

# Contractual acrobatics: a configurational analysis of outcome specifications and payment in outcome-based contracts

Clare FitzGerald, Stefanie Tan, Eleanor Carter & Mara Airoidi

**To cite this article:** Clare FitzGerald, Stefanie Tan, Eleanor Carter & Mara Airoidi (2023): Contractual acrobatics: a configurational analysis of outcome specifications and payment in outcome-based contracts, Public Management Review, DOI: [10.1080/14719037.2023.2244501](https://doi.org/10.1080/14719037.2023.2244501)

**To link to this article:** <https://doi.org/10.1080/14719037.2023.2244501>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 07 Aug 2023.



[Submit your article to this journal](#)



Article views: 406



[View related articles](#)



[View Crossmark data](#)

# Contractual acrobatics: a configurational analysis of outcome specifications and payment in outcome-based contracts

Clare FitzGerald <sup>a</sup>, Stefanie Tan <sup>b</sup>, Eleanor Carter <sup>c</sup> and Mara Airoidi <sup>c</sup>

<sup>a</sup>Department of Public Services Management & Organisation Group, King's College London, King's Business School, London, United Kingdom; <sup>b</sup>North American Observatory for Health Systems and Policy, Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, Canada; <sup>c</sup>Government Outcomes Lab, Blavatnik School of Government, University of Oxford, Oxford, United Kingdom

## ABSTRACT

Outcome-based contracting (OBC) seeks to improve public services by paying for service outcomes rather than service activities. This article explores the link between how outcomes are contractually specified and how much is paid for their achievement. Using fuzzy-set Qualitative Comparative Analysis, we test a framework for assessing the strength of outcome specifications in 34 UK-based social impact bonds, a particular form of OBC. Results show that contract features which define intended participant cohorts and include deadweight estimation approaches help constrain suppliers' ability to appropriate value and thus reduce the likelihood that public managers pay for social outcomes of questionable value.


**ARTICLE HISTORY** Received 27 October 2022; Accepted 31 July 2023

**KEYWORDS** social impact bonds; outcome-based contracting; qualitative comparative analysis; opportunism; social value

## Introduction

The UK central government has promoted outcome-based contracting (OBC) for over a decade. OBC seeks to improve public services by making payment to providers dependent on the achievement of outcomes rather than the delivery of activities. Blending managerialism and marketization, OBC supporters suggest that payment on outcomes, in lieu of specification of activities, allows contract principals to give strategic direction and ensure accountability, whilst enabling service flexibility and innovation for delivery agents (FitzGerald et al. 2019; Lazzarini 2022). This, in turn, aligns financial incentives and social goals enabling better cross-sector collaboration (Mulgan et al. 2011; Cohen, 2011). Hence, advocates contend that OBCs are unique in their ability to incentivize adaptation in pursuit of service efficiency and effectiveness,

**CONTACT** Clare FitzGerald  [clare.fitzgerald@kcl.ac.uk](mailto:clare.fitzgerald@kcl.ac.uk)

 Supplemental data for this article can be accessed at <https://doi.org/10.1080/14719037.2023.2244501>

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

whilst ensuring public accountability through generating evidence of the achievement of pre-defined social outcomes for an agreed price.

The allure of outcomes orientation is nothing new in public management. Rewarding on the achievement of goals is undoubtedly familiar to scholars of performance management in public service organizations (Heinrich 1999; Moynihan et al. 2011; Swiss 1992). A key difference, however, is the application of intra-organizational theory to inter-organizational activity. Albeit a deliciously straightforward idea to drive behaviour, outcome-orientation poses a particular conundrum when contractualised (Tse and Warner 2020a). OBCs are complex and incomplete contracts (Hart 1988), with commissioned services often targeting entrenched, multiply disadvantaged cohorts in fragmented and densely networked service delivery partnerships (Carter et al. 2018). In the UK, OBCs also operate in a highly constrained local funding environment (FitzGerald et al. 2021). Because of these complexities, delivery partners are often unable to provide full contract specifications at proportional verification costs (Brown, Potoski, and Slyke 2018; FitzGerald et al. 2019). Consequently, a perennial fear within OBCs is that, due to under-specification, delivery partners will prioritize profit over societal benefits or social value, for example, by failing to support all intended service users, delivering insufficient service quality, or compromising longer-term policy objectives by specifying and paying on the basis of shorter term measures of success (FitzGerald et al. 2021; Heinrich and Kabourek 2019; Lazzarini 2020).

A recurrent problem within OBCs is incentivizing providers to support all programme participants effectively considering their differing support costs and varied likelihoods of realizing specified social outcomes and thus triggering payment (Carter and Whitworth 2015; Newhouse 1984; van de Ven and van Vliet 1992). International literature makes clear that programme design and payment structures can drive delivery agent behaviours in ways that can promote or undermine the social value of service arrangements (Isaksson, Blomqvist, and Winblad 2018; Lazzarini 2020; Maurya and Srivastava 2019). The implication being that through careful specification and oversight, it may be possible to constrain the ability of service providers to appropriate economic rents and ensure that providers support the full range of programmatic objectives rather than advancing only those which maximize profit (Lazzarini 2020). Consequently, an early challenge for those purchasing services through OBCs is to configure contracts that retain the perceived benefits of OBC – namely the flexibility and adaptiveness associated with minimal service prescription – whilst ensuring efficiency, effectiveness, and equity.

Using the population of completed UK social impact bonds (SIBs), a particular form of OBC, this article investigates which aspects of contractual outcome specifications mitigate against value appropriation and thus potential opportunism. In the sections that follow, we provide background on UK SIBs and a review of transaction cost economics and complex contracting literature. We suggest that conventional routes to overcome opportunism in the form of service requirements (Williamson 1985) runs counter to the relational ethos and flexibility of outcome contracts and instead, suggest and operationalize a framework of contractual conditions for further analysis, building on FitzGerald et al.'s framework of requisite contracts (FitzGerald et al. 2019).

## Social Impact Bonds – a ‘novel’ form of OBC

SIBs are outcome-based contracts between outcome payors (usually government) and service providers (usually charities), with investors covering the costs of service delivery to providers to make a return if predefined social outcomes are achieved (FitzGerald et al. 2019). Conditioning payment upon outcome achievement is the mechanism through which SIBs are meant to transfer the financial risks of complex social services to investors, enabling flexibility *and* public accountability in delivery.

SIBs make a compelling example of OBC to study given their use of high-powered financial incentives, unique bundling of contemporary public management reform approaches – including performance contracting, public–private partnerships, and evidence-informed policymaking – and growing adoption globally (Heinrich and Kabourek 2019). Since the first SIB launched at HMP Peterborough in 2010, over 270 SIBs have launched worldwide with 93 of those projects operating or completed within the UK (INDIGO 2023). While political support for OBCs has remained relatively consistent internationally, SIBs have been the subject of polarized academic debate. As identified by Fraser and colleagues (Fraser et al., 2018), narratives about SIBs take various forms: an optimistic ‘narrative of promise’ where SIBs are believed to be a solution to a lack of public sector innovation and entrepreneurship and a ‘narrative of caution’ where SIBs promulgate a dangerous process of financialization with governmental and voluntary endeavours being subordinated to profit seeking by the investor class (Warner, 2013; Lake, 2015; Tse and Warner 2020b). Through their use of private contracting models and performance management within service networks, SIBs at once embody New Public Management reform tenants alongside the multi-party collaborative forms more associated with New Public Governance (FitzGerald et al. 2023; French et al. 2022). Further, evidence shows that a key pinch-point in the development of SIB projects is the setting of monetary incentives, where it is noted that greater commitment from project partners is needed to ensure ‘payments for outcomes are methodologically rigorous and based on long-term outcomes that cannot be easily “gamed”’ (Heinrich and Kabourek 2019, 867).

The larger body of work on public sector contracting, meanwhile, is largely focused on general terms and contextual factors within contractual relationships rather than the analysis of specific contractual features which link to contract performance (Romzek and Johnston, 2005; Klijn and Koppenjan, 2016; Maurya and Srivastava 2019). Detailed empirical work investigating alternative specifications and contractual arrangements particular to OBC projects is rare internationally and detailed, systematic comparative assessments are particularly scarce. Academic work on SIBs is still largely conceptual, based on accounts of one to a few projects within discreet geographies or policy areas (FitzGerald, Fraser, and Kimmitt 2020). The research offered here therefore makes a novel contribution, by presenting a comparative configurational analysis of contract features and outcome payments across 34 UK SIBs completed as of January 2021.

## Specifying outcomes to achieve social value

Transaction Cost Economics (TCE) views the relationship between a purchaser and a supplier of goods or services as an exercise in reducing the ‘transaction costs’ of negotiating and managing the relationship, acknowledging the human characteristics

**Table 1.** Williamson’s two dimensions of opportunism and bounded rationality: a typology of contracting approaches (adapted from Williamson 1985.).

| Opportunism∨ Bounded rationality> | Absent                      | Admitted                         |
|-----------------------------------|-----------------------------|----------------------------------|
| Absent                            | Bliss                       | ‘General Clause’ contracting     |
| Admitted                          | ‘Comprehensive’ contracting | Serious contracting difficulties |

of *opportunism* (i.e. ‘self-interest seeking with guile’ (Williamson 1985, 47)) and *bounded rationality* (i.e. purchaser’s incomplete knowledge of the current and future situation (Simon 1957)) (see Table 1). Hence, a well-specified or ‘complete’ contract is considered the purchaser’s best defence against supplier opportunism. However, because the purchaser is boundedly rational, they cannot predict all the risks and contingencies required to fully specify a contract. Where a purchaser is unconcerned with supplier opportunism and acknowledges the incompleteness of their information, they may rely on a ‘general clause’ contract specifying only that both parties will act in good faith. Thus, the uncertainty that a contract will ‘work’ stems from two critical and persistent knowledge gaps: a ‘lack of technical knowledge’ about the good or service being contracted and ‘a lack of knowledge about how the other will behave in the areas where the contract permits discretion’ (Brown, Potoski, and Slyke 2018).

Because opportunism and bounded rationality are never wholly absent (Williamson 1985, 67), purchasers must perform a tightrope walk. While a general clause contract may be inexpensive to arrange, it leaves the purchaser at risk of supplier opportunism, particularly if the supplier takes advantage of information asymmetry about the product or service they offer. And, because of bounded rationality, attempts to specify a ‘complete contract’ are costly and not guaranteed to guard against unforeseen types of opportunism. The literature on complex contracting suggests that rules can solve this dilemma by offering ‘incentives to favor consummate behavior’ (Brown, Potoski, and Slyke 2016, 301). Consummate behaviour ‘decreases the performer’s gains but by a smaller amount than the gains it creates for the other side’ (Brown, Potoski, and Slyke 2016, 300). Its opposite is perfunctory behaviour which, like opportunism, produces ‘small gains for the performer but imposes greater losses on the other side’ (Brown, Potoski, and Slyke 2016, 300).

As a form of OBC, SIBs prioritize monetary performance incentives as a tool for delivering a ‘win-win-win’ for stakeholders and it is the outcome specification within OBCs that promises to lower the risk of opportunism by simplifying the exchange (Lazzarini 2022). From the purchaser’s perspective however, specifying outcomes instead of inputs and activities means that the traditional routes of completing a contract are less applicable: purchasers cannot simultaneously dictate the means of delivery through minimum service specifications or intervention selection and then insist on holding suppliers responsible for the outcomes those requirements produce. Hence, the question for OBC purchasers – i.e. those paying for outcomes – is how specify outcomes such that they complete the contract, ensuring adequate flexibility to suppliers *and* adequate protection of public expenditure for purchasers.

**Opportunism and appropriation in OBCs**

The logic of contract incompleteness is clear in its implications for public managers contracting with profit-motivated suppliers: while efficiency may improve, the

tendency of profit-oriented entities to ‘appropriate excessive economic rents’ reduces the value gains managers can expect (Lazzarini 2020). In the strategic management tradition, value has been thought of as the delta between a consumer’s willingness to pay for a good or service minus the costs associated with its production (Garcia-Castro and Aguilera 2015; Lazzarini 2020). In the public administration tradition, value has taken on social dimensions with social value conceptualized as ‘new and appropriable benefits to society for which it directly – as consumers – or indirectly – as taxpayers – is able and prepared to pay’ (Bryson, Crosby, and Bloomberg 2014; Kroeger and Weber 2014, 514; Lazzarini 2020, 623; Stoker 2006).

Because of their use of social outcomes as payment triggers, OBCs financialize social value such that its appropriation has direct financial cost for outcome payors. This means that it is crucial for outcome payors to be alive to the ability of suppliers to appropriate social value as the success of OBCs financially can – *and does* – decouple from their success socially (Del Giudice and Migliavacca 2019; Hevenstone et al. 2023; Tse and Warner 2020b). The decoupling of contract payment from social success is of critical importance for OBCs, especially those that leverage profit seeking investment as public managers can find themselves paying a premium for specified outcomes – due to a combination of elevated transaction costs associated with developing OBCs and the inclusion of investor returns in outcome prices (FitzGerald, Fraser, and Kimmitt 2020; Lowe et al. 2019; Pandey et al. 2018) – without having made a durable difference in the lives of individuals they intended to help (FitzGerald et al. 2023; Hevenstone et al. 2023).

As tools to transfer the risk of complex social services away from government, OBC suppliers have two avenues for appropriating financialized social value – a supply-side approach of economizing on noncontractible service elements or a demand-side approach of focusing on more profitable service users (Lazzarini 2020). In OBCs, this means that public managers may end up failing to reach the population they are most eager to support with services and paying for outcomes which would have been achieved anyway or are poor indications of positive changes in service user lives (FitzGerald et al. 2019).

### **Outcome specifications: cohort definition, outcome alignment, pricing**

In outcome-based contracts for complex, person-centred public services, suppliers broadly have three mechanisms for prioritizing services to more profitable cohort segments: ‘cherry-picking’ is when service providers refer easier-to-help participants into programmes, ‘cream-skimming’ and ‘parking’ are when service providers concentrate post-referral efforts on easier-to-help participants at the expense of cohort members for whom positive outcomes are likely to be more challenging and costly to achieve (Carter 2021; Carter and Whitworth 2015).

Contract rules to mitigate *cherry-picking* strive for independence within the referral process such that suppliers cannot target and refer easier-to-help individuals onto services. For example, contract rules which circumscribe or identify a participant cohort using routine or independently held data, stipulate independent professional bodies as exclusive referral agents, or include external auditing of referrals ensure that services reach intended cohorts rather than substituting with individuals for whom an outcome payment is easier to achieve. Contract rules to mitigate *creaming-and-parking*

target post-referral behaviours amongst suppliers, incentivizing support to each member of the referred cohort.

Examples of such specification include differentiated payments to compensate for harder-to-reach cohort members, population segmentation to identify and personalize support ex-ante, minimum service offers for all referred participants, capping the number of referrals, and limiting the number of outcomes by type where rate cards (i.e. menus of social outcomes with articulated 'willingness to pay' values from outcome payors) are deployed. Differentiation, segmentation, and personalization all constitute attempts to capture and design for homogeneity within referred cohorts, setting incentives to steer consummate supplier behaviour towards supporting the full range of service users. Capping the number of referrals means that suppliers are prevented from over-recruiting as a way of generating a sizable easier-to-treat subpopulation post-referral. Capping the number of outcomes by type ensures that suppliers are encouraged to support cohort members to achieve the full array of specified outcomes rather than simply generating payments for shorter term outcomes (e.g. service completion; entry into employment) at the expense of longer-term outcomes (e.g. employment sustained for 6 months).

False assumptions around the interchangeability or causal validity of outcome measures mean that suppliers can manipulate proxy outcomes for financial benefit while leaving the underlying social needs unaddressed and potentially eroding positive treatment effects. Purchasers must ensure, then, that payable outcomes meaningfully distinguish successful programmes from failed ones. Public managers can ensure value by specifying outcome measures which are valid and well aligned to their policy objectives. This means that outcomes should be *conceptually aligned* to the underlying social objectives outcomes payors seek (FitzGerald et al. 2019). An example of such alignment can be seen in the HMP Peterborough SIB where the policy objective of the UK Ministry of Justice was to reduce reoffending, and the payable outcome was a cohort-based percentage reduction in reconvictions (FitzGerald et al. 2019).

Contract rules which support conceptual alignment would include minimal use of activity and output payments in outcome specifications (i.e. payments based on engaging cohort members, doing assessments, or entering employment) and minimal use of *renegotiation*, which evidence shows often results in the introduction of activity and output payments after an initial period of underperformance (Ronicle, Stanworth, and Wooldridge 2022).

Finally, if the attribution of outcomes to the intervention is not robust, then those providing services can claim payment for outcomes that would have occurred anyway, the so-called *deadweight* measured through comparison with a 'counterfactual' – a control or comparison group. For programmes which result in well-defined proximate benefits aligned to a purchaser's policy goals (e.g. stepping-down children from residential care into home placements), robust outcome validation may be less critical for reducing opportunism. Rules for outcome validation detail the process by which projects demonstrate the achievement of predefined outcome metrics. This can include building in the results of an experimental or quasi-experimental evaluation into the payable outcomes, benchmarking current performance against a historical baseline, or even relying on self-reported survey data.



## Methods

A configurational approach is used to test whether key features of outcome specifications prevent purchasers from overspending on OBCs (FitzGerald et al. 2019). We explore patterns across the population of completed UK SIBs as of January 2021: 34 SIBs delivering services from 2010 to 2020 (see A2). We systematically compared cases across conditions of cohort definition (*cherry-picking* and *cream-skimming*), outcome alignment (*conceptual alignment* and *renegotiation*), and outcome validation (*deadweight*) as they related to our outcome of interest (Fiss 2011; Rioux 2006; Rioux and Ragin 2008): appropriation as demonstrated by investor return and projects hitting their maximum contracted value across all 34 cases. A strength of using this configurational approach is its ability to account for equifinality, where multiple paths can lead to the same outcome (Rioux and Ragin 2008) and identify conjunctural causation as it does not assume that the causal pathways for achieving an outcome are the same as those that result in failing to achieve the outcome (Kane et al. 2014).

We apply fuzzy-set QCA (fsQCA) to create distinct configurations across our 34 cases. Unlike crisp-set QCA, which requires researchers to code simply the presence or absence of a condition, fsQCA allows researchers to consider the *degree* to which a condition is present in a given case (i.e. present as ‘fully in’ and absence as ‘fully out’) and thus explains the outcome of interest as well as whether an observation is consistent with the presence or absence of appropriation. We explore five conditions amongst our 34 cases – cohort definition as features to mitigate against *cherry-picking* and *cream-skimming*; outcome alignment as *conceptual alignment* and *renegotiation*; and outcome validation as *deadweight* – well within the recommended maximum of seven conditions for this number of cases (Marx and Dusa 2011).

## Data collection, measurement, thresholds, and calibration

To develop the case knowledge, a crucial step in configurational analysis, the 34 SIB cases were divided amongst the five-person research team (Schneider and Wagemann 2010). Two team members retained sight of the total population of completed UK SIBs with the remaining three researchers covering subsets. Using grey and academic literature, policy documents, publicly available data, the research team generated rich project descriptions across the configurational conditions of interest. Documents included programmatic strategy documents and press releases, invitations to tender and procurement documents, including published rate cards, and independent evaluations. To validate researchers’ interpretations of case information, a further six interviews were undertaken with individuals possessing detailed knowledge which combined and covered the totality of SIBs included in the analysis. Interviews involved representatives from two intermediary organizations, two third-party evaluators, a not-for-profit service provider, and a grants and funding administrator.

Then, with reference to project descriptions and using customized survey scales, three research team members were asked to assign values of 1–5 indicating the presence of each condition of interest on each project. For example, the highest score for features to mitigate cherry-picking (5) would include evidence that a given project employs administrative data to generate a named cohort and the lowest score (1) would include projects that allow anyone to make referrals onto the service. Regarding outcomes alignment, high conceptual alignment (5) is reserved for projects



whose payable outcomes are the same as the stated goals of the project. Low conceptual alignment would apply to SIBs that pay primarily on service engagement or process measures. Renegotiation, meanwhile, is a binary variable where 1 represents cases with a formal reprofiling of outcome targets and 0 where reprofiling was not planned or purported to have occurred.

Using averages of these scores for each condition, the remaining two researchers then iteratively calibrated the data provided by other members of the research team applying the recoding calibration method whereby raw data from the survey is used in combination with further case knowledge to establish whether a particular condition is present in each case (Schneider and Wagemann 2010; Thomann, Ege, and Paustyan 2022). This constituted a second round of validation to ensure calibration was robust. Hence, the thresholds used in this analysis principally stem from the scores assigned to cases with calibration adjustments to account for meaningful clusters in the data and variation in case knowledge amongst the research team (Berg-Schlosser and Cronqvist 2012; Casady 2021).

For the outcome of interest for this analysis, we identify instances where payments have been maximized and the possibility of undercutting social value is present. Captured in fsQCA as the absence of value appropriation amongst SIB suppliers – we are interested in whether the maximum contracted value of specified outcomes was paid by public commissioners alongside indications of whether investors also received a return. Whilst it is not possible to directly observe appropriation in our project data, we know that the maximum contracted value or ‘contract cap’ of a SIB is often used by public managers as a tool to prevent excessive and unexpected payment levels in OBCs and that SIB business cases often target a mid-performance scenario to establish feasibility. Likewise, because SIBs are risk-holding instruments, we expect variable performance. Combined, this analysis sheds light on whether patterning associated with enhanced ability for SIB suppliers to appropriate is also associated with projects paying at their maximum value. Thus, a signal of less value appropriation would be understood as a contract paying out below the maximum contracted value of outcomes and too low to provide a return on investment to investor (i.e. ‘more out than in’). Where the maximum contracted value is not hit, but investors still receive a return, projects are considered ‘more in than out’, with ‘fully in’ projects being those where a contract hits its maximum contracted value. While this does not necessarily convey the motivational aspect of appropriative, opportunistic, or perfunctory behaviour, it does indicate where SIB investors and service providers were able to maximize their pay-out and whether those instances overlap with specifications proving the greatest ability for them to appropriate value. The analysis thus suggests instances where commissioners paid for outcomes with a value that is at-best over-estimated or at-worst inflated. See Table 2 for detail on calibration thresholds and supplemental material for raw survey data and calibrated cases.

## Analysis

Upon calibrating conditions to run fsQCA, we established consistency and frequency thresholds for analysis. Consistency captures how well a case that has a particular configuration is associated with the outcome of interest. Here, consistency refers to the percentage of cases within a configuration where there is no evidence of appropriation. Our threshold for analysis is 0.800 as our truth table indicates distinct configurations

**Table 2.** Conditions operationalized for analysis.

|                     | Characteristics                             | Key Condition  | Operationalization   |
|---------------------|---|--|--|
| Cohort Definition   | Features to mitigate 'cherry-picking'       | Impartiality/independence in referral process            | Fully out = open and self-referral<br>More out than in = referral from professional bodies with vested interest in project success<br>More in than out = referral from professional multi-stakeholder panel<br>Fully in = cohort identified and referred using routine/independently managed database<br>More out than in = referral from professional bodies with vested interest in project success<br>More in than out = referral from professional multi-stakeholder panel<br>Fully in = cohort identified and referred using routine/independently managed database                         |
|                     | Features to mitigate 'creaming-and-parking' | Incentives to provide support to each member of cohort   | Fully out = none or only individual payment caps<br>More out than in = only referral number or outcome type caps present<br>More in than out = named cohort, differential pricing, or some cohort-based payments<br>Fully in = majority cohort-based payments and named cohort or claimant leverage  |
| Outcome Alignment   | Conceptual alignment                        | Paid outcomes are congruent to overarching policy intent | Fully out = attachment/assessment payments present with reliance on process payments<br>More out than in = reliance on process payments with few payments tied to policy intent<br>More in than out = process payments marginal with substantial payment tied to policy intent<br>Fully in = payable outcomes congruent with policy intent<br>More out than in = reliance on process payments with few payments tied to policy intent<br>More in than out = process payments marginal with substantial payment tied to policy intent<br>Fully in = payable outcomes congruent with policy intent |
|                     | Renegotiation                               | Short- and long- term outcomes not interchangeable       | Fully out = formal point for reconsideration/reprofiling anticipated or reported<br>Fully in = reprofiling not formally part of project plans or not reported  |
| Outcomes validation |   | Consideration of non-intervention outcomes               | Fully out = no consideration of non-intervention outcomes and use of self-reported or administrative data<br>Fully in = deadweight estimate of any quality (i.e. benchmark, baseline, or comparison) used in price setting and validation or deadweight demonstrably close to 0<br>Fully in = deadweight estimate of any quality (i.e. benchmark, baseline, or comparison) used in price setting and validation or deadweight demonstrably close to 0  |

*(Continued)*

**Table 2.** (Continued).

| Characteristics | Key Condition | Operationalization   |
|-----------------|---------------|--|
| Appropriation   |               | Fully out = project does not hit maximum contracted value and investors lose money<br>More out than in = project does not hit maximum contracted value and investors do not make a return<br>More in than out = project does not hit cap and investors do make a return<br>Fully in = project hits maximum contracted value and investors make a return<br><br>More out than in = project does not hit maximum contracted value and investors do not make a return<br>More in than out = project does not hit cap and investors do make a return<br>Fully in = project hits maximum contracted value and investors make a return |

with low appropriation cases consistent with these configurations. Looking to frequency, because our sample size is small, we consider only configurations with at least one observed case (as recommended by Rihoux and Ragin,2008). Further, we use the intermediate solution for analysis as it accounts for all possible logical combinations whether observed in our sample or not (Fiss 2011). Because fsQCA does not assume equifinality, we run the analysis two ways. First, we explore configuration pathways which avoid hitting their maximum contracted value to understand pathways which mitigate appropriation. Then, we explore configurational pathways which do hit their maximum contracted value to understand what conditions do not mitigate appropriation.

**Configurations and illustrative cases**

Table 3 shows the results of the analysis of conditions resulting in projects not hitting their maximum contracted valued. Our interpretation follows convention: necessary and central conditions are represented by (N) and central conditions by (C). Necessary conditions are those which appear in all configurations consistent with an outcome. As such, they can be thought of as a pre-requisite for hitting – or not hitting – maximum

**Table 3.** Pathways mitigating appropriation.

| Conditions           |                      | Path 1       | Path 2       |
|----------------------|----------------------|--------------|--------------|
| Cohort definition    | Cherry Picking       | Presence     | Absence (C)  |
|                      | Creaming-and-parking | Presence (N) | Presence     |
| Outcome alignment    | Conceptual alignment | Presence (N) | Presence     |
|                      | Renegotiation        | Presence (N) | Presence     |
| Pricing              | Deadweight           | Absence      | Presence (C) |
| Consistency          |                      | 0.853448     | 1            |
| Raw coverage         |                      | 0.238842     | 0.160434     |
| Unique coverage      |                      | 0.238842     | 0.160434     |
| Number of cases      |                      | 1            | 1            |
| Solution consistency |                      | 0.906849     |              |
| Solution coverage    |                      | 0.399276     |              |

contracted value. Central conditions, meanwhile, are those conditions which appear in both intermediate and parsimonious solutions. Conditions not marked by (N) or (C) are contributing conditions, while blanks indicate that a condition is not important within that configuration (i.e. a 'don't care') (Lazzarini et al. 2020).

Looking at Table 3, we see two paths for avoiding maximum contracted value. Path 1 differs from Path 2 in two important ways. First, features to mitigate cream-skimming, conceptual alignment, and renegotiation are present and necessary, meaning neither path includes projects with formal renegotiation points or evidence of reprofiling payable outcomes. Deadweight estimation, meanwhile, is absent from Path 1. In Path 2, deadweight estimation is present and central, whilst features to mitigate creaming-and-parking, conceptual alignment, and renegotiation are all present. However, the absence of features to mitigate cherry-picking is central.

An illustrative example of Path 1 is Futureshapers. Launched in 2015 in Sheffield as part of the UK Department for Work and Pensions (DWP) Youth Engagement Fund (YEF), Futureshapers supported young people aged 14–17 for up to 3 years who were deemed 'most-disadvantaged' and at risk of being long term out of education, training, or employment. The intended cohort of participants included young people with poor school attendance, who were excluded from school, experienced the youth justice system, were under local authority care, were diagnosed with special needs or disability, or who were teen parents. To identify eligible young people, the Council generated an initial list of individuals using internal data sources which were then cross-referenced by the service provider – a local non-profit, Sheffield Futures – and representatives from local schools to further target the list (features to mitigate cherry-picking). Importantly, as a YEF project, once Futureshapers had engaged the number of young people they targeted in their project proposal to the DWP, they were not able to support additional young people. Payable outcomes were also set by the DWP, and each outcome could only be claimed once per participant with a total cost payable to each participant capped at £11,800 – notionally the total cost of unemployment benefits to a young person over 3 years – although the unit costs as outlined in bids were below this ceiling (features to mitigate creaming-and-parking). Flexible outcome caps were also in place: Futureshapers could not treat profiled outcomes interchangeably: it was not possible to be paid for 'over-performance' on some metrics whilst, underperforming on others and maintaining the overall payment envelope (conceptual alignment). Self-reported data was used to validate outcome payments (absent deadweight).

Reconnections, from the Commissioning Better Outcomes Fund, illustrates Path 2. Based in Worcestershire, this SIB provided one-to-one support for lonely individuals over 50 years old from 2015 to 2020. Referrals into the service were made by primary care providers, social care services, social landlords, other VCSE organizations, family, friends, and self-referrals (absent features to mitigate cherry-picking). To be referred, individuals had to have an elevated score in a validated loneliness scale or experience at least five risk factors for loneliness including living alone, being single, divorced, never married or low income. Payable outcomes included decreases in the validated loneliness scale across the treatment cohort at months six and 18 (deadweight estimate based on cohort benchmark).

In Table 4, we observe four pathways associated with projects hitting maximum contracted payments. Path 3 includes the absence of features to mitigate cream-skimming, conceptual alignment, and deadweight estimation whilst renegotiation

**Table 4.** Pathways allowing appropriation.

| Conditions           |                      | Path 3      | Path 4a     | Path 4b     | Path 5       |
|----------------------|----------------------|-------------|-------------|-------------|--------------|
| Cohort definition    | Cherry Picking       | Absence (C) | Absence     | Presence    | Presence     |
|                      | Creaming-and-parking |             |             | Presence    | Absence (C)  |
| Outcome alignment    | Conceptual alignment | Absence     | Absence     | Presence    | Presence (C) |
|                      | Renegotiation        | Absence (C) | Presence    |             | Presence     |
| Pricing              | Deadweight           |             | Absence (C) | Absence (C) | Presence     |
| Consistency          |                      | 1           | 0.950599    | 1           | 1            |
| Raw coverage         |                      | 0.545702    | 0.493971    | 0.19331     | 0.0517308    |
| Unique coverage      |                      | 0.18203     | 0.117075    | 0.0132244   | 0.0517308    |
| Number of cases      |                      | 16          | 15          | 2           | 1            |
| Solution consistency |                      | 0.966514    |             |             |              |
| Solution coverage    |                      | 0.740957    |             |             |              |

and features to mitigate cherry picking are ‘don’t care’. Path 4a also includes an absence of deadweight estimation and conceptual alignment whilst including the presence of renegotiation. Both aspects of cohort definition are ‘don’t care’. Like Paths 3 and 4a, deadweight estimation is absent from Path 4b, and like Path 4a, renegotiation is present, but so too are features to mitigate cherry picking and creaming-and-parking. Conceptual alignment is ‘don’t care’. Paths 4a and 4b are neutral permutations in that they share the same central conditions but have different contributing conditions. Consequently, we interpret them as the same path because their variations do not affect the ‘overall performance of the configuration’ (Fiss 2011, 398). In Path 5, meanwhile, deadweight estimation, renegotiation, and features to mitigate cherry picking are present, conceptual alignment is present and central, and features to mitigate creaming-and-parking is absent and central.

The DWP Innovation Fund (IF) project Links4Life illustrates Path 3. A precursor fund to the YEF, IF projects like Links4Life specified that self-reported data would be used to authenticate outcome achievement, including payment for early-term outputs like improved school attendance and behaviour (absent deadweight). In Links4Life specifically, link workers in East London offered personalized one-to-one support to 14–19-year-olds to either stay in or re-enter education, training, or employment. The referral pathways for Links4Life – but not all IF projects – included schools, local authority mental health and leaving care teams, local youth service providers and self-referrals. Unlike in Futureshapers projects, the DWP allowed Links4Life to refer participants in excess of targeted cohort numbers and reprofile outcome achievement – i.e. over deliver on improved school attendance but under deliver on higher level qualification attainment – such that outcomes were viewed as interchangeable within the same per capita payment envelope (absent features to mitigate cherry-picking and creaming-and-parking as well as absent conceptual alignment).

Paths 4a and 4b are illustrated by St. Basil’s and Mayday respectively. St. Basil’s ran from 2015 to 2018 in the West Midlands as part of the Ministry of Housing, Communities and Local Government’s (MHCLG) Fair Chance Fund (FCF) supporting 18–24-year-olds who were not in education, employment or training that were also identified as homeless but not deemed priority need for local authority supported accommodation. With outcomes defined by MHCLG, St. Basil’s was paid on the basis of completing a series of assessments with service users as well as entry and sustainment of

accommodation, education/training, and volunteering and employment as well as achievement of qualifications through self-reported data (absent conceptual alignment and deadweight). Mayday, inspired by the FCF design, targeted a very similar cohort in Northamptonshire with an expanded age range (18–30 years old) and slightly broader eligibility conditions with referrals coming through professional support providers granting access to individuals with significant mental health issues, substance misuse, low/medium learning disabilities or personality disorders not eligible for support under the Fair Access to Care Services criteria, or individuals simply unable to be accommodated in supporting housing because of previous difficulties or a lack of available specialist support in Northamptonshire (features to mitigate cherry-picking).

Manchester Multi-dimensional Treatment Foster Care-Adolescents (MTFC-A) exemplifies Path 5. Running from 2014 to 2017, this SIB trained foster carers to provide MTFC-A, an evidence-based intervention, as an alternative to residential care for 11–14-year-olds currently in local authority care. Young people were referred to the programme by Manchester City Council. Upon receipt of a referral, programme staff contacted key professionals already supporting the young person to gather information about their ability to understand the service and be committed to it. For young people with positive assessments, the referral was then discussed with potential foster carers who could give agreement for the placement (features to mitigate cherry-picking). Payable outcomes included continued service engagement and increased movement out of residential and into family setting compared to a historic baseline. Payments were also made for the number of weeks spent out of residential care compared to a historic baseline over a 2.5-year period and the achievement of well-being outcomes at programme graduation and 12-months post completion (present deadweight).

## Discussion

Paths 1 and 2 suggest something of a specification trade-off for public purchasers: effort in integrating deadweight estimation in the outcome specification allows for looser restrictions on cohort definition, particularly features to mitigate cherry-picking. Likewise, where deadweight cannot be estimated, public managers can instead rely on definitions of cohort and outcome alignment to mitigate the likelihood that social value is appropriated by suppliers. Paths 3, 4a, 4b, and 5, meanwhile, underscore the importance of including a deadweight estimate as part of outcome validations: in three paths, the absence of deadweight estimation is a central condition for maximizing payment. In sum, across the five conditions, we see that the presence of features to mitigate creaming and parking, conceptual alignment, and deadweight estimation do diminish the likelihood that public managers pay for social outcomes of questionable value.

Extant evaluation material supports these findings and shows that in some instances, financially successful projects failed to generate durable social benefits. Results from the quasi-experimental evaluation of Innovation Fund (IF) projects – in which most projects hit their contract cap – show that the aggregate impact estimate of IF projects is negative for later-term outcomes (e.g. higher level qualifications) but positive for nearer-term outputs (e.g. entry-level qualifications) as compared to a propensity-score matched comparison group. Analyses found that overall, the

Fund did not achieve value-for-money as many of the outcomes would have been achieved regardless (Salis, Wishart, and McKay 2018). Notably, these projects consistently lacked contract features associated with clear cohort definitions, outcome alignment, and deadweight estimates in pricing. Results from HMP Peterborough and Rough Sleeping projects provide an interesting counterpoint to this narrative. While neither project hit their contract cap, impact estimates of the overarching policy aim – reduced reconvictions and rough sleeping – were positive (Anders and Dorsett 2017; Spurling 2017).

Results of the fsQCA in tandem with extant evaluation suggest that outcome specifications characterized by a greater ability for SIB suppliers to appropriate value are associated with projects hitting their contract cap and investors making a return. The corollary is also shown to be true: configurations which do not hit their contract cap are characterized by a lesser ability for SIB suppliers to appropriate value. While this study goes some way in revealing the linkages between outcome specifications and outcome payment, value as delivered by OBCs and SIBs remains somewhat impenetrable. Critically, designing robust outcome specifications is not limited to simply selecting outcome metrics and setting financial incentives, but extends to specifying the referral pathways through which individuals access services. Public managers must consider supply-side *and* demand-side avenues for value appropriation, designing with contract incompleteness *and* cohort profitability in mind. The case of IF projects should be cautionary – failing to cap referrals whilst allowing suppliers to claim for outcome payments interchangeably created a scenario wherein overperformance on near-term outputs obscured under-performance on late-term, more conceptually aligned outcomes.

## Conclusion

OBCs are not an immediate or automatic route to unlock social value and, as this analysis shows, the intricacies of outcome specifications and contract design are particularly important as they can exacerbate or constrain perfunctory behaviour amongst investors and providers. Evaluations and the fsQCA results show links between outcome specifications, value appropriation, and the financial and nonfinancial impacts of interventions. Whilst the full impact of appropriation is not known, this analysis shows that projects lacking features to mitigate against it can be associated with intervention *ineffectiveness* and may result in overpayment, corroding social value overall.

Importantly, this article does have limitations. Whilst case knowledge was developed across the research team, uniform information was not available across all SIB projects. This was especially true for information about their performance, a known limitation of research on SIBs (Fraser and FitzGerald 2019) and an awkward irony given their espoused strengths in data-led performance management and proclivity to be regarded as successful by stakeholders (Carter et al. 2018; Hevenstone et al. 2023). There are also challenges capturing appropriation. While we have endeavoured to operationalize this by documenting patterns based on variations in outcome specification and outcome payments, we acknowledge that we may only have a partial signal of this phenomenon. For instance, our analysis does not capture specific appropriative behaviours nor individual or organizational motivations amongst stakeholders. We are instead reliant on signals of divergence between financial success and delivery of social



value as conveyed through publicly available material and semi-structured interviews. Likewise, a lack of high-quality impact and economic evaluation across studied projects makes it impossible to unpick where financially successful projects also delivered credible social value as conveyed through positive treatment effects. This speaks to a wider pattern of absconding from ‘gold-standard’ evaluative methods to trigger payment within UK OBCs due in dual parts to ethical concerns around randomizing or restricting treatment access and catering to investor appetite. Consequently, the generalizability of these findings may be limited outside of the UK given the particularities of the domestic social investment market and frequent use of rate cards in the early stages of developing the model (Williams 2019).

We maintain, however, that the lessons on outcome specifications derived from the analysis are broadly transferable given the pioneering role the UK has played in the development and transference of OBCs internationally. This article is the first to provide a systematic comparison of OBC projects across policy areas and makes a demonstrable contribution by offering a novel analysis of an entire population of OBCs enacted within the UK. Future research should consider further analyses which include non-contractual conditions which undoubtedly influence the behaviour of providers delivering against outcome-based contracts but are not captured here. For example, implementation factors including managerial approach or oversight capacity of contract managers and the legal form and makeup of partner organizations may prove influential. Relational contracting features remain of interest, including whether governance routines meaningfully promote an articulation of shared aims and better balance financial and social value while advancing more holistic and collaborative performance management. With the publication of further performance information, new operationalizations of the outcome of interest may also be possible which, in tandem with expanding the pool of cases to include other OBCs and SIBs as they come to term, will build the explanatory value of similar analyses and better illuminate the internal mechanisms of these arrangements.

## Acknowledgments

The authors would like to thank Dr Christi Economy for her help with data collection and analysis as well as Dr Alec Fraser, Dr Thomas Elston, and Prof Sergio Lazzarini for their comments on earlier versions of this paper.


The Government Outcomes Lab (GO Lab) and the Blavatnik School of Government, University of Oxford, is supported by the UK Department for Digital, Culture, Media and Sport (DCMS) under grant number 360 G-cabinetoffice-GA-094926. The views expressed in this publication are those of the author and not necessarily those of the DCMS, its arm’s length bodies, or other government departments.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## ORCID

Clare FitzGerald  <http://orcid.org/0000-0002-1021-8041>

Stefanie Tan  <http://orcid.org/0000-0002-6152-9851>

Eleanor Carter  <http://orcid.org/0000-0001-5774-7728>

## References

- Anders, J., and R. Dorsett. 2017. *Peterborough Social Impact Bond: Final Report on Cohort 2 Analysis*. Ministry of Justice. Accessed October 26, 2022. <https://www.gov.uk/government/publications/final-results-for-cohort-2-of-the-social-impact-bond-payment-by-results-pilot-at-hmp-peterborough>.
- Berg-Schlosser, D., and L. Cronqvist (2012) 'Calibrating and Aggregating Multi-Dimensional Socio-Economic and Political Indices with Fuzzy Set Scores'. Rochester, NY. Available at: <https://papers.ssrn.com/abstract=2108797> (Accessed: 4 August).
- Brown, T. L., M. Potoski, and D. M. V. Slyke. 2018. "Complex Contracting: Management Challenges and Solutions." *Public Administration Review* 78 (5): 739–747. <https://doi.org/10.1111/puar.12959>.
- Brown, T. L., M. Potoski, and D. V. Slyke. 2016. "Managing Complex Contracts: A Theoretical Approach." *Journal of Public Administration Research and Theory* 26 (2): 294–308. <https://doi.org/10.1093/jopart/muv004>.
- Bryson, J. M., B. C. Crosby, and L. Bloomberg. 2014. "Public Value Governance: Moving Beyond Traditional Public Administration and the New Public Management." *Public Administration Review* 74 (4): 445–456. <https://doi.org/10.1111/puar.12238>.
- Carter, E. 2021. "More Than Marketised? Exploring the Governance and Accountability Mechanisms at Play in Social Impact Bonds." *Journal of Economic Policy Reform* 24 (1): 78–94. <https://doi.org/10.1080/17487870.2019.1575736>.
- Carter, E., C FitzGerald, R Dixon, C Economy, T Hameed, and M Airoidi. 2018 *Building the Tools for Public Services to Secure*. University of Oxford, p. 21. <https://waystowellness.org.uk/site/assets/files/1317/2018-bsg-golab-evidencereport.pdf>.
- Carter, E., and A. 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): 277–296. <https://doi.org/10.1017/S0047279414000841>.
- Casady, C. B. 2021. "Examining the Institutional Drivers of Public-Private Partnership (PPP) Market Performance: A Fuzzy Set Qualitative Comparative Analysis (fsQca)." *Public Management Review* 23 (7): 981–1005. <https://doi.org/10.1080/14719037.2019.1708439>.
- Cohen, R. (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>
- Del Giudice, A., and M. Migliavacca. 2019. "Social Impact Bonds and Institutional Investors: An Empirical Analysis of a Complicated Relationship." *Nonprofit and Voluntary Sector Quarterly* 48 (1): 50–70. <https://doi.org/10.1177/0899764018797480>.
- Fiss, P. C. 2011. "Building Better Causal Theories: A Fuzzy Set Approach to Typologies in Organization Research." *The Academy of Management Journal* 54 (2): 393–420. <https://doi.org/10.5465/amj.2011.60263120>.
- FitzGerald, C., E. Carter, R. Dixon, and M. Airoidi. 2019. "Walking the Contractual Tightrope: A Transaction Cost Economics Perspective on Social Impact Bonds." *Public Money & Management* 39 (7): 458–467. <https://doi.org/10.1080/09540962.2019.1583889>.
- FitzGerald, C., A. Fraser, and J. Kimmitt. 2020. "Tackling the Big Questions in Social Impact Bond Research Through Comparative Approaches." *Journal of Comparative Policy Analysis: Research and Practice* 22 (2): 85–99. <https://doi.org/10.1080/13876988.2020.1726177>.
- FitzGerald, C., A. Fraser, J. Kimmitt, L. Knoll, and J. Williams. 2023. "'Outcomes-Based Contracting and Public Management Reform: Lessons from a Decade of Experimentation'." *International Public Management Journal* 26 (3): 329–338. Accessed: 14 April 2023. [https://kclpure.kcl.ac.uk/portal/en/publications/outcomesbased-contracting-and-public-management-reform-lessons-from-a-decade-of-experimentation\(5868a975-1523-4da0-af13-4211988e1ec8\).html](https://kclpure.kcl.ac.uk/portal/en/publications/outcomesbased-contracting-and-public-management-reform-lessons-from-a-decade-of-experimentation(5868a975-1523-4da0-af13-4211988e1ec8).html).
- FitzGerald, C., T. Hameed, F. Rosenbach, J R Macdonald, and R Dixon. 2021. "Resilience in Public Service Partnerships: Evidence from the UK Life Chances Fund." *Public Management Review*: 1–21. <https://doi.org/10.1080/14719037.2021.2015186>.

- FitzGerald, C., F. Rosenbach, T. Hameed, R. Dixon, and J. Blundell. 2021. "New Development: Rallying Together—The Rationale for and Structure of Collaborative Practice in England." *Public Money & Management*: 1–4. <https://doi.org/10.1080/09540962.2021.1981628>.
- Fraser, A., et al. (2018) 'Narratives of Promise, Narratives of Caution: A Review of the Literature on Social Impact Bonds', *Social Policy & Administration*, 52(1), pp. 4–28. Available at: <https://doi.org/10.1111/spol.12260>.
- Fraser, A., and C. FitzGerald. 2019. "Social Impact Bonds Could Benefit from Greater transparency'." *Public Finance*. Accessed November 22, 2022. <https://www.publicfinance.co.uk/opinion/2019/12/social-impact-bonds-could-benefit-greater-transparency>.
- French, M., J. Kimmitt, R. Wilson, D. Jamieson, and T. Lowe. 2022. "Social Impact Bonds and Public Service Reform: Back to the Future of New Public Management?" *International Public Management Journal*: 1–20. <https://doi.org/10.1080/10967494.2022.2050859>.
- Garcia-Castro, R., and R. V. Aguilera. 2015. "Incremental Value Creation and Appropriation in a World with Multiple Stakeholders." *Strategic Management Journal* 36 (1): 137–147. <https://doi.org/10.1002/smj.2241>.
- Hart, O. D. 1988. "Incomplete Contracts and the Theory of the Firm Conference Papers to Celebrate the Fiftieth Anniversary of the Nature of the Firm'." *Journal of Law, Economics, & Organization* 4 (1): 119–140.
- Heinrich, C. J. 1999. "Do Government Bureaucrats Make Effective Use of Performance Management Information?" *Journal of Public Administration Research and Theory* 9 (3): 363–394. <https://doi.org/10.1093/oxfordjournals.jpart.a024415>.
- Heinrich, C. J., and S. E. Kabourek. 2019. "Pay-For-Success Development in the United States: Feasible or Failing to Launch?" *Public Administration Review* 79 (6): 867–879. <https://doi.org/10.1111/puar.13099>.
- Hevenstone, D., A. Fraser, L. Hobi, W. Przepiorka, and Gemma G. M. Geuke. 2023. "The Impact of Social Impact Bond Financing'." *Public Administration Review* 83 (4): 930–946. [https://kclpure.kcl.ac.uk/portal/en/publications/the-impact-of-social-impact-bond-financing\(9b8be82f-73f2-4a5f-b0fe-e8cee91a36a2\).html](https://kclpure.kcl.ac.uk/portal/en/publications/the-impact-of-social-impact-bond-financing(9b8be82f-73f2-4a5f-b0fe-e8cee91a36a2).html).
- INDIGO. 2023. *Impact Bond Dataset*. Accessed March 31 2023. <https://golab.bsg.ox.ac.uk/knowledge-bank/indigo/impact-bond-dataset-v2/>.
- Isaksson, D., P. Blomqvist, and U. Winblad. 2018. "Privatization of Social Care Delivery – How Can Contracts Be Specified?" *Public Management Review* 20 (11): 1643–1662. <https://doi.org/10.1080/14719037.2017.1417465>.
- Kane, H., M. A. Lewis, P. A. Williams, and L. C. Kahwati. 2014. "Using Qualitative Comparative Analysis to Understand and Quantify Translation and Implementation." *Translational Behavioral Medicine* 4 (2): 201–208. <https://doi.org/10.1007/s13142-014-0251-6>.
- Klijn, E. H., and J. Koppenjan "The Impact of Contract Characteristics on the Performance of Public–Private Partnerships (PPPs)." *Public Money & Management*, 36 (6): 455–462. 2016. Available at. <https://doi.org/10.1080/09540962.2016.1206756>.
- Kroeger, A., and C. Weber. 2014. "Developing a Conceptual Framework for Comparing Social Value Creation." *The Academy of Management Review* 39 (4): 513–540. <https://doi.org/10.5465/amr.2012.0344>.
- Lake, R. W. "The Financialization of Urban Policy in the Age of Obama." *Journal of Urban Affairs*, 37 (1): 75–78. 2015. Available at. <https://doi.org/10.1111/juaf.12167>.
- Lazzarini, S. 2020. "The Nature of the Social Firm: Alternative Organizational Forms for Social Value Creation and Appropriation." *Academy of Management Review* 45 (3): 620–645. <https://doi.org/10.5465/amr.2018.0015>.
- Lazzarini, S. G. 2022. *The Right Privatization: Why Private Firms in Public Initiatives Need Capable Governments*. Cambridge University Press.
- Lazzarini S G, Pongeluppe L S, Ito N C, Oliveira F de and Ovanessoff A. (2020). Public Capacity, Plural Forms of Collaboration, and the Performance of Public Initiatives: A Configurational Approach. *Journal of Public Administration Research and Theory*, 30(4), 579–595. [10.1093/jopart/muaa007](https://doi.org/10.1093/jopart/muaa007)
- Lowe, L., J. Kimmitt, R. Wilson, M. Martin, and J. Gibbon. 2019. "The Institutional Work of Creating and Implementing Social Impact Bonds." *Policy & Politics* 47 (2): 353–369. <https://doi.org/10.1332/030557318X15333032765154>.

- Marx, A., and A. Dusa. 2011. "Crisp-Set Qualitative Comparative Analysis (csQca), Contradictions and Consistency Benchmarks for Model Specification." *Methodological Innovations Online* 6 (2): 103–148. <https://doi.org/10.4256/mio.2010.0037>.
- Maurya, D., and A. K. Srivastava. 2019. "Managing Partner Opportunism in Public–Private Partnerships: The Dynamics of Governance Adaptation." *Public Management Review* 21 (10): 1420–1442. <https://doi.org/10.1080/14719037.2018.1559341>.
- Moynihan, D. P., S. Fernandez, S. Kim, K. M. LeRoux, S. J. Piotrowski, B. E. Wright, K. Yang, et al. 2011. "Performance Regimes Amidst Government Complexity." *Journal of Public Administration Research and Theory: J-Part* 21 (Supplement 1): 141–155. <https://doi.org/10.1093/jopart/muq059>.
- Mulgan, G., N. Reeder, M. Aylott, and L. Bo'sher. 2011. *Social Impact Investment: The Challenge and Opportunity of Social Impact Bonds*. London: Young Foundation.
- Newhouse, J. P. 1984. "Cream Skimming, Asymmetric Information, and a Competitive Insurance Market." *Journal of Health Economics* 3 (1): 97–100. [https://doi.org/10.1016/0167-6296\(84\)90030-4](https://doi.org/10.1016/0167-6296(84)90030-4).
- Pandey, S., J. J. Cordes, S. K. Pandey, and W. F. Winfrey. 2018. "Use of Social Impact Bonds to Address Social Problems: Understanding Contractual Risks and Transaction Costs." *Nonprofit Management and Leadership* 28 (4): 511–528. <https://doi.org/10.1002/nml.21307>.
- Rihoux, B. 2006. "Qualitative Comparative Analysis (QCA) and Related Systematic Comparative Methods: Recent Advances and Remaining Challenges for Social Science Research." *International Sociology* 21 (5): 679–706. <https://doi.org/10.1177/0268580906067836>.
- Rihoux, B., and C. C. Ragin. 2008. *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. SAGE Publications.
- Romzek, B. S., and J. M. Johnston. "State Social Services Contracting: Exploring the Determinants of Effective Contract Accountability." *Public Administration Review*, 65:436–449. 2005. Available at.
- Ronicle, J., N. Stanworth, and R. Wooldridge. 2022. *Commissioning Better Outcomes Third Update Report*. <https://golab.bsg.ox.ac.uk/documents/CBO-3rd-update-report.pdf>.
- Salis, S., R. Wishart, and S. McKay (2018) Evaluation of the Innovation Fund Pilot: Quantitative Assessment of Impact and Social Return on Investment. 956. London: UK Department for Work & Pensions. Available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/737021/evaluation-of-the-innovation-fund-pilot-quantitative-assessment-of-impact-and-social-return-on-investment.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/737021/evaluation-of-the-innovation-fund-pilot-quantitative-assessment-of-impact-and-social-return-on-investment.pdf).
- Schneider, C. Q., and C. Wagemann. 2010. "Standards of Good Practice in Qualitative Comparative Analysis (QCA) and Fuzzy-Sets." *Comparative Sociology* 9 (3): 397–418. <https://doi.org/10.1163/156913210X12493538729793>.
- Simon, H. A. 1957. *Models of Man: Social and Rational; Mathematical Essays on Rational Human Behavior in Society Setting*. New York: Wiley.
- Spurling, L. (2017) The Impact Evaluation of the London Homelessness Social Impact Bond. London: Uk Department for Communities and Local Government. Available at: [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/658881/SIB\\_Impact\\_evaluation\\_report.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/658881/SIB_Impact_evaluation_report.pdf).
- Stoker, G. 2006. "Public Value Management: A New Narrative for Networked Governance?" *The American Review of Public Administration* 36 (1): 41–57. <https://doi.org/10.1177/0275074005282583>.
- Swiss, J. E. 1992. "Adapting Total Quality Management (TQM) to Government." *Public Administration Review* 52 (4): 356–362. <https://doi.org/10.2307/3110395>.
- Thomann, E., J. Ege, and E. Paustyan. 2022. "Approaches to Qualitative Comparative Analysis and Good Practices: A Systematic Review." *Swiss Political Science Review* 28 (3): 557–580. <https://doi.org/10.1111/spsr.12503>.
- Tse, A. E., and M. E. Warner. 2020a. "A Policy Outcomes Comparison: Does SIB Market Discipline Narrow Social Rights?" *Journal of Comparative Policy Analysis: Research and Practice* 22 (2): 134–152. <https://doi.org/10.1080/13876988.2019.1609789>.
- Tse, A. E., and M. E. Warner. 2020b. "The Razor's Edge: Social Impact Bonds and the Financialization of Early Childhood Services." *Journal of Urban Affairs* 42 (6): 816–832. <https://doi.org/10.1080/07352166.2018.1465347>.
- Van De Ven, W. P. M. M., and R. C. J. A. Van Vliet (1992) 'How can we prevent cream skimming in a competitive health insurance market?: The great challenge for the 90's', in P. Zweifel and H.E. Frech (edited by) *Health Economics Worldwide*. Dordrecht: Springer Netherlands (Developments

- in Health Economics and Public Policy), 23–46. Available at: [https://doi.org/10.1007/978-94-011-2392-1\\_2](https://doi.org/10.1007/978-94-011-2392-1_2).
- Warner, M. E. “Private Finance for Public Goods: Social Impact Bonds.” *Journal of Economic Policy Reform*, 164:303–319. 2013. Available at.
- Williams, J. W. 2019. “From Visions of Promise to Sign of Struggle: Exploring Social Impact Bonds and the Funding of Social”. Accessed March 31, 2023. <https://policycommons.net/artifacts/1736812/from-visions-of-promise-to-sign-of-struggle/2468428/>.
- Williamson, O. E. 1985. *The Economic Institutions of Capitalism*. New York: Simon and Schuster.