Association 109 (February): 80–90. https://doi.org/10.1016/j.jada.2008.10.018.

Harvie, David. 2019. “(Big) Society and (Market) Discipline: Social Investment and the Financialisation of Social Reproduction.” Historical Materialism 27 (1): 92–124. https://doi.org/10.1163/1569206X-00001573.

Hawe, Penelope, Alan Shiell, and Therese Riley. 2004. “Complex Interventions: How ‘out of Control’ Can a Randomised Controlled Trial Be?” BMJ : British Medical Journal 328 (7455): 1561–63.

Heinrich, Caroline. 2018. “Keynote Speech: Comparing SIBs and Outcome-Based Approaches.” Social Outcomes Conference, Blavatnik School of Government, University of Oxford.

Hevenstone, Debra, and Matthias von Bergen. 2020. “Public–Private Partnerships in Social Impact Bonds: Facilitating Competition or Hindering Transparency?” Public Money & Management 40 (3): 205–12. https://doi.org/10.1080/09540962.2020.1714304.

Hill, Michael, and Peter Hupe. 2009. Implementing Public Policy: An Introduction to the Study of Operational Governance. London: Sage Publications Ltd.

Holmstrom, Bengt. 1979. “Moral Hazard and Observability.” The Bell Journal of Economics 10 (1): 74. https://doi.org/10.2307/3003320.

Hood, Christopher. 1991. “A Public Management for All Seasons?” Public Administration 69 (1): 3–19. https://doi.org/10.1111/j.1467-9299.1991.tb00779.x.

Hsieh, Hsiu-Fang, and Sarah E. Shannon. 2005. “Three Approaches to Qualitative Content Analysis.” Qualitative Health Research 15 (9): 1277–88. https://doi.org/10.1177/1049732305276687.

Huxham, Chris. 2003. “Theorizing Collaboration Practice.” Public Management Review 5 (3): 401–23. https://doi.org/10.1080/1471903032000146964.

INDIGO. 2020. “INDIGO Data Dictionary 2020.” International Network for Data on Impact and Government Outcomes. https://indigo-data-standard.readthedocs.io/en/latest/data-dictionary/index.html

INDIGO. 2021. “Impact Bond Dataset.” International Network for Data on Impact and Government Outcomes, The Government Outcomes Lab. 2021. https://golab.bsg.ox.ac.uk/knowledge-bank/indigo/impact-bond-dataset-v2/

Instiglio. 2021. “Impact Bonds.” What Is an Impact Bond? (blog). 2021. https://www.instiglio.org/en/impact-bonds/.

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. https://doi.org/10.1080/15575330.2013.854258.

Johnson, Jeffrey, Joyce Fortune, and Jane Bromley. 2017. “Systems, Networks, and Policy.” In Non-Equilibrium Social Science and Policy: Introduction and Essays on New and Changing Paradigms in Socio-Economic Thinking, edited by Jeffrey Johnson, Andrzej Nowak, Paul Ormerod, Bridget Rosewell, and Yi-Cheng Zhang, 111–34. Understanding Complex Systems. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-42424-8_8.

Joy, Meghan, and John Shields. 2013. “Social Impact Bonds: The Next Phase of Third Sector Marketization?” Canadian Journal of Nonprofit and Social Economy Research 4 (2). https://doi.org/10.22230/cjnser.2013v4n2a148.

Kenton, Will. 2020. “Organisation for Economic Co-Operation and Development (OECD).” Investopedia. November 22, 2020. https://www.investopedia.com/terms/o/oecd.asp.

Kisby, Ben. 2010. “The Big Society: Power to the People?” The Political Quarterly 81 (4): 484–91. https://doi.org/10.1111/j.1467-923X.2010.02133.x.

La Torre, Mario, Annarita Trotta, Helen Chiappini, and Alessandro Rizzello. 2019. “Business Models for Sustainable Finance: The Case Study of Social Impact Bonds.” Sustainability 11 (7): 1887. https://doi.org/10.3390/su11071887.

Lagarde, Mylene, Michael Wright, Julie Nossiter, and Nicholas Mays. 2013. “Challenges of Payment-for-Performance in Health Care and Other Public Services – Design, Implementation and Evaluation.” London: Policy Innovation Research Unit, London School of Hygiene and Tropical Medicine.

Lake, Robert W. 2015. “The Financialization of Urban Policy in the Age of Obama.” Journal of Urban Affairs 37 (1): 75–78. https://doi.org/10.1111/juaf.12167.

Leventhal, Rebecca. 2012. “Effecting Progress: Using Social Impact Bonds to Finance Social Services Articles and Comments Arising from the 2012 Fall Conference: The Law & Finance of Social Enterprise.” New York University Journal of Law and Business 9: 511–34.

LGA. 2019. “LGA Briefing: Debate on Local Government Funding, House of Commons.” Local Government Association. https://www.local.gov.uk/parliament/briefings-and-responses/lga-briefing-debate-local-government-funding-house-commons.

Liang, Max, Brian Mansberger, and Andrew Spieler. 2014. “An Overview of Social Impact Bonds.” Journal of International Business and Law 13 (2). https://scholarlycommons.law.hofstra.edu/jibl/vol13/iss2/3.

Liebman, Jeffrey B. 2011. “Social Impact Bonds.” Center for American Progress. https://www.americanprogress.org/issues/general/reports/2011/02/09/9050/social-impact-bonds/.

Lipsky, Michael. 1980. Street-Level Bureaucracy, 30th Ann. Ed.: Dilemmas of the Individual in Public Service. 2nd Revised Edition. New York: Russell Sage Foundation.

Loder, John. 2011. “The Promise of Social Impact Bonds.” Social Space, 42–47.

Lowe, Toby, Jonathan Kimmitt, Rob Wilson, Mike Martin, and Jane Gibbon. 2019. “The Institutional Work of Creating and Implementing Social Impact Bonds.” Policy & Politics 47 (2): 353–70. https://doi.org/10.1332/030557318X15333032765154.

Lowe, Toby, and Rob Wilson. 2017. “Playing the Game of Outcomes-Based Performance Management. Is Gamesmanship Inevitable? Evidence from Theory and Practice.” Social Policy & Administration 51 (7): 981–1001. https://doi.org/10.1111/spol.12205.

Lowndes, Vivien, and Lawrence Pratchett. 2012. “Local Governance under the Coalition Government: Austerity, Localism and the ‘Big Society.’” Local Government Studies 38 (1): 21–40. https://doi.org/10.1080/03003930.2011.642949.

Maier, Florentine, Gian Paolo Barbetta, and Franka Godina. 2018. “Paradoxes of Social Impact Bonds.” Social Policy & Administration 52 (7): 1332–53. https://doi.org/10.1111/spol.12343.

Marinetto, Michael. 2011. “A Lipskian Analysis of Child Protection Failures from Victoria Climbié to ‘Baby P’: A Street-Level Re-Evaluation of Joined-up Governance.” Public Administration 89 (3): 1164–81.

Matland, Richard. 1995. “Synthesizing the Implementation Literature: The Ambiguity-Conflict Model of Policy Implementation.” Journal of Public Administration Research and Theory, April. https://doi.org/10.1093/oxfordjournals.jpart.a037242.

May, Peter J., and Søren C. Winter. 2009. “Politicians, Managers, and Street-Level Bureaucrats: Influences on Policy Implementation.” Journal of Public Administration Research and Theory 19 (3): 453–76. https://doi.org/10.1093/jopart/mum030.

McHugh, Neil, Stephen Sinclair, Michael Roy, Leslie Huckfield, and Cam Donaldson. 2013. “Social Impact Bonds: A Wolf in Sheep’s Clothing?” Journal of Poverty and Social Justice 21 (3): 247–57. https://doi.org/10.1332/204674313X13812372137921.

McKay, Kyle. 2013. “Debunking the Myths Behind Social Impact Bond Speculation.” Stanford Social Innovation Review. https://ssir.org/articles/entry/debunking_the_myths_behind_social_impact_bond_speculation.

Merriam-Webster. 2021. “Definition of GOVERNMENT.” Merriam-Webster Dictionary. 2021. https://www.merriam-webster.com/dictionary/government.

Minogue, Martin. 2001. The Internationalization of New Public Management. Cheltenham: Edward Elgar Publishing. https://www.elgaronline.com/view/1840641819.00008.xml.

Moe, Ronald C. 2001. “The Emerging Federal Quasi Government: Issues of Management and Accountability.” Public Administration Review 61 (3): 290–312.

Mongkol, Kulachet. 2011. “The Critical Review of New Public Management Model and Its Criticisms.” 2011. https://doi.org/10.3923/rjbm.2011.35.43.

Morley, Julia. 2019. “The Ethical Status of Social Impact Bonds.” Journal of Economic Policy Reform 0 (0): 1–17. https://doi.org/10.1080/17487870.2019.1573681.

Moss, Ian. 2010. “The State of Commissioning.” The Institute for Government. https://www.instituteforgovernment.org.uk/publications/state-commissioning.

MST Program Manager. In-person Interview, August 2018.

MST Supervisor 1. In-person Interview, August 2018.

MST Supervisor 2. In-person Interview, August 2018.

MST Therapist 1. Interview, 2018.

MST Therapist 2. Interview, 2018.

MST Therapist 3. Interview, 2018.

MST Therapist 4. Interview, 2018.

MST Therapist 5. Interview, 2018.

MST Therapist 6. Interview, 2018.

MST UK. 2018. “Multisystemic Therapy | MST.” 2018. http://www.mstuk.org/about/about-2.

Mueller, Benjamin. 2019. “What Is Austerity and How Has It Affected British Society?” The New York Times, February 24, 2019. https://www.nytimes.com/2019/02/24/world/europe/britain-austerity-may-budget.html.

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.

Musante (DeWalt), Kathleen, and Billie R. DeWalt. 2010. Participant Observation: A Guide for Fieldworkers. Lanham: Rowman Altamira.

Nagel, Jack H. 1997. “Editor’s Introduction.” Journal of Policy Analysis and Management 16 (3): 349–56.

Nethercote, J.R. 1989. “Public Service Reform: Commonwealth Experience.” paper presented to the Academy of Social Sciences of Australia 25 February 1989, Canberra, Australian National University.

Neyland, Daniel. 2018. “On the Transformation of Children At-Risk into an Investment Proposition: A Study of Social Impact Bonds as an Anti-Market Device.” The Sociological Review 66 (3): 492–510. https://doi.org/10.1177/0038026117744415.

Nicholls, Alex. 2010. “The Legitimacy of Social Entrepreneurship: Reflexive Isomorphism in a Pre–Paradigmatic Field.” Entrepreneurship Theory and Practice 34 (4): 611–33. https://doi.org/10.1111/j.1540-6520.2010.00397.x.

Nicholls, Alex, and Alex Murdock. 2012. “The Nature of Social Innovation.” In Social Innovation, 1–30. Palgrave Macmillan, London. https://doi.org/10.1057/9780230367098_1.

O’Flynn, Janine. 2007. “From New Public Management to Public Value: Paradigmatic Change and Managerial Implications.” Australian Journal of Public Administration 66 (3): 353–66. https://doi.org/10.1111/j.1467-8500.2007.00545.x.

O’Flynn, Peter, and Chris Barnett. 2017. “Evaluation and Impact Investing: A Review of Methodologies to Assess Social Impact.” Brighton: Institute of Development Studies. https://www.ids.ac.uk/publications/evaluation-and-impact-investing-a-review-of-methodologies-to-assess-social-impact/.

Ofsted. 2010. “Inspection of Safeguarding and Looked after Children Services - Essex County Council.” https://files.api.ofsted.gov.uk/v1/file/50002491.

Osborne, Stephen P. 2006. “The New Public Governance?” Public Management Review 8 (3): 377–87. https://doi.org/10.1080/14719030600853022.

Painter, Gary. 2020. “Are Social Impact Bonds an Innovation in Finance or Do They Help Finance Social Innovation?” Presented at the Social Outcomes Conference 2020, Blavatnik School of Government, University of Oxford.

Painter, Martin. 1988. “Editorial Public Management: Fad or Fallacy?” Australian Journal of Public Administration 47 (1): 1–3. https://doi.org/10.1111/j.1467-8500.1988.tb01041.x.

Pawson, Ray, and Nicholas Tilley. 1997. Realistic Evaluation. London Sage Publications Ltd. https://uk.sagepub.com/en-gb/eur/realistic-evaluation/book205276.

Pearson, T.A., C. Lewis, S. Wall, P.L. Jenkins, A. Nafziger, and L. Weinehall. 2001. “Dissecting the ``black Box’ ’ of Community Intervention: Background and Rationale.” Scandinavian Journal of Public Health 29: 5–12. https://doi.org/10.1177/14034948010290021301.

Pollitt, Christopher. 1990. Managerialism and the Public Services: The Anglo-American Experience. Oxford: Basil Blackwell.

Pollitt, Christopher, and Geert Bouckaert. 2004. Public Management Reform: A Comparative Analysis. 2nd edition. Oxford; New York: Oxford University Press.

Rees, James, Adam Whitworth, and Elle Carter. 2014. “Support for All in the UK Work Programme? Differential Payments, Same Old Problem.” Social Policy & Administration 48 (2): 221–39. https://doi.org/10.1111/spol.12058.

Roberts, Lauren, and Genevieve Cameron. 2014. “Evaluation of the Essex Multi-Systemic Therapy Social Impact Bond.” OPM. https://traverse.ltd/application/files/9515/2285/2105/Interim-report-Essex-MST-SIB-Evaluation.pdf.

Roberts, Tamsyn. 2013. “The Centre for Social Impact Bonds.” Civil Service Quarterly. https://quarterly.blog.gov.uk/2013/10/22/the-centre-for-social-impact-bonds/.

Ronicle, James, and Kate Smith. 2020. “Youth Engagement Fund Evaluation.” GOV.UK. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/886650/YEF_Evaluation_Report_.pdf

Rosenman, Mark. 2014. “How Social Impact Bonds Put Private Profit Ahead of Public Good | PBS NewsHour.” PBS News Hour, February 19, 2014. https://www.pbs.org/newshour/nation/social-impact-bonds-put-private-profit-ahead-public-good.

Rothstein, Bo, and Dietlind Stolle. 2008. “Social Capital and Street-Level Bureaucracy: An Institutional Theory of Generalized Trust,” Comparative Politics 40 (4): 441-459. https://doi.org/10.2307/20434095

Rowe, Mike. 2012. “Going Back to the Street: Revisiting Lipsky’s Street-Level Bureaucracy.” Teaching Public Administration 30 (1): 10–18. https://doi.org/10.1177/0144739411435439.

Roy, Michael, Neil McHugh, and Stephen Sinclair. 2018. “A Critical Reflection on Social Impact Bonds.” Stanford Social Innovation Review. https://ssir.org/articles/entry/a_critical_reflection_on_social_impact_bonds.

Salis, Sergio, Robert Wishart, and Stephen McKay. 2018. “Evaluation of the Innovation Fund Pilot: Quantitative Assessment of Impact and Social Return on Investment.” National Centre for Social Research: Department for Work and Pensions.

Schinckus, Christophe. 2017. “Financial Innovation as a Potential Force for a Positive Social Change: The Challenging Future of Social Impact Bonds.” Research in International Business and Finance, Special Issue articles on Finance Reconsidered edited by Dr. Thomas Lagoarde-Segot and Dr. Bernard Paranque, Special issue articles on Recent trends and challenges in financial and commodity markets Edited by Prof. Fredj Jawadi and Prof. Benoît Sevi & Special Issue articles on Recent Topics in Banking and Finance: New Findings and Implications Edited by Prof. Fredj Jawadi and Prof. Wael Louhichi, 39 (January): 727–36. https://doi.org/10.1016/j.ribaf.2015.11.004.

Schneider, M, WJ Hall, AE Hernandez, K Hindes, G Montez, T Pham, L Rosen, et al. 2009. “Rationale, Design and Methods for Process Evaluation in the HEALTHY Study.” International Journal of Obesity (2005) 33 (Suppl 4): S60–67. https://doi.org/10.1038/ijo.2009.118.

Schreier, Margrit. 2014. “Qualitative Content Analysis.” In The SAGE Handbook of Qualitative Data Analysis. London: SAGE Publications Ltd. https://doi.org/10.4135/9781446282243.

Seidman, Harold. 1988. “The Quasi World of the Federal Government.” The Brookings Review 6 (3): 23–27. https://doi.org/10.2307/20080040.

SIB Board Advisor. In-person Interview, August 2018.

SIB Intermediary Lead. In-person Interview, July 2018.

SIB Investor 1. Phone Interview, September 2018.

SIB Investor 2. In-person Interview, August 2018.

SIB Investor 3. In-person Interview, August 2018.

Sinclair, Stephen, Neil McHugh, Leslie Huckfield, Michael Roy, and Cam Donaldson. 2014. “Social Impact Bonds: Shifting the Boundaries of Citizenship.” In Social Policy Review 26, edited by Kevin Farnsworth, Zoë Irving, and Menno Fenger, 119–36. Policy Press. https://doi.org/10.1332/policypress/9781447315568.003.0007.

Sinclair, Stephen, Neil McHugh, and Michael J. Roy. 2019. “Social Innovation, Financialisation and Commodification: A Critique of Social Impact Bonds.” Journal of Economic Policy Reform: 1–17. https://doi.org/10.1080/17487870.2019.1571415.

Skelcher, Chris. 2007. “Public–Private Partnerships and Hybridity.” In The Oxford Handbook of Public Management. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199226443.003.0016.

Smith, Peter. 1995. “Performance Indicators and Outcome in the Public Sector.” Public Money & Management 15 (4): 13–16. https://doi.org/10.1080/09540969509387889.

Social Finance. 2011. “Social Impact Bonds, the One Service, One Year on: Overview of the Peter-Borough Social Impact Bond.” London: Social Finance. http://www.socialfinance.org. uk/social-impact-bonds-the-one-service-one-year-on/.

Ssengooba, Freddie, Barbara McPake, and Natasha Palmer. 2012. “Why Performance-Based Contracting Failed in Uganda – An ‘Open-Box’ Evaluation of a Complex Health System Intervention.” Social Science & Medicine 75 (2): 377–83. https://doi.org/10.1016/j.socscimed.2012.02.050.

Steckler, Allan, and Laura Linnan, eds. 2002. Process Evaluation for Public Health Interventions and Research. 1st edition. San Francisco: Jossey-Bass Inc.

Stoesz, David. 2014. “Evidence-Based Policy: Reorganizing Social Services Through Accountable Care Organizations and Social Impact Bonds.” Research on Social Work Practice 24 (2): 181–85. https://doi.org/10.1177/1049731513500827.

Stout, Lynn A. 2014. “Killing Conscience: The Unintended Behavioral Consequences of ‘Pay for Performance.’” The Journal of Corporation Law 39: 38.

Tan, Stefanie, Alec Fraser, Chris Giacomantonio, Kristy Kruithof, Megan Sim, Mylene Lagarde, Emma Disley, Jennifer Rubin, and Nicholas Mays. 2015. “An Evaluation of Social Impact Bonds in Health and Social Care,” London: PIRU, London School of Hygiene & Tropical Medicine and RAND Europe.

Tan, Stefanie, Alec Fraser, Neil McHugh, and Mildred Warner. 2019. “Widening Perspectives on Social Impact Bonds.” Journal of Economic Policy Reform: 1–10. https://doi.org/10.1080/17487870.2019.1568249.

Thiel, Sandra van, and Frans L. Leeuw. 2002. “The Performance Paradox in the Public Sector.” Public Performance & Management Review 25 (3): 267–81. https://doi.org/10.1080/15309576.2002.11643661.

Thynne, Ian. 2003. “Making Sense of Public Management Reform.” Public Management Review 5 (3): 449–59. https://doi.org/10.1080/1471903032000146982.

Trelstad, Brian. 2009. “The Nature and Type of ‘Social Investors.’” Acumen Fund Investors. Acumen Fund. https://thegiin.org/assets/binary-data/RESOURCE/download_file/000/000/47-1.pdf.

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: 1–17. https://doi.org/10.1080/07352166.2018.1465347.

Urban Institute. 2019. “Denver Supportive Housing Social Impact Bond Initiative.” Urban Institute. April 25, 2019. https://www.urban.org/policy-centers/metropolitan-housing-and-communities-policy-center/projects/denver-supportive-housing-social-impact-bond-initiative.

US Treasury. 2018. “SIPPRA Legislation | U.S. Department of the Treasury.” February 2018. https://home.treasury.gov/services/social-impact-partnerships/sippra-pay-for-results/sippra-legislation.

Van Urk, Felix. 2016. “Function-Focused Implementation Fidelity for Complex Interventions: The Case of Studio Schools.” University of Oxford. https://ora.ox.ac.uk/objects/uuid:5c73b308-efbf-48aa-91b1-f8c06b7eb885.

Vera. 2015. “Impact Evaluation of the Adolescent Behavioral Learning Experience (ABLE) Program at Rikers Island.” Vera Institute of Justice. https://www.vera.org/publications/impact-evaluation-of-the-adolescent-behavioral-learning-experience-able-program-at-rikers-island.

Warner, Mildred. 2013. “Private Finance for Public Goods: Social Impact Bonds.” Journal of Economic Policy Reform 16 (4): 303-319. https://doi.org/10.1080/17487870.2013.835727.

Warner, Mildred. 2015. “Profiting from Public Value? The Case of Social Impact Bonds.” In *Creating Public Value in Practice,* ed. By Bryson, John, Barbara Crosby and Laura Bloomberg. New York: CRC Press, Taylor and Francis.

Webb, Calum J. R., and Paul Bywaters. 2018. “Austerity, Rationing and Inequity: Trends in Children’s and Young Peoples’ Services Expenditure in England between 2010 and 2015.” Local Government Studies 44 (3): 391–415. https://doi.org/10.1080/03003930.2018.1430028.

Weston, Cynthia, Terry Gandell, Jacinthe Beauchamp, Lynn McAlpine, Carol Wiseman, and Cathy Beauchamp. 2001. “Analyzing Interview Data: The Development and Evolution of a Coding System.” Qualitative Sociology 24 (3): 381–400. https://doi.org/10.1023/A:1010690908200.

Yin, Robert K. 2013. Case Study Research: Design and Methods. Vol. 5. Thousand Oaks, CA, US: Sage Publications, Inc.

Appendices

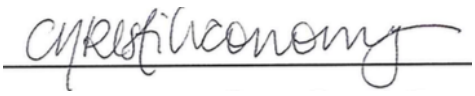
Appendix 1: Co-authorship declaration

Article Three (Have we 'stretched' Social Impact Bonds too far? An empirical analysis of SIB design in practice) of this thesis is co-authored by Christina Economy, Dr. Eleanor Carter, and Dr. Mara Airoldi. Drs. Carter and Airoldi confirm that Christina Economy is the primary author of this article and give permission for it to be included in her DPhil thesis.

The contributions of each author are described below:

- **Christina Economy (DPhil student):** Ms. Economy contributed to defining the overall research question and conceptualizing the study design. Ms. Economy proposed the use of qualitative content analysis as the methodology, and collaboratively all three co-authors worked together to construct the paper's coding frame. Ms. Economy was primarily responsible for data collection, data analysis, and interpretation, with guidance and assistance from the two co-authors. Ms. Economy drafted the paper and revised it according to co-authors comments.
- **Dr. Eleanor Carter:** Dr. Carter proposed the original idea for the research question and helped to conceptualize the paper design alongside the other authors. Dr. Carter worked collaboratively with the other authors to construct and test the coding frame, as well as to interpret the results. Dr. Carter also provided a substantive review of the paper draft and proposed improvements.
- **Dr. Mara Airoldi:** Dr. Airoldi helped to conceptualize the research question and design alongside the other two authors. Dr. Airoldi also helped to construct the coding frame and came up with the idea to create composite scores for the two dimensions of interest (payment for impact and nature of the capital). Dr. Airoldi assisted in interpretation of results, and provided a substantive review of the paper draft.

Signed:



Christina Economy Date: April 2, 2021

ECarter

Eleanor Carter Date: 13 April 2021



Mara Airoldi Date: April 13, 2021

Appendix 2: Ethics approval, Essex case study



Ms Christine Economy
Dr Mara Airoldi
Blavatnik School of Government

09 May 2018

Dear Ms Economy,

Research Ethics Approval

Ref No: SSD/CUREC1A/BSG_C1A-18-008

The role of social impact bond financing on delivery of multi-systemic therapy.

The above application has been considered on behalf of the Blavatnik School of Government's Departmental Research Ethics Committee (DREC) in accordance with the procedures laid down by the University for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to the DREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly approval has been granted.

Should there be any subsequent changes to the project, which raise ethical issues not covered in the original application, you should submit details to the DREC for consideration.

Please note that this approval is valid until the 01 April 2024 as requested in your CUREC application. Should you need an extension, please request this from the committee before your validity period ends.

Yours sincerely,

Dr. Thomas Elston
BSG DREC Chair
Blavatnik School of Government

Blavatnik School of Government, University of Oxford
ethics@bsg.ox.ac.uk | www.bsg.ox.ac.uk

Appendix 3: Participant information sheet, Essex case study

Christina Economy
Blavatnik School of Government
Radcliffe Observatory Quarter
Woodstock Road
Oxford OX2 6GG
Tel: 01865 614330
Email: christina.economy@stcatz.ox.ac.uk

The role of social impact bond financing on delivery of interventions in children's social care

PARTICIPANT INFORMATION SHEET

Ethics Approval Reference: BSG-C1A-18-008

1. ***What is the purpose of this research?***

This study aims to contribute to our understanding of how social impact bond (SIB) projects could support children and families. In particular, the research seeks to investigate key mechanisms through which SIBs may yield improved intervention outcomes and the impact that this financing structure may have on front-line staff.

2. ***Why have I been invited to take part?***

Your team has been invited because you have been identified as front-line workers involved with a multi-systemic therapy (MST) or an MST-related project. MST is one of the interventions in children's social care that has been funded using a SIB.

3. ***Do I have to take part?***

No. You and/or any members of your team have no obligation to take part. You can ask questions about the study before deciding whether or not to participate. If you do agree to participate, you may withdraw yourself and the information you provided from the study at any time up until publication, without giving a reason and without penalty, by advising the researchers of this decision. If you do decide to withdraw, any information you provide will be permanently deleted from the researcher's records.

4. ***What will happen to me if I take part in the research?***

If you are happy to participate in the research, one or two researchers will spend time shadowing and observing your team. You/members of your team may also be asked to take part in an interview that will last approximately one hour. At this point, front-line staff will be asked to sign a consent form to participate. After the interview ends, the researcher may contact you for follow-ups or clarification questions. These follow-ups will be limited.

5. ***Are there any potential risks in taking part?***

The proposed research will be highly unlikely to cause harm to participants. However, there is the potential risk to reveal opportunistic behaviour or negative attitudes, or to reveal opportunistic behaviour

of superiors or upper-level management during the interview process. To mitigate these risks, all interviews with front-line staff will be anonymized. No one other than the principal researchers will be able to link the information provided in the interview with front-line staff. Participants are also free to withdraw at any time, and free to skip any of the questions they don't wish to answer.

6. Are there any benefits in taking part?

There will be no direct benefit to you from taking part in this research.

7. What happens to the data provided? [23]

All data will be stored on secure University systems, and research data will be stored for a 3-year period after publication or public release of the work of the research according to University guidelines.

Interviews will be recorded if participants give their consent, otherwise brief notes will be made during the interview and written up immediately afterwards. Recordings will be downloaded as soon as possible to a secure University computer and then deleted from the portable device. The portable device used for recordings will be encrypted. If hand-written notes are taken, these will not include any names or identifying information. Recordings will not be kept beyond the end of the research project without permission from interviewees, and transcripts will be archived anonymously. Publications arising from the study will not include any information or quotations that allow the speaker to be identified, unless this consent is given.

8. Will the research be published?

The University of Oxford is committed to the dissemination of its research for the benefit of society and the economy and, in support of this commitment, has established an online archive of research materials. This archive includes digital copies of student theses successfully submitted as part of a University of Oxford postgraduate degree programme. Holding the archive online gives easy access for researchers to the full text of freely available theses, thereby increasing the likely impact and use of that research.

The research will be written up as a thesis. On successful submission of the thesis, it will be deposited both in print and online in the University archives, to facilitate its use in future research. The thesis will be openly accessible. The research may also be published in academic journals, which will be available online and may also be accessible in print.

9. Who is organising and funding the research?

The Government Outcomes Lab (GO Lab) is organizing and funding the research. The GO Lab is a joint partnership between the University of Oxford and the U.K. Department for Digital, Culture, Media & Sport.

10. Who has reviewed this study?

This study has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee (Reference number: BSG-C1A-18-008).

[23] Please refer to CUREC's Best Practice Guidance on Data Collection and Management (BPG 09)

11. Who do I contact if I have a concern about the study or I wish to complain?

If you have a concern about any aspect of this study, please speak to Christina Economy [01865 614330] or Dr. Mara Airoldi [01865 614393]. The researcher should acknowledge your concern within 10 working days and give you an indication of how they intend to deal with it. If you remain unhappy or wish to make a formal complaint, please contact the relevant chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter in a reasonably expeditious manner: Chair, **Social Sciences & Humanities Inter-Divisional Research Ethics Committee**; Email: ethics@socsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD

12. Further Information and Contact Details

If you would like to discuss the research or if you have questions at any time, please contact:

Christina Economy
Tel: 01865 614330
Email: christina.economy@stcatz.ox.ac.uk

Appendix 4: Written consent forms, Essex case study

Consent form for elite interviews

Blavatnik School of Government
University of Oxford,
Radcliffe Observatory Quarter,
Woodstock Road,
Oxford OX2 6GG,
United Kingdom

Christina Economy
Christina.economy@stcatz.ox.ac.uk
DPhil Candidate in Public Policy
Oxford telephone number: 01865 614330

PARTICIPANT CONSENT FORM

CUREC Approval Reference: **BSG-C1A-18-008**

Research Study: The role of social impact bond financing on delivery of multi-systemic therapy

	Please initial each box
I understand that this project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee. I also understand how to raise a concern or make a complaint.	☐
I consent to being audio recorded.	☐
I give permission to be identified by name in this research and any related publications.	☐
I agree to take part in the study.	☐

Name of participant	Date	Signature

Name of person taking consent	Date	Signature

Consent form for front-line staff interviews

Blavatnik School of Government
University of Oxford,
Radcliffe Observatory Quarter,
Woodstock Road,
Oxford OX2 6GG,
United Kingdom

Christina Economy
Christina.economy@stcatz.ox.ac.uk
DPhil Candidate in Public Policy
Oxford telephone number: 01865 614330

PARTICIPANT CONSENT FORM

CUREC Approval Reference: **BSG-C1A-18-008**

Research Study: The role of social impact bond financing on delivery of interventions in children's social care

	Please initial each box
I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	☐
I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without any adverse consequences or penalty.	☐
I understand that this project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee. I also understand how to raise a concern or make a complaint.	☐
I understand that all of the information I provide in this interview will be anonymized.	☐
I consent to being audio recorded.	☐
I agree to take part in the study	☐

Name of Participant	Date	Signature
Name of person taking consent	Date	Signature

1

Consent form for front-line shadowing

Blavatnik School of Government
University of Oxford,
Radcliffe Observatory Quarter,
Woodstock Road,
Oxford OX2 6GG,
United Kingdom

Christina Economy, DPhil Candidate in Public Policy
Christina.economy@stcatz.ox.ac.uk
Oxford telephone number: 01865 614330

PARTICIPANT CONSENT FORM

CUREC Approval Reference: **BSG-C1A-18-008**

Research Study: The role of social impact bond financing on delivery of interventions in children's social care

	Please initial each box
I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	☐
I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without any adverse consequences or penalty.	☐
I understand that this project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee. I also understand how to raise a concern or make a complaint.	☐
I understand that all of the information provided by myself or members of my team during the observation/"shadowing" process will be anonymized.	☐
I consent to being audio recorded during the formal interview portion of this study.	☐
I give permission to be identified by name in this research and any related publications.	☐
I agree to take part in the study	☐

Name of participant	Date	Signature
Name of person taking consent	Date	Signature

Appendix 5: List of key stakeholders, Essex case study

Title	Description
ECC Head of Children's Services	Directs the entire array of Essex County Council's children and family services
ECC Project Lead	Has a leadership role in Essex's children services agency and served as the primary government lead on the SIB project
ECC Service Manager	Has a leadership role in Essex's children social care work, and interacted often with the MST service
ECC Former Head of Partnership	Had a leadership role in Essex's commissioning unit and helped plan the SIB prior to its launch
ECC Strategy and Innovation Manager	Works in Essex's commissioning unit and worked directly on the SIB project's implementation
ECC Front-line Staff	Serve Essex's children and families through programs similar to MST
Provider Operational Director	Oversees a number of the provider's service delivery projects, including MST in Essex
MST Program Manager	Employed by the provider, and worked directly on the Essex project with MST supervisors, council leadership, and the SIB Board
MST Supervisors	Employed by the provider, acted as direct supervisors for the therapists providing MST in Essex
MST Therapists	Employed by the provider, delivered MST to children and families in Essex
SIB Investors	Contributed to the upfront funding of the project and remained engaged in managing the project throughout service delivery
SIB Advisor	Served on the SIB Board with the SIB Investors, played an advisory role based on previous experiences of MST delivery in other jurisdictions
Intermediary Leads	Served as a link between all project stakeholders, responsible for data monitoring and management related to the outcome payments and project performance

Appendix 6: Interview topic guides, Essex case study

Government stakeholders topic guide

Topic	Prompts/Issues to discuss
Introductions	- Nature of role in relation to Essex MST project - Motivation for involvement in SIB project
Overall SIB effect	- Comparison of SIB project and similar government services - Value-add of the SIB model - Drawbacks/frustrations with SIB model - Key SIB ingredients/components that led to value-add or drawbacks
Multi-sector partnership	- Key players in SIB project and relational dynamics between them o Examples of stakeholder cooperation or conflict - Frequency/nature of interaction with investor, intermediary, and provider partners - Fit of SIB project within wider scope of children's services - Process/governance structure of designing/implementing the SIB
Data monitoring/ Performance management	- Description of SIB project management structure - Frequency/content of meetings with investor or provider staff o Outcomes discussed/consequences for missing targets - Biggest obstacles/challenges in MST delivery - Process for identifying operational or service quality challenges

Flexibility/Resource allocation	- Changes to service delivery over course of the project
	  o Process that led to these changes (level of gov't discretion)
	  o Resources spent on making changes (level of gov't discretion)

Outcome focus	- Sentiment around project's 'payable outcomes' - Concerns around not meeting outcome or project goals o Action/response to these concerns (if any)
Wind-down	- Reasons for not continuing with MST or SIB - Additional/final reflections - Referrals to other project stakeholders

SIB investor/advisor/intermediary topic guide

Topic	Prompts/Issues to discuss
Introductions	- Nature of role in relation to Essex MST project - Motivation for investment/involvement in project
Overall SIB effect	- Value-add of SIB model - Drawbacks of SIB model - Key SIB ingredients/components that led to value-add or drawbacks
Multi-sector partnership	- Key players in SIB project and relational dynamics between them o Examples of stakeholder cooperation or conflict - Frequency/nature of interaction with government and provider partners
Data monitoring/ Performance management	- Description of SIB project management structure - Frequency/content of meetings with provider staff o Outcomes discussed/consequences for missing targets

	- Process for identifying operational or service quality challenges

Flexibility/Resource allocation	- Changes to service delivery over course of the project
	  o Process that led to these changes
	  o Resources spent on making changes

Outcome focus	- Sentiment around project's 'payable outcomes' - Concerns around not meeting outcome goals o Action/response to these concerns (if any) - Incentives provided to meet targets/goals
Wind-down	- Additional/final reflections - Referrals to other project stakeholders

Service provider (management-level) topic guide

Topic	Prompts/Issues to discuss
Introductions	- Nature of role in relation to Essex MST project - Previous experience (including with SIBs/PbR/FFS) - Familiarity with SIB structure
Overall SIB effect	- Comparison of Essex MST to other MST teams - Familiarity/opinion on payable outcomes - Value-add of the SIB model - Drawbacks of SIB model - Biggest challenges in service delivery - Components that would make job easier

Performance management/Outcome focus/Incentives	- Frequency/content of meetings with front-line therapists
	  o Outcomes discussed/consequences for missing targets
	- Frequency/content of interactions with other project stakeholders
	  o Process for reporting back on outcomes
	  o Consequences or incentives for achieving/missing targets

Collaboration/Support networks	- Relationship with SIB investor board/intermediary/government partner/wider service provider organization
Flexibility	- Changes to service delivery over course of the project o Process that led to these changes - Level of discretion over referrals, changes to service delivery, action plans for working with families, etc. - Goodness of fit of MST referrals
Wind-down	- Additional/final reflections - Referrals to other project stakeholders

Service provider front-line staff topic guide

Topic	Prompts/Issues to discuss
Introductions	- Length of time on this project - Previous experience in children's social care o Comparison to previous roles
Performance management/Outcome focus/Incentives	- Motivations to do this job - Main areas of stress on the job - Typical daily activities - Process for keeping track of family's progress - Frequency/content of meetings with supervisors/management o Outcomes discussed/consequences for missing targets
Collaboration/Support networks	- Relationship with SIB investor board/intermediary/government partner/wider service provider organization - Relationship with other stakeholders also working with a family
Flexibility	- Process for deciding what to work on/towards with a family - Process if MST service isn't meeting family's needs - Level of discretion over referrals, changes to service delivery, action plans for working with families, etc. - Goodness of fit of MST referrals

SIB effect	- Familiarity with SIB structure (if at all)
	  o Value-add of the SIB model
	  o Drawbacks of SIB model
	- Biggest challenges in service delivery
	- Components that would make job easier

Wind-down	- Additional/final reflections - Referrals to other project stakeholders

Appendix 7: NVivo thematic coding nodes, Essex case study

SIB effect on public service delivery

Major Theme (Node)	Sub-Themes (Sub-Nodes)
Multi-sector partnership	- Collaboration - Culture clashes - Governance structure - MST service integration - Stakeholder relationships - Sustained engagement - System linkages
Data monitoring system	- Administrative burden - Identifying challenges - Performance management system - Real-time changes
External investment	- Additional resources - Democratic accountability - Disproportionality - Provider discretion - Returns motive - Tailored support - Transparency
Outcomes contract	- Changing context - Outcome alignment - Outcome focus - Provider incentives - Scope of services

SIB effect on front-line staff

Major Theme (Node)	Sub-Themes (Sub-Nodes)
Support networks	- Availability of additional resources - Collaboration - Conflict - Job Security - Lines of accountability - MST service integration
Performance management	- Administrative burden - Identifying challenges - Pressure to meet targets - Use of performance data
Flexibility	- Innovative service approaches - Provider discretion - Resource allocation - Tailored support
Incentives	- Consequences for missed targets - Financial rewards - Non-financial rewards - Outcome alignment - Outcome focus - Outcome gaming

Appendix 8: Coding frame definitions, qualitative content analysis

Term	Definition
Payment model	The percentage of contract payments that are made based on outcomes. **Options**: 0% payment on outcomes; Less than 50% (but more than 0%) payment on outcomes; 50% or more (but less than 100%) payment on outcomes; 100% payment on outcomes
Outcome	Results of service users achieved once the beneficiary population uses the tangible goods/service/intervention provided by the project (Gertler et al. 2016, 34). Outcomes are not directly under the control of a program-implementing agency (Ibid, p. 330)
0% payment on outcomes	None of the contract payments are made for outcomes
Less than half (but more than 0%) payment on outcomes	Less than half, but more than 0%, of the contract payments are made for outcomes
More than half (but less than 100%) payment on outcomes	More than half, but less than 100%, of the contract payments are made for outcomes
100% payment on outcomes	All of the contract payments (100%) are made for outcomes
Validation method	The approach taken to assessing whether the specified outcomes have been achieved and the basis on which 'outcome payments' will be made. **Options**: Randomized Control Trial; Quasi-experimental study; Historical comparison/Pre-Post analysis; Simple ex-post verification; No validation method
No validation method	There is no approach taken to assess whether specified outcomes have been achieved
Simple ex-post verification	An approach that assesses whether program beneficiaries have achieved outcomes after the fact, without measuring changes over time or using a comparison group
Pre/post analysis or historical comparison	An approach that tracks changes in outcomes for program beneficiaries over time, using measurements before and after the program or policy is implemented (Gertler et al. 2016, 326). Before measurements are either from program participants themselves (pre/post) or an earlier group (historical comparison)
Quasi-experimental study	Impact evaluation methods (i.e. that make a causal link between a program and set of outcomes) that use a counterfactual (i.e. estimate of what the outcome would have been for program participants if they had not participated) but are not based on randomized assignment of the intervention (Gertler et al. 2016).
Randomized control trial	Impact evaluation method whereby every eligible individual or unit has a probability of being selected for treatment by a program. With a sufficiently large number of units, ensures equivalence in both observed and unobserved characteristics between the treatment group and the comparison group, thereby ruling out selection bias. Considered the most robust method for estimating counterfactuals and is often referred to as the gold standard of impact evaluation (Gertler et al. 2016, 332).

Primary investor type	The category of organization that contributed the most money to upfront project costs. **Options**: Government; Provider; Philanthropy; Social Investor; Commercial Investor
Government	A local, state, or national agency through which a political unit exercises authority and performs functions (Merriam-Webster 2021).
Provider	An organization that has been contracted as part of the project to deliver the goods/services/intervention to intended beneficiaries
Philanthropy	An organization or individual that provides money to address social issues, and whose primary mission is to maximize social returns (Trelstad 2009).
Social Investor	An organization or individual that provides upfront money for a project, and whose primary mission is blended - to achieve social impact and a financial return (Trelstad 2009).
Commercial Investor	An organization or individual that provides upfront money for a project, and whose primary mission is to maximize financial return (Trelstad 2009).
Repayment Structure	A description of how the performance risk is shared between the investors of upfront capital and the service provider. **Options**: Provider holds all performance risk; Provider primarily holds performance risk; Provider and investor split performance risk; Investor primarily holds performance risk; Investor holds all performance risk
Performance risk	The possibility of losing upfront capital investment if payable project outcomes are not met
Provider holds all performance risk	Investors are paid principal plus return regardless of performance, and provider is paid based on outcomes
Provider primarily holds performance risk	Investor is paid back the principal regardless but returns vary by performance. The provider is paid based on outcomes
Provider and investor split performance risk	Both investor and provider groups are paid based on outcomes
Investor primarily holds performance risk	Investor are paid based on outcomes, but can withhold payment from the provider based on meeting certain key performance indicators
Investor holds all performance risk	Investor is paid based on outcomes, and the provider is paid regardless of outcome performance

Appendix 9: Sensitivity analysis, qualitative content analysis

	Nature of Capital Score (% of Projects)					Total # of Projects
	0	**1**	**2**	**3**	**4**	
	ORIGINAL VALUES (IN PAPER)					
UK	1%	1%	1%	25%	72%	81
US	0%	0%	9%	36%	55%	22
Totals	1%	1%	3%	27%	68%	103
	SENSITIVITY ANALYSIS 1: FULLY VERIFIED REPAYMENT STRUCTURE					
UK	0%	0%	7%	21%	71%	14
US	0%	0%	10%	33%	57%	21
Totals	0%	0%	9%	29%	63%	35
	SENSITIVITY ANALYSIS 2: DEFAULT LOWER FOR PRIMARY INVESTOR OF CAPITAL					
UK	1%	6%	0%	28%	64%	81
US	5%	0%	9%	45%	41%	22
Totals	2%	5%	2%	32%	59%	103

Sensitivity analysis 1: In many cases in the UK the specific details of the repayment structure between investors and provider partners are treated as commercially sensitive and are unavailable. However, for many projects, there are indications of whether providers shared in performance risk, by either holding the outcomes contract or contributing some of the upfront capital (or both). For the purposes of Article 3, in cases where the repayment structure was unable to be verified, projects that had evidence of provider risk share scored a 2 and projects that did not have evidence of provider risk share scored a 3. The results of this analysis are reflected in the 'Original Values' section of the chart above. To test these assumptions, Sensitivity Analysis 1 shows the results when only coding projects that have been fully verified in their repayment structure. The resulting proportions are similar across the scores.

Sensitivity analysis 2: For the purposes of Article 3, if the data is unclear on how much upfront capital different groups contributed, projects were scored the highest of the contributing types (for example, if both philanthropy and social investors provided upfront investment but with no detail on what quantities, the project would score a 3 along primary investor type). The results of this analysis are reflected in the 'Original Values' section of the chart above. To test these assumptions, Sensitivity Analysis 2 shows the results when projects are scored the lowest of the contributing types (for example, if both philanthropy and social investors provided upfront investment but with no detail on what quantities, the project would score a 2 along primary investor type). The resulting score proportions remain relatively similar, although more projects would have scored a 3 rather than a 4 for nature of the capital if we had defaulted to the lower score for primary investor of capital.