

Alternate Strategies for Incomplete Public Contracting - Assessing Cost Effectiveness

Elaine De Gruyter¹, Maria Patouna¹, Eleanor Carter^{1*}

¹ Blavatnik School of Government
Government Outcomes Lab
University of Oxford
120 Walton St
Oxford OX2 6GG
UK

* Presenting author and author for SIOE correspondence: Eleanor Carter:
eleanor.carter@bsg.ox.ac.uk

ABSTRACT

Incomplete contracts are a core challenge faced by every government buyer. Developing contracts with the right incentives and safeguards so public money is wisely spent and citizens receive quality services, while keeping transaction costs proportionate is a difficult balancing act. One innovation is the use of outcomes-based contracts (OBC). Unlike traditional contracts like fee-for-service (FFS) where payment is based on inputs or outputs, OBCs respond to contract incompleteness by shifting the focus to outcomes which may be more readily verifiable and more congruent with overarching contract objectives. This means that government buyers can focus on outcomes achieved and providers have greater flexibility to adapt service delivery to achieve the public principal's goals.

Despite growing policy interest, empirical evidence on the performance consequences of alternative contracting modalities remains limited (Schepker et al., 2014), particularly with respect to comparative assessments of cost effectiveness. We contribute to this emerging evidence base by empirically comparing the cost effectiveness of a conventional FFS contract against an OBC in the context of a complex human service; housing-related support provision in Kirklees, England. Our 'cost' data is collated through systematic analysis of contract data and semi-structured key informant interviews. Our 'effect' estimates are derived through a quasi-experimental analysis comparing individuals' housing and employment outcomes under both contract types, drawing on linked provider management information and UK government administrative data. We find that the OBC delivers lower per-person costs than the FFS contract (£3,236 vs £4,856 cost in real 2024 GBP). We also find greater effectiveness in the OBC, with a significant increase in participants sustaining independence from housing-related benefits (by 6.15 percentage points) and employment entry (6.20 percentage points). The OBC is therefore less costly and more effective than the FFS contract, with an incremental cost effectiveness ratio of -£12 per additional person sustaining independence from housing-related benefits; and -£11 per additional person entering employment.

Alternate Strategies for Incomplete Public Contracting - Assessing Cost Effectiveness

1. Introduction

Governments channel large amounts of taxpayers' money to private sector and not-for-profit suppliers when they contract for complex goods and services (Cabinet Office, 2019; OECD, 2025). From a public management perspective, service performance, and public value more broadly, are largely contingent on contract performance (Kettl, 2002). Managing the contractual relationship with these non-governmental agents is difficult, however. Operating under limited information, finite time, and cognitive constraints (Simon, 1957), public buyers cannot anticipate all future states or specify enforceable provisions for every relevant contingency; nor can they fully observe or verify key dimensions of supplier effort, service quality, and performance – especially in complex sectors such as human services (Hart et al., 1997). Contracts are inevitably incomplete (Coase, 1937; Segal, 1999). Although Oliver Hart and others suggest that it is a fool's errand to attempt to complete the contract, public managers must take action to guide their agents against opportunism and minimise the extent to which contractual gaps are exploited to appropriate (non-public) value from the exchange. How such market-inspired instruments can effectively and efficiently steer public service delivery remains a persistently elusive question for both scholars and practitioners.

The puzzle, then, is how to best design public sector contracts for complex services in the face of incompleteness. The traditional response is to apply a formal, rules-based approach, emphasising detailed *ex ante* specification of exchange terms to constrain discretion (Brown et al., 2013). Input- and output-based contracts, such as fee-for-service (FFS) contracts exemplify this logic by enforcing and rewarding suppliers according to observable and verifiable activities. This arrangement reduces uncertainty about what the supplier must deliver, but does not ensure that specified activities translate into improved outcomes, potentially undermining contract objectives (Bertelli & Smith, 2010; Domberger & Jensen, 1997; FitzGerald et al., 2023). As an alternative response, outcomes-based contracts (OBCs) seek to steer supplier behaviour by enforcing and rewarding performance against measurable outcomes, while allowing discretion over service delivery inputs and activities. In public settings, discretion cannot rely on a purely trust-based relational contracting approach (Zheng et al., 2008) and is typically scaffolded by formal performance and accountability mechanisms that mitigate risks of ambiguity, capture, and the political or perverse use of performance information (Carter et al., 2026b). In principle, OBCs better align suppliers' effort with contract objectives by channelling discretionary behaviour toward effective and tangible delivery improvements (Bel et al., 2025). OBCs have therefore been promoted as a model for achieving better value for money in public service commissioning (Battye, 2015; Finkenstadt & Cummins, 2026; Office of Management and Budget, 2013).

A further challenge lies in determining which contracting modality is best suited to a given public purchase. While theory suggests that effectiveness depends on the fit between contractual form and contextual conditions, rather than any single universally optimal arrangement (Warsen et al., 2019), systematic comparative evidence on the appropriate match between conditions and contractual form – especially for complex human services – remains

scarce. A recent meta-analysis helpfully summarises the performance of public performance-based contracting, finding that the true effect of performance-linked (in other words, outcome-based contracting) “is positive but small and not statistically significant” (Bel et al., 2025, p. 2) and cautions against the assumption that all incentive contracts deliver genuine performance improvement. Bel et al.’s study finds comparatively weaker performance of outcome-linked schemes in complex human services, noting the need to carefully consider context and sector-specific characteristics in identifying which contracting strategy is most apt. Moreover, the meta-analysis looks at standardised effects only, and does not encompass cost or cost effectiveness analysis, simply noting that performance-based contracting “has been linked to increased administrative costs for the government and providers, given the need to improve data collection and monitoring...” (Bel et al., 2025, p. 1).

The academic literature that credibly surfaces the transaction costs of public contracting is remarkably thin. As FitzGerald et al. (2019, p. 460) observe, “transaction costs are rarely systematically reported by commissioning departments due to commercial sensitivities and measurement difficulties”. This opacity reflects a deeper methodological gap; the field lacks a well-developed approach for systematically collecting and categorising contracting costs, particularly where expenditure must be reconstructed retrospectively across multiple stakeholders and contract cycles. Where cost evidence does exist, it is largely confined to siloed sub-literatures (e.g., PPPs and private finance initiatives, see De Schepper et al., 2015; Pandey et al., 2018; Whittington, 2012). Likewise, comparative estimates of cost effectiveness across contracting approaches remain similarly scarce (Fernandez, 2007, 2009; Hevenstone et al., 2023; Lu, 2016), and rarely account simultaneously for contracting costs and effects (Petersen, 2019; Potoski et al., 2025). The result is a fragmented and methodologically underdeveloped evidence base that cannot reliably inform public management decisions about which contractual modality offers superior value for money.

To begin addressing this gap, we systematically compare the cost effectiveness of a FFS contract against an OBC contract for commissioning a social housing support service in Kirklees, a local authority in northern England. Unlike existing studies, our analysis accounts for both governance costs and performance effects across two contracting modalities within the same service and geographical context. Our transaction and delivery ‘cost’ data is collated through systematic analysis of contract data and semi-structured key informant interviews. Our ‘effect’ estimates are derived through a quasi-experimental analysis comparing individuals’ housing and employment outcomes under the FFS versus the OBC, drawing on linked provider management information and administrative data from the UK Department for Work and Pensions and HM Revenue and Customs. Our results indicate that the OBC delivers lower per-person costs than the FFS contract (£3,236 vs £4,856 cost in 2024 prices). The OBC is associated with improved housing and employment outcomes, with positive effects on sustained housing-related benefit independence and employment entry; impacts on time in work and earnings are modest and not statistically significant. We conclude that the OBC is both less costly and more effective than the FFS contract.

Our study advances ongoing discussions in public sector contracting and performance literature by developing an integrated framework for comparing contract modalities that synthesises incomplete contracts theory, principal-agent theory and transaction cost economics. Analytically, we move the empirical needle by estimating the relative value-for-money of two alternative contractual modalities for the same complex social service. Providing this piece in the puzzle is important for both scholars and practitioners (Gustafsson-Wright & Osborne, 2020; Shiva et al., 2024). For scholarship, cumulative comparative inference has been constrained by the lack of large-scale, standardised data capable of supporting systematic comparisons with robust counterfactuals (Bel et al., 2025). While practitioners suffer from a painfully weak evidence base to inform contract design and commissioning decisions.

The article proceeds as follows. Section 2 sets out the conceptual scaffolding for our work, developing theoretical expectations for how FFS and OBCs structure incentives under contractual incompleteness. Section 3 outlines the empirical setting. Section 4 describes the data, methods and our empirical approach to exploring cost effectiveness. Section 5 reports the results. Section 6 concludes with implications for contract design, limitations, and directions for future research.

2. Theoretical expectations

When governments contract out service provision, they shift from hierarchical, bureaucratic control to governing at arm's length through contracts. *Incomplete contracts* theory (Grossman & Hart, 1986; Hart et al., 1997; Hart & Moore, 1988) explains why this governance approach is inherently challenging; as contract drafters are boundedly rational, they cannot foresee every state of the world or fully specify contractual mechanisms for managing possible contingencies, such as cost-sharing rules or renegotiation provisions (Simon, 1957). Public contracts are therefore inherently incomplete (Coase, 1937; Segal, 1999).

A central implication is that contract success depends not only on what is written in the contract (*specific rights*) but also on who holds *residual control rights* – i.e., decision authority over issues that are not (and often cannot be) specified ex ante (Grossman & Hart, 1986). In outsourced public services, residual control typically lies with the suppliers as the asset owners, which gives them discretion over how services are delivered, including choices about staffing, processes, and other inputs that affect quality. In complex human services, it is particularly challenging to credibly specify dimensions of service quality ex ante, so these contracts tend to leave more residual control, and greater discretion, with suppliers.

Agency theory sharpens this account by specifying the resulting governance hazards. When the exchange between suppliers (agents) and public buyers (principals) is characterised by incomplete information and informational asymmetries, parties may exploit these asymmetries strategically (Arrow, 1984; Eisenhardt, 1989; Jensen & Meckling, 1979; Miller & Whitford, 2006; Sappington, 1991). Suppliers may use contractual gaps (Williamson, 1985) to underinvest in quality unless incentives make service improvement privately worthwhile (Hart et al., 1997). These frictions create scope for “quality shading” – subtle reductions in quality or strategic non-cooperation that are difficult to detect – which lower supplier costs while

eroding public value (Domberger & Jensen, 1997; Esteve & Reyes-Gonzalez, 2020; Levin & Tadelis, 2010; Sclar, 2000). Governments need to design contracts that mitigate these hazards by incentivising suppliers to exercise their residual control in line with public value and overarching contractual objectives.

Furthermore, *transaction cost economics* provides deeper insight into the severity of these hazards, and the relative appropriateness of governance choices, focusing on service attributes and the wider operating environment (Brown & Potoski, 2003; Brunjes, 2020; Hefetz & Warner, 2012; Williamson, 1981). Complex services tend to require greater ex ante specification and ex post evaluation effort (Petersen et al., 2015, 2019). For example, human services, are characterised by heterogeneous tasks and uncertainty about “what works” across clients (*product complexity*). They also involve high *asset specificity* as suppliers make specialised, hard-to-redeploy investments, such as staff upskilling and training (Brown & Potoski, 2003; Hefetz & Warner, 2012; Malatesta & Smith, 2011), increasing exposure to opportunism and hold-up through ex post renegotiation (Williamson, 1985). Meanwhile, contract outcomes are often multidimensional and realised over long horizons, limiting *measurability* and verifiability (Malatesta & Smith, 2014; Piatak & Jensen, 2024). Accordingly, contracting under conditions of high asset-specificity and low measurability is expected to involve greater information asymmetry and higher transaction costs (Williamson, 1996).

The severity of these contracting challenges is also moderated by market structure and environmental conditions. Competition typically disciplines supplier behaviour by sharpening incentives and reducing information asymmetries (Cunat & Guadalupe, 2005; Sklivas, 1987). In high-asset-specificity settings such as human services, thin markets may nonetheless support service quality by enabling scale economies, longer contracting horizons, and more relational governance (DeHoog, 1990; Girth et al., 2012; Warner & Bel, 2008). Environmental volatility introduces further complexity; shocks such as labour-cost inflation or large-scale service disruptions can rapidly render contractual terms obsolete, increasing adaptation costs and renegotiation pressures regardless of contract type (FitzGerald et al., 2021, 2024). Combined with incomplete contracting, these conditions create what Brown et al. (2018) describe as a complex contracting challenge.

Governments rely on a portfolio of responses to navigate this complex contracting challenge including supplier selection and screening, contractual guardrails, monitoring and audit regimes, and incentive provisions to align service provision with public value under uncertainty (Alchian & Demsetz, 1972; Sappington, 1991). Conceptually, we group these responses into two broad strategies for navigating contractual incompleteness:

i. Input and process-based contractual incentives

The conventional response to contractual incompleteness is a formal, rules-based governance strategy that seeks to reduce uncertainty through detailed ex ante specification of exchange terms (Brown et al., 2013). The strategy’s premise is specificity: precise, comprehensive requirements are intended to make suppliers’ behaviour “legible” to public buyers, clarify

expectations, and strengthen accountability (Romzek & Johnston, 2005; Savas, 2000). In practice, this involves codifying deliverables and procedures, screening bidders' capacity, and maintaining close oversight through diligent monitoring of activities and near-term output measures and active contract management, backed by legal enforcement (Kettl, 1993; Marlin, 1984; Savas, 1987, 2000).

Conventional fee-for-service (FFS) contracts embody this rules-based approach to steering provider behaviour. FFS contracts specify inputs and delivery processes in detail – for example, the required equipment and technologies are specified, and contracts may set out prescribed staffing levels and time allocations. Suppliers are paid per unit of activity they provide (such as sessions delivered, clients served, or hours worked) rather than for outcomes achieved (Heinrich & Choi, 2007; Tomkinson, 2016). This modality prioritises procedural compliance, holding suppliers accountable for the means of delivery, and makes expenditure relatively predictable, with routine administrative oversight and stable provider cash-flow.

However, it is difficult to predict *ex ante* which activities will reliably produce the desired outcomes. This challenge is particularly acute in human services, where delivery is labour intensive and hinges on suppliers' discretionary judgement (Hasenfeld, 1983; Lipsky, 1980). Moreover, because compensation is independent of outcomes achievement, suppliers face no incentives to innovate or enhance service quality, as better quality is likely to entail economic inefficiency (Wulczyn, 2005). For example, quality improvements often require recruiting and training specialised staff, while more effective services may reduce future demand. Overall, under FFS contracts, the means and ends of service delivery are weakly coupled, encouraging sub-optimal allocations of effort toward measurable outputs and compliance with specified inputs/procedures rather than harder-to-observe service outcomes (Baker et al., 1994).

ii. Outcomes as contractual incentives

An alternative response to contractual incompleteness is to contract for outcomes. As Lazzarini (2022, p. 99) observes, under an OBC strategy “we are essentially supposing that it is possible to select measures of social value creation and compensate private agents according to those measures: thus contracts become complete.” While completeness remains unattainable in practice, OBCs represent a more direct alignment between contractual specification and residual control than FFS modalities, narrowing the gap between what is specified and what is socially valued. This alignment creates a shared focal point around which principals and agents coordinate effort, with greater supplier discretion over how outcomes are achieved channelling residual control toward experimentation and the adoption of improved delivery methods (Carter et al., 2025; Considine et al., 2020; Patouna et al., 2025). Accordingly, OBCs are theorised to spur innovation and enable more personalised, responsive services (Bovaird & Davies, 2011; FitzGerald et al., 2020; Olson et al., 2024). This may be particularly valuable in heterogeneous, high-uncertainty settings, such as human services, where professional judgement and flexible adaptation are essential to dynamic client needs (Albertson et al., 2018; Arena et al., 2016; Carter et al., 2018; Cox, 2011; Joy & Shields, 2013). From a behavioural perspective, outcome targets may also function as “high goals” – clear, challenging, and meaningful – strengthening goal alignment and ownership relative the compliance logic of

process-based contracts (Bel et al., 2025; Locke & Latham, 1990). Consistent with this mechanism, high-goal elements have been linked to improved performance outcomes (Favero et al., 2016).

OBCs introduce their own set of governance challenges, however. Outcomes are often intangible, multifaceted, long-term, and context-dependent, making them costly to define, measure, and verify (Brown et al., 2006; Hart & Moore, 1988; Hevenstone et al., 2023; Larsen & Wright, 2014). Even where measurable, isolating providers' contributions from participant characteristics and external shocks is difficult, exposing providers to downside risk for factors beyond their control (Gerrish, 2016; Heinrich & Choi, 2007; Hodgkin et al., 2020; Lazzarini et al., 2022). These features can promote narrow target compliance and goal displacement, shifting attention away from harder-to-observe but socially valuable dimensions of quality (Brown & Potoski, 2003; Radin, 2006). In mission-driven sectors, where the principal's objective is not easily "contractible," high-powered incentives may crowd out intrinsic motivation and elicit strategic responses (Baker, 1992, 2002; Dias & Maynard-Moody, 2006; Perry et al., 2010). Consistent with this, 'gaming' (Bevan & Hood, 2006) is well documented in results-based human service contracting, including cynical 'creaming' and 'parking' that inflate measured performance relative to true impact (Carter & Whitworth, 2015; Heinrich, 2007; Koning & Heinrich, 2013; Lowe & Wilson, 2017). This 'performance paradox' (Thiel & Leeuw, 2002) undermines OBC's effectiveness in human services.

Beyond contract success, a complete comparative assessment of contracting modalities must account for the costs of governing the contractual exchange. Transaction costs – defined as "all costs borne by the consumer that are not transferred to the seller of the good" (Wallis & North, 1986, p. 98) – encompass the design, negotiation, monitoring, and enforcement of contractual agreements (Williamson, 1985). Conceptually, there is an ambiguity regarding the relative costs of FFS and OBCs. Bel et al (2025, p. 2) note that "contracts based on measurable outcomes lower transaction costs" (Williamson, 1996), by replacing detailed behavioural monitoring of supplier processes with incentive alignment around desired results but also acknowledge that performance-linked contracts are associated with increased administrative costs for government and providers, given the need to improve data collection and monitoring. Amirkhanyan et al (2010) note that transaction costs for alternative contracting strategies may vary dependent on the time horizon. The net transaction cost implications of OBC relative to FFS are therefore theoretically indeterminate.

Empirical evidence on these cost differentials is also limited. Existing studies suggest that OBC governance costs can be substantial in practice; defining measurable outcomes, establishing credible counterfactuals, pricing attributable outcomes, and maintaining the evidentiary infrastructure required for performance verification impose significant ex ante and ex post burdens on public buyers. FitzGerald et al. (2019) caution that public commissioners must walk a "contractual tightrope," balancing the protective value of tighter outcomes specification against the escalating costs of producing and implementing such contracts. Their analysis of two UK social impact bonds – a prominent form of OBC – finds that neither contract successfully struck this balance. Estimates from another social impact bond in the United States

suggests that transaction costs can reach 10% of total contract value (Pandey et al., 2018), compared with 2-3% typically observed in conventional public-private infrastructure contracts (Dudkin & Väilä, 2006). These figures, however, derive from different empirical contexts and direct comparisons across contracting modalities within the same service and geographical context remain scarce.

Overall, FFS and OBC constitute two alternative strategies for navigating contractual incompleteness, allocating residual control and shifting suppliers' discretion to different parts of the exchange. The former specifies the *means* of delivery, which reduces ambiguity over required delivery resources and process, but cannot ensure that prescribed activities translate into meaningful outcome achievements. The latter formalises the *ends* of delivery, which strengthens accountability for results, yet may also expand the scope for opportunistic behaviour that undermines public value.

While theory-driven arguments have been made for the benefits of OBC in complex contract settings (Brown et al., 2013, 2018; Frydinger & Hart, 2024), much of the conceptual development has emerged from public procurement contexts centred on complex goods and projects (Brown et al., 2013), rather than complex human services. Systematic empirical comparisons of cost effectiveness across contracting modalities – accounting for both performance effects and governance costs within a common empirical setting – are limited. We respond to this gap by offering a rare empirical investigation of cost effectiveness across two contracting modalities for the same complex human service.

3. Empirical setting: Kirklees, UK

3.1. Fee-for-service (FFS) contract

Our empirical setting is Kirklees, UK as this site brings a rare opportunity to investigate two alternate contracting modalities that have been applied to the same set of social services. Before 2019, Kirklees Council (local government) commissioned housing-related support under a FFS contract through a legacy model of Floating Support services, comprising flexible, non-accommodation-linked services. The aim was to prevent homelessness by helping individuals sustain tenancies and live independently. The eligibility criteria were broad and included individuals aged 16 and over who have vulnerabilities or support needs that may impact on their ability to live independently (e.g. mental health conditions, repeat offending behaviour, substance misuse, experience of domestic violence, disabilities).

The FFS contracts were held between Kirklees Council and the non-profit social service providers, with an average value of £4 million per year. This involved 15 individual contracts managed by three local government contract managers¹, where services were delivered by nine non-profit provider organisations. Floating Support was delivered on a one-to-one basis between a vulnerable adult and a dedicated case worker, for a specified number of hours per

¹ It should be noted that these roles do not represent full-time equivalent (FTE) allocation to contract management but were undertaken alongside other administrative responsibilities.

week. The intervention duration was limited to 12 months. The payment to providers was made monthly in advance as a block fee, based on forecasted caseloads rather than verified results. Monitoring under this model was primarily process-oriented (Rosenbach et al., 2023; Rosenbach & Carter, 2020). While providers were expected to help service users achieve ‘independent living’, this outcome lacked a shared operational definition and was not tied to performance indicators. Evidence of participant progress was captured through individual support plans which were subject to periodic auditing, but outcome sustainment was not routinely measured or incentivised.

3.2. Outcomes-based contract (OBC)

From 1 September 2019 onwards, Kirklees Council transitioned to a new outcomes-based model of service delivery specified under an OBC co-funded by the UK central government (Department for Culture, Media and Sport, or DCMS), as part of the Life Chances Fund.² The OBC aimed to improve accommodation, employment, wellbeing, and financial resilience among vulnerable adults, and had comparable eligibility criteria to the services specified under the FFS contract. Under the OBC there was a single contract held between Kirklees Council and a newly established entity known as the Kirklees Better Outcomes Partnership (KBOP), the “social prime” – meaning a “prime” contractor responsible for managing a network of subcontracted service providers, constituted as an investor-owned “social” enterprise with a dual objective of social impact and financial return. This OBC defined the conditions and outcome measures that informed payments to the KBOP social prime. Outcome payments were based on a pre-defined sum for each outcome achievement measured at the level of individual service participants (such as achieving sustained accommodation for 12 months). The OBC covered a 4.5-year period (until 30 March 2025) and total maximum outcomes payments was £22.3 million. Kirklees Council received a partial contribution of 30% of the total contract value towards the outcomes payments through co-funding from the central government.

The KBOP social prime held bilateral contracts with eight non-profit provider organisations – all of whom had also delivered services under the FFS – and was responsible for program design, oversight and managing service delivery. Under these bilateral contracts, key performance indicators were tailored to each provider, and providers were paid on a full cost recovery basis according to their agreed annual budgets. All delivery costs were covered for the duration of the contract, while KBOP was remunerated through performance-based payments from Kirklees Council. Under the OBC, participants were allocated to service providers through a central referral hub, managed by KBOP with oversight from Kirklees Council. Participant data including outcomes achievements, referral assessment and support plans were stored on a central intelligence system. The outcome claims and verification process involved providers uploading the evidence for outcomes onto this system under the supervision

² The Life Chances Fund (LCF) was an outcomes fund and was delivered through 29 locally commissioned social outcomes partnerships with an outcomes-based contract. The Kirklees Better Outcomes Partnership was one of these 29 projects.

of KBOP, followed by Kirklees Council verifying the evidence and providing pre-defined outcome payments to the KBOP social prime as appropriate.

Under the OBC, there was a shift in service provision to a person-centred model of housing-related support. This approach reflected asset-based principles, tailoring support to the individual's strengths, preferences and priorities, rather than focusing on an individual's deficits.³ The aim was to offer participants greater autonomy, choice, and ownership over their support journey, while improving coordination across local services. Under this model, support was co-designed with participants where providers were granted greater flexibility in support provision. There was no prescribed length, frequency, or intensity of support, and outcomes were defined jointly by participants and reflected in the contractual rate card. Importantly, the 'outcomes' specified for payment spanned a continuum from near-term, output-type indicators, such as entry into suitable accommodation, to longer-term measures of sustained housing stability, reflecting participants' broader trajectory towards independence. These outcomes were tracked over an extended time horizon, effectively embedding a long observation window into the contract design and enabling the assessment of both immediate effects and the durability of those effects over time. Cases remained open until the outcomes each participant sought to achieve had been realised.

While the OBC contractually commenced on 1 September 2019, the transition to the new service delivery model occurred incrementally over an 18-month period. This involved the appointment of new staff, and a series of organisational reforms including structured training for support workers, new governance and operational frameworks, and the implementation of a centralised data system. In addition, the onset of the COVID-19 pandemic in 2020 significantly disrupted implementation and service delivery, with temporary suspension of the outcomes-based funding model and a shift back to FFS payments. National lockdown measures also constrained participants' access and engagement with services. For our analysis, full delivery of the outcomes-based delivery model under the OBC is therefore considered as 19 July 2021 onwards, which reflects when COVID-19 restrictions were completely lifted across England (Cabinet Office, 2021).

3.3. Comparison of key contract features

The shift from the FFS contract to the OBC marked a significant shift in the commissioning and delivery of housing-related support in Kirklees (Carter et al., 2025). A summary of the key differences in the contract features is provided in Table 1. The FFS contract involved 15 individual contracts between Kirklees Council and the non-profit provider organisations, in contrast to the single OBC held between Kirklees Council and the KBOP social prime. While the FFS was structured around forecasted caseloads, process compliance, and administratively defined risk categories, the OBC introduced a performance-based model that prioritised personalised outcomes and relational support. Unlike the FFS contract, there was greater flexibility for providers under the OBC with no prescribed length, frequency or intensity of

³ This was based on the "Person-led Transitional and Strength-based Support" (PTS) developed by the Mayday Trust (Mayday Trust, 2025).

support, and the contract period covered a longer duration. Under the FFS contract, evidence of participant progress was subject to periodic auditing; under the OBC, the outcome claims and verification process involved providers uploading evidence for outcomes onto a centralised data system which was verified by Kirklees Council.

Table 1: Comparison of key contract features

Contract features	Fee-for-service contract	Outcomes-based contract
Contract parties	Kirklees Council and social service provider organisations	Kirklees Council and KBOP social prime (investor-owned special purpose vehicle)
Contract management responsibility	Kirklees Council	Kirklees Council
Payment mechanism	Monthly advance block payment	Monthly outcomes payment (i.e. payment is contingent on achieved outcome number and type)
Key performance indicators (KPIs)	Service utilisation; throughput; independent living	Accommodation; education, training and employment; health and wellbeing; financial resilience
KPIs require sustainment of outcome achievements	No	Yes
Auditing	No pre-defined evidence requirements; spot checks of qualitative evidence (e.g. workbooks) or presence of procedural documents	Pre-defined evidence requirements; Kirklees Council audits every outcome
Contract duration	Maximum 2 years	5 years

4. Data and methods

To determine whether the OBC is cost effective compared to the services funded under the FFS contract, we undertake a cost-effectiveness analysis. The approach to collating data on costs and effects, and the cost effectiveness analysis is outlined in the following sub-sections.

4.1. Costs

In order to comprehensively and comparably capture transaction costs and implementation costs for both contracting modalities we prepared a cost activity assessment framework, as outlined in Table 2. Three categories are identified: (1) intervention / service costs, defined as the cost of service provision; (2) transaction costs, defined as government costs involved in finding/selecting vendors, negotiating contract terms, monitoring performance and ensuring delivery meets specifications (Petersen, Baekkeskov, Potoski, & Brown, 2019; Williamson, 1996); and (3) other costs which seek to capture any additional costs. These cost categories were further validated with local and central government stakeholders (Kirklees Council and Department for Culture, Media and Sport respectively) to ensure that all relevant costs had been included.

Cost data was collated from a variety of administrative sources capturing both contractual commitments and realised expenditures. Data on intervention / service costs was collated from contracts provided by Kirklees Council dating back to April 2015/16 for the FFS contracts, and from central government end of grant reconciliation forms and commissioner payments from The National Lottery Community Fund (TNLCF) for the OBC. These are administrative records generated through formal contracting and grant management processes. Kirklees contract data reflect legally binding agreements subject to internal governance and audit, while OBC data derive from financial returns submitted to the DCMS Data Portal and verified by Kirklees Council in its role as co-commissioner. TNLCF guidelines further require end-of-grant financials to be agreed across delivery partners, introducing collective validation, limiting reporting bias, and strengthening the reliability of the cost estimates.

To estimate buyer-side transaction costs, semi-structured interviews (attached in Appendix A) were undertaken with local and central government stakeholders. The interview questions were developed in accordance to the transaction cost categories in Petersen et al. (2019) and were tailored to reflect the FFS and OBC contexts. Prior to the interview, a 30-minute information session was conducted, and a briefing document containing information on transaction costs, and the interview questions was provided. This gave research participants the opportunity to review prior correspondence and retrieve archived records relevant to transaction costs that may have been accrued several years prior to the interview. A 90-minute interview was then undertaken and used to estimate transaction costs under each contracting modality, as well as to identify any additional costs not captured within the primary categories.

Table 2: Cost assessment framework

Cost category: by activity	Definition	Data collation approach	
		FFS	OBC
1 Intervention / service costs	Cost of direct service provision	Contract value	End of grant final reconciliation forms, commissioner payments from The National Lottery Community Fund
2 Transaction costs	Comprises*: 1. Ex ante: costs associated with pre-contract signing search and information, contract set-up, negotiation 2. Ex post: costs associated with governance, monitoring, evaluation and enforcement	Semi-structured interviews: Kirklees Council	Semi-structured interviews: Kirklees Council and DCMS (central government)
3 Other costs	For example, IT and overheads	Semi-structured interviews: Kirklees Council	Semi-structured interviews: Kirklees Council and DCMS (central government)

*The definition of transaction costs is based on Williamson (1996). Key activities are further broken down into categories outlined by Petersen et al. (2019).

4.2. Effects

4.2.1. Data and sample

To estimate the relative effectiveness of the different contracting modalities, management information from non-profit service providers was integrated with administrative records held by the UK Department for Work and Pensions (DWP). Provider data encompassed unique participant identifiers – including name, date of birth, National Insurance Number, and postcode – alongside supplementary information such as reason for referral and presenting needs for individuals receiving Floating Support services under both the FFS and OBC arrangements. Following cleaning and standardisation, this dataset was submitted to DWP for record linkage. Using deterministic (fuzzy) matching,⁴ corresponding administrative records were identified across three core systems: (1) Registration and Population Interaction Database (RAPID); (2) Universal Credit (UC); and (3) Single Housing Benefit Extract (SHBE). The integration of these systems enabled the construction of a longitudinal dataset capturing

⁴ This was undertaken in accordance with established DWP procedures (Department for Work and Pensions, 2025). The deterministic (fuzzy) matching involved a series of match-keys, which are pre-defined combinations of personal data variables. Records from the inbound dataset were systematically compared with mater administrative DWP files and assigned a match status based on alignment across the selected match keys.

participants’ demographic characteristics, benefit receipt histories, and employment and earnings trajectories.

The sample was defined with reference to the OBC implementation date of July 2021 (refer to Section 3.2), with a two-year follow-up period to permit assessment of longer-term outcomes. The final sample comprised 2,360 individuals; 1,639 individuals receiving Floating Support services under the FFS contract and 721 receiving KBOP’s services under the OBC (Table 3).

It should be acknowledged that inconsistent data recording practices among some legacy providers limited the completeness of the sample. Nonetheless, by grounding effect estimates in administrative rather than self-reported or provider-collected data, this approach strengthens the credibility and independence of the findings.

Table 3: Final sample

	Total number of participants
Total FFS contract	1,639
Total OBC	721
Total sample	2,360

4.2.2. Outcome measures

To construct comparable participant-level effect measures across the alternative contracting modalities, we undertook a systematic mapping exercise across administrative datasets held by DWP, conducted in consultation with data analysts in the Employment Data Lab within DWP’s Central Analysis and Science Strategy Unit. The variables identified below represent the most credible available proxies for the socioeconomic outcomes the contracts were designed to produce.⁵

Housing-related benefit receipt is derived from monthly indicators drawn from the Universal Credit (UC) and Single Housing Benefit Extract (SHBE) datasets, identifying whether individuals received either Housing Benefit (HB) or the Housing Cost Element of Universal Credit (UCHE) in a given month. Together, these instruments capture the extent to which an individual is in receipt of direct state assistance with housing costs – that is, publicly subsidised contributions to rent or accommodation expenses – and represent the primary mechanisms through which such support is delivered to low-income households in England.⁶ Our key outcome measure is sustained independence from all housing-related benefits over a

⁵ Refer to Central Analysis and Science Strategy Unit (2024) for further information on RAPID.
⁶ Housing Benefit (HB) is a means-tested benefit assisting low-income individuals with rental costs. Following the introduction of Universal Credit (UC) – a single consolidated monthly payment replacing six legacy working-age benefits – under the Welfare Reform Act 2012, HB was progressively phased out for most working-age claimants and replaced by UC’s Housing Cost Element (UCHE), which provides equivalent assistance toward rent and eligible service charges. Since the study population spans both systems, both indicators are retained to ensure comprehensive coverage of housing-related state assistance across the sample. Refer to GOV.UK (2025) for further information.

continuous 18-month period following intervention commencement, which we consider a credible proxy for the contracts' intended objective of supporting individuals to achieve and maintain stable, independent housing.

Time in employment is measured as the estimated number of days an individual was in paid work – encompassing both employment and self-employment – within a given financial year, based on their labour market status (Barclay & Gray, 2025). *Earnings* from employment are measured using taxable income, reflecting total gross pay from all employment sources within the financial year, net of pension contributions and non-taxable deductions. Together, these variables reflect the extent to which individuals move toward financial independence.

4.2.3. Sample characteristics

Table 4 outlines the sample characteristics of individuals under the FFS contract and OBC, based on administrative records. The two groups are broadly comparable across key demographic and labour market indicators and any observed variation in baseline characteristics does not indicate meaningful selection bias between the two cohorts. Differences in benefit histories, particularly relatively higher reported receipt of Employment and Support Allowance (ESA) and Jobseeker's Allowance (JSA) among FFS participants, primarily reflect the phased transition from legacy, income-based benefits to Universal Credit over the study period. As such, these differences are best understood as administrative reclassification associated with benefit migration, rather than as evidence of substantive differences in participant need or circumstances (Mackley et al., 2026).

Table 4: Sample characteristics

Variable	FFS contract		OBC	
	Mean	SD	Mean	SD
Demographics				
Age	36.4	12.1	35.9	12.0
Presence of children (%)	10.7	31.0	6.1	24.0
Male (%)	52.3	50.0	47.7	50.0
History of benefit receipt				
History of Employment & Support Allowance (%)	33.5	47.2	19.7	39.8
History of Jobseeker's Allowance (%)	9.7	29.6	2.2	14.7
History of Personal Independence Payment (%)	23.1	42.2	24.0	42.7
History of Disability Living Allowance (%)	7.2	25.9	2.4	15.2
Housing-related benefits receipt at intervention start				
Received Universal Credit	59.5	50.0	72.2	44.8
Received Universal Credit Housing Element	29.6	46.5	35.8	48.5
Received Housing Benefit	36.1	48.9	25.8	44.2
Labour market status at intervention start				
Employed	18.1	39.2	20.5	40.9
Looking for work	38.8	49.6	41.4	49.8
Earnings from paid employment at intervention start				
Annual earnings (real 2024 GBP)*	2,232	6,473	2,905	7,773
Total observations	1,639		721	

*Converted to real 2024 pounds using Consumer Price Index values from the Office for National Statistics (2025).

4.2.4. Methodology to estimate effects

Propensity Score Matching (PSM) is adopted to estimate the effect of the OBC relative to the FFS contract on participants' employment and housing outcomes. The approach constructs a comparable group of FFS participants in order to approximate what outcomes under the OBC might have looked like under the alternative contracting modality. This is achieved by balancing the distributions of observed covariates across the two groups, matching individuals based on the similarity of their predicted probabilities of receiving support services under the OBC (i.e. their propensity scores).

Two groups are identified: individuals who received services under the OBC (denoted $D_i = 1$ for individual i) and those who received services under the FFS contract (denoted $D_i = 0$). Individuals from the OBC group are matched to those in the FFS contract group on the basis of their propensity score:

$$P(x_i) = \text{Prob}(D_i = 1|x_i) \quad (0 < P(x_i) < 1)$$

where x_i is a vector of pre-intervention control variables. Based on Rosenbaum and Rubin (1983), if D_i is independent over all i and outcomes are independent of participation given x_i , the conditional independence assumption holds and differences in outcomes can be causally attributed to the intervention under either contract type.

The propensity score is estimated for each participant using a logit model. Nearest neighbour matching (NNM) is adopted to construct a comparable control group. NNM involves selecting for each OBC participant the closest eligible participants from the FFS contract. This involved a one-to-many matching approach where each case under the OBC was matched to the 100 closest cases from the FFS contract. This method reduces the variance of the estimated treatment effect by leveraging more information per treated individual, while maintaining sufficient similarity between matched pairs (Stuart, 2010). To ensure that matches were of high quality, a caliper is applied where a participant under the FFS contract is selected as a match only if their propensity score lies within the pre-defined caliper of 0.01 and is the closest possible match to the participant under the OBC (Caliendo & Kopeinig, 2008). Matching is conducted with replacement so participants under the FFS contract could serve as a match for more than one participant under the OBC, enabling improved match quality (Stuart, 2010). The pre-intervention control variables used to match participants includes a range of demographic and socioeconomic indicators including age, gender, prior benefit receipt and employment trajectories two years preceding service participation (e.g. Hevenstone, Fraser, Hobi, Przepiorka, & Geuke, 2023). Labour market history may act as a proxy for unobserved attributes such as intrinsic motivation, which may influence the likelihood of participation and subsequent employment outcomes (Caliendo & Kopeinig, 2008). The full list of matching variables is provided in Appendix B (Table B1). The matching process is implemented using the “psmatch2” package in Stata (Leuven & Sianesi, 2003).

We assess balance between the matched FFS contract and OBC groups to ensure that participants are sufficiently comparable. Our balance statistics indicate a high degree of covariate balance between both groups (refer to Appendix C, Table C1 for further details). We also verify the common support condition, ensuring adequate overlap in propensity scores between treated and control units. We undertake robustness checks on our model by varying the choice of matching estimators. We find that our model is robust and provides the optimal trade-off between covariate balance and sample retention (refer to Appendix D for further details).

4.3. Cost effectiveness analysis

The cost effectiveness analysis is undertaken across a nine-year time period, from April 2015/16 to April 2024/25. This time period was selected based on data availability, key milestones and implementation dates. The base year for the analysis is 2024, so costs are converted to real 2024 pounds using Consumer Price Index values from the Office for National Statistics (2025). The key metric estimated is the incremental cost effectiveness ratio (ICER), which measures the incremental cost per unit of outcome for the OBC compared to the FFS contract.

5. Results

5.1. Costs

5.1.1. Annual costs: fee-for-service contract

Annual costs under the FFS contract are outlined in Table 5. Given that the OBC commenced on 1 September 2019, costs in the period April 2019-20 are not incurred for the full financial year.

Based on semi-structured interviews, the following staffing involved in transaction cost activities are estimated:

- 1 Full Time Equivalent (FTE) Commissioning Manager (grade 14): approximately 50% of time
- 1 FTE Contract Manager (grade 13): approximately 60% of time
- 2.46 FTE Contract Officers (grade 11): approximately 60% of time.

Based on these staffing levels across the pre-KBOP period, transaction costs are estimated using the mid-point of salary estimates in Kirklees Council (2025), amounting to £90,015 – £99,512 during a full financial year.

Overall, total costs of services under the FFS contract reduce over time from £4.4m to £3.4m (excluding the partial financial year of April 2019-20), driven by the merging of contracts and managed reduction of contract values in response to austerity (Rosenbach & Carter, 2020).

Table 5: Annual costs – FFS contract

Financial year	Apr 15/16	Apr 16/17	Apr 17/18	Apr 18/19	Apr 19/20
<i>Intervention/service costs</i>					
Community Links	301,173	301,173	301,173	286,109	120,385
Connect Housing	379,965	379,965	355,476	280,648	118,087
Foundation Housing	820,746	820,746	718,434	615,560	259,006
Fusion	1,336,192	1,336,192	1,336,192	1,002,144	421,667
Home Group	425,000	344,100	344,100	344,100	144,785
Horton	556,662	556,662	556,662	417,497	175,668
Making Space	35,007	35,007	35,007	35,007	14,730
Richmond Fellowship	271,322	271,322	237,500	203,491	85,622
Pennine Domestic Violence Group	138,368	138,368	112,216	91,950	43,676
Total	4,264,434	4,183,534	3,996,759	3,276,505	1,383,626
<i>Transaction costs</i>					
Kirklees Council	90,015	93,076	96,240	99,512	41,444
TOTAL COSTS	4,354,449	4,276,610	4,093,000	3,376,018	1,425,070

5.1.2. Annual costs: outcomes-based contract

Annual costs under the OBC are outlined in Table 6. While the OBC commenced in September 2019, the OBC was part of the UK central government's Life Chances Fund (refer to Section 3), so set-up occurred in the four years preceding this. Overall, total annual costs under the OBC are around £6m once fully implemented.

In comparison to the FFS contract, there are now commissioner payments from both local and central government, with local government paying for the majority of total outcomes (70%) compared to central government (30%). Staffing levels associated with transaction cost activities changed when the OBC was implemented and is estimated to comprise:

- 1 FTE Senior Contracts Manager (grade 15), approximately 10% of time;
- 1 FTE Contract Manager (grade 13), approximately 50% of time; and
- 1.67 FTE Contract Officers (grade 11), approximately 80% of time.

These transaction costs are estimated to be £78,389 – £89,606 during a full financial year. The Kirklees Council also incurred set-up and business case development costs for the OBC prior to KBOP implementation totalling £50,196, which are understood as additional ex ante transaction costs, beyond the staffing costs.

Under the OBC there are also three new cost categories. This includes IT costs for the central intelligence system collating participant information, comprising an initial licence fee and a recurrent annual licence fee; DCMS transaction costs; and evaluation costs incorporating a process and impact evaluation of KBOP. While evaluation expenditures would not ordinarily be classified as transaction costs under prevailing frameworks (e.g., Petersen et al.), their inclusion here reflects the design of the Life Chances Fund, under which such evaluative activities are contractually mandated and therefore constitute a necessary cost of transacting under the OBC modality.

Table 6: Annual costs – OBC

Financial year	Apr 15/16	Apr 16/17	Apr 17/18	Apr 18/19	Apr 19/20	Apr 20/21	Apr 21/22	Apr 22/23	Apr 23/24	Apr 24/25
<i>Intervention/service costs</i>										
Commissioner payments	-	-	-	-	397,500	4,537,508	5,430,600	6,282,674	5,905,515	-
IT costs	-	-	-	-	40,700	12,246	16,981	17,660	18,367	19,101
Total intervention/service costs	-	-	-	-	438,200	4,549,754	5,447,581	6,300,334	5,923,882	19,101
<i>Transaction costs</i>										
Kirklees Council	-	-	-	50,196	45,066	78,389	81,054	83,810	86,660	89,606
DCMS* & evaluation costs	3,415	7,020	20,002	68,170	46,368	71,728	65,202	70,824	77,664	236,600
Total transaction costs	3,415	7,020	20,002	118,366	91,434	150,117	146,256	154,634	164,323	326,206
TOTAL COSTS	3,415	7,020	20,002	118,366	529,634	4,699,871	5,593,837	6,454,968	6,088,205	345,307

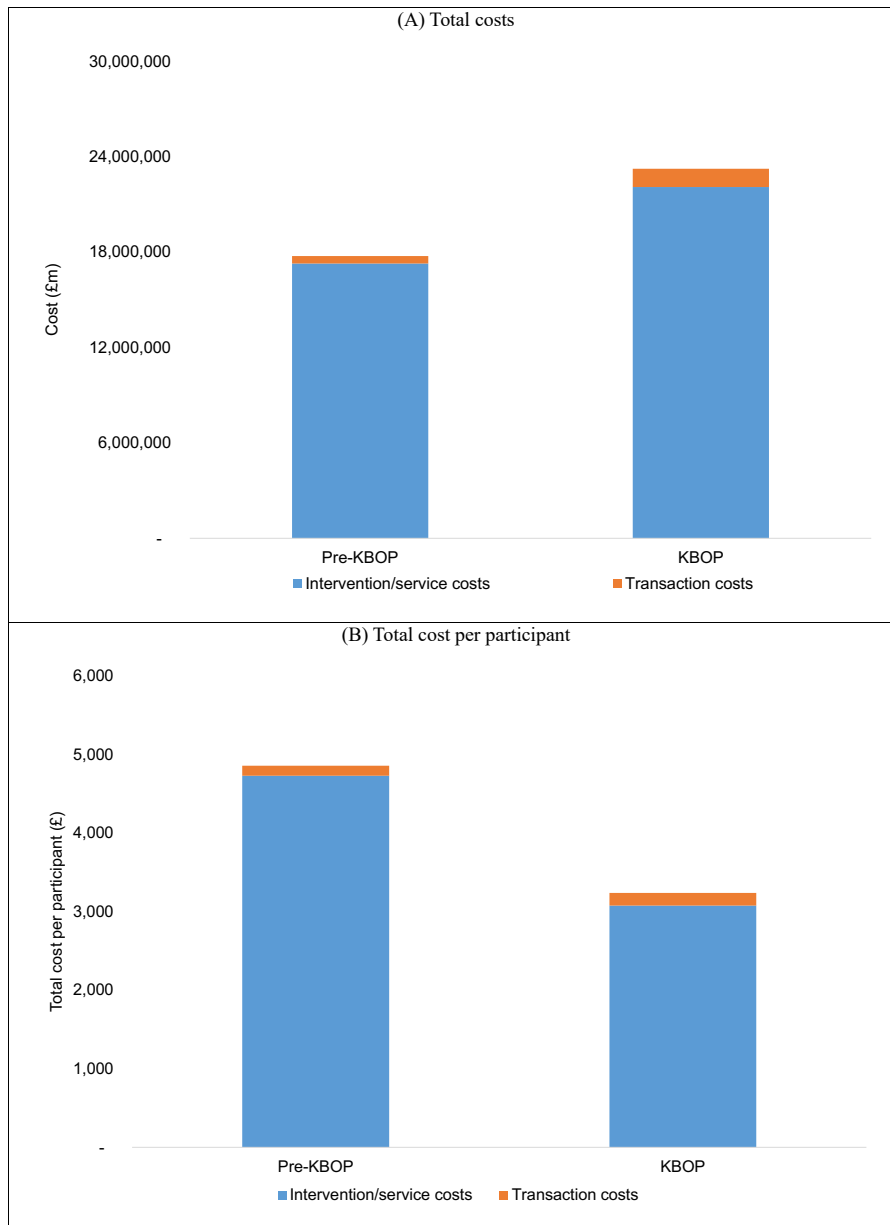
*UK Department for Culture, Media and Sport (central government).

5.1.3. Total costs

Total costs (real 2024£) under the OBC are higher than under the FFS contract (£23.3m compared to £17.8m), as illustrated in Figure 1 (Panel A). When costs are considered as a percentage of total cost, intervention / service costs are higher under the FFS contract (97%) compared to the OBC (95%), while transaction costs were lower under the FFS contract (3%) compared to the OBC (5%).

However, while the OBC had higher total costs, this modality supported a higher number of participants (n=7,185) compared to the FFS contract (n=3,657). To account for this difference, costs are standardised at the per participant level. This results in a lower total cost per participant under the OBC compared to the FFS contract (£3,236 compared to £4,856) as illustrated in Figure 1 (Panel B).

Figure 1: Total costs (real 2024 £)



5.2. Effects

Estimated impacts for housing and employment outcomes are outlined in Table 7. OBC participants are significantly more likely to sustain independence from housing-related benefits for 18 consecutive months post intervention commencement compared to participants under the FFS contract (by 6.20 percentage points). This indicates that the OBC contributed to participants' broader independence from housing support benefits.

OBC participants are also significantly more likely to record an employment spell within the first two years following programme commencement (by 6.15 percentage points). While days in employment and average earnings increased post commencement for OBC participants relative to the FFS contract (by 5 days and £539 respectively), these results were not statistically significant.

Table 7: Estimated impacts

Outcome variable	(1) OBC	(2) FFS	(3) Treatment effect	(4) 95% CI
Housing				
Sustained independence from housing-related benefits for 18 months post intervention start (%)	26.91	20.76	6.15***	1.38, 10.93
Employment				
Employment in the first two years post intervention (%)	38.20	32.00	6.20**	0.90, 11.60
Days in employment	78	73	5	-9, 20
Earnings (real 2024 £)	3,256.56	2,717.48	539.09	-392.16, 1,470.33

*** p<0.01, ** p<0.05, * p<0.1

Notes: (3) The treatment effect is calculated as the difference between OBC (1) and FFS (2). (4) 95% CI shows the 95% confidence interval for the treatment effect in (3).

5.3. Cost effectiveness

Table 8 outlines the incremental cost effectiveness ratio (ICER) for each of the outcome variables. The OBC is generally less costly and more effective compared to the FFS contract. The ICER is -£12 per additional person sustaining independence from housing-related benefits; and -£11 per additional person entering employment in the first two years post intervention commencement.

Table 8: Incremental cost effectiveness ratio (ICER)

Outcome variable	(1) ICER
Housing	
Sustained independence from housing-related benefits for 18 months post intervention start	-11.61***

a supprimé: The insignificant impact estimates for days in employment and earnings resulted in an ICER of -£324 and -£3.00 respectively.

Outcome variable	(1) ICER
Employment	
Employment in the first two years post intervention	-11.42**
Days in employment	-323.91
Earnings (real 2024 £)	-3.00

*** p<0.01, ** p<0.05, * p<0.1

Notes: The ICER (1) is based on incremental cost per participant estimated in Section 5.1 and the effect estimates in Section 5.2.

6. Discussion and conclusions

Overall, we find that the OBC is less costly per person, and more effective than the FFS contract. The OBC had a lower cost per participant than the FFS contract (£3,236 compared to £4,856). The OBC also resulted in improved outcomes relative to the FFS contract, with a significant increase in participants sustaining independence from housing-related benefits for 18 months (6.15 percentage points) and being employed in the first two years (6.20 percentage points). Participants under the OBC also had greater days in employment and earnings, but these results were not statistically significant. These results indicate that the OBC represents value for money in comparison to the FFS contract. The incremental cost effectiveness ratio (ICER) is -£12 per additional person sustaining independence from housing-related benefits; and -£11 per additional person entering employment in the first two years post intervention commencement. While the ICER for days in employment and earnings indicate greater cost effectiveness under the OBC, these results were not statistically significant.

The relative cost-effectiveness of the OBC appears to stem from several reinforcing mechanisms, which are further corroborated by qualitative analyses of the KBOP model (Carter et al., 2025; Rosenbach et al., 2023, 2025). First, formalising the ends of delivery created a shared focal point around which principals and agents could coordinate effort, while granting providers greater flexibility to tailor services to meet individual user needs (Carter et al., 2025). Unlike the FFS contract, the OBC did not rigidly specify the length, frequency, or intensity of support, instead offering space for a person-centred support model that enabled outcomes achievement (Carter et al., 2025; Patouna et al., 2025). Second, the KBOP social prime played a core role in coordinating and overseeing service delivery, mediating between service providers and Kirklees Council, and ensuring strong performance oversight through a centralised data system – proactively monitoring providers’ performance, applying strengthened quality standards, and supporting underperforming providers. There was also greater collaboration among service providers, with the social prime facilitating trust-based relationships, transparent governance, and structured opportunities for knowledge exchange and collective learning (Carter et al., 2026a; Rosenbach et al., 2025). Although total programme costs were higher under the OBC, reflecting both greater transaction costs and a larger overall delivery spend, these costs were offset by greater efficiency in terms of participants served, yielding better value for money.

The present study has limitations that warrant acknowledgement. First, the ICER is based on aggregate programme costs, and we have not attempted to estimate the proportion of

intervention/service costs attributable to each outcome, therefore the reported ICERs are likely conservative estimates. Second, data recording inconsistencies among some legacy providers mean the FFS sample does not capture the full participant population, introducing selection effects that cannot be fully adjudicated. Third, the PSM identification strategy balances observed pre-intervention covariates but does not account for differences in macroeconomic and labour market conditions across the two contracting periods. Given that the OBC cohort was observed during a period of post-pandemic labour market recovery, some portion of the estimated employment effects may reflect broader economic trends rather than the contracting modality per se; data limitations precluded further investigation of this temporal confound. Further research could consider the use of alternative identification strategies to control for these differences.

In the situation of complex homelessness prevention services in northern England, we find that an outcomes-based contracting arrangement offers superior value for money as a contracting strategy when compared to more conventional fee-for-service arrangements. The service efficiencies and elevated effectiveness seen in this OBC more than offset the higher costs. This has implications for public managers who may otherwise be daunted by the allegedly 'high' transaction costs associated with developing outcomes specifications, embedding these within contracts and giving service providers the latitude to innovate in implementation whilst holding firm to accountability for outcomes that align with public value.

REFERENCES

- Albertson, K., Fox, C., O'leary, C., Painter, G., Bailey, K., & Labarbera, J. (2018). Outcome-based commissioning: Theoretical underpinnings. In K. Albertson, C. Fox, C. O'Leary, & G. Painter (Eds.), *Payment by Results and Social Impact Bonds: Outcome-Based Payment Systems in the UK and US* (p. 0). Policy Press.
<https://doi.org/10.1332/policypress/9781447340706.003.0002>
- Alchian, A. A., & Demsetz, H. (1972). Production, Information Costs, and Economic Organization. *The American Economic Review*, 62(5), 777–795.
<https://www.jstor.org/stable/1815199>
- Amirkhanyan, A. A., Kim, H. J., & Lambright, K. T. (2010). Do Relationships Matter?: Assessing the Association Between Relationship Design and Contractor Performance. *Public Performance & Management Review*, 34(2), 189–220.
<https://doi.org/10.2753/PMR1530-9576340203>
- Arena, M., Bengo, I., Calderini, M., & Chiodo, V. (2016). Social Impact Bonds: Blockbuster or Flash in a Pan? *International Journal of Public Administration*, 39(12), 927–939.
<https://doi.org/10.1080/01900692.2015.1057852>
- Arrow, K. (1984). The economics of the agency. In *Principals and agents: The structure of business*. Harvard Business School Press.
<https://cir.nii.ac.jp/crid/1970304959832378934>
- Baker, G. (1992). Incentive Contracts and Performance Measurement. *Journal of Political Economy*, 100(3), 598–614. <https://www.jstor.org/stable/2138733>
- Baker, G. (2002). Distortion and Risk in Optimal Incentive Contracts. *The Journal of Human Resources*, 37(4), 728. <https://doi.org/10.2307/3069615>

- Baker, G., Gibbons, R., & Murphy, K. J. (1994). Subjective Performance Measures in Optimal Incentive Contracts. *The Quarterly Journal of Economics*, 109(4), 1125–1156. <https://doi.org/10.2307/2118358>
- Barclay, L., & Gray, H. (2025). *Employment Data Lab: Methodology report*. <https://www.gov.uk/government/publications/employment-data-lab-information-and-guidance/employment-data-lab-methodology-report>
- Battye, F. (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>
- Bel, G., Espaillat, P., & Esteve, M. (2025a). The performance of performance-based contracting in public outsourcing: A meta-regression analysis. *Journal of Public Administration Research and Theory*, muaf037. <https://doi.org/10.1093/jopart/muaf037>
- Bertelli, A. M., & Smith, C. R. (2010). Relational Contracting and Network Management. *Journal of Public Administration Research and Theory*, 20(suppl_1), i21–i40. <https://doi.org/10.1093/jopart/mup033>
- Bevan, G., & Hood, C. (2006). WHAT'S MEASURED IS WHAT MATTERS: TARGETS AND GAMING IN THE ENGLISH PUBLIC HEALTH CARE SYSTEM. *Public Administration*, 84(3), 517–538. <https://doi.org/10.1111/j.1467-9299.2006.00600.x>
- Bovaird, T., & Davies, R. (2011). Chapter 7 Outcome-Based Service Commissioning and Delivery: Does it make a Difference? In S. Groeneveld & S. Van De Walle (Eds.), *Research in Public Policy Analysis and Management* (Vol. 21, pp. 93–114). Emerald Group Publishing Limited. [https://doi.org/10.1108/S0732-1317\(2011\)0000021011](https://doi.org/10.1108/S0732-1317(2011)0000021011)

- Brown, T. L., & Potoski, M. (2003). Transaction Costs and Institutional Explanations for Government Service Production Decisions. *Journal of Public Administration Research and Theory*, 13(4), 441–468. <https://doi.org/10.1093/jopart/mug030>
- Brown, T. L., Potoski, M., & Slyke, D. M. V. (2013). *Complex Contracting*. Cambridge University Press.
- Brown, T. L., Potoski, M., & Van Slyke, D. M. (2006). Managing Public Service Contracts: Aligning Values, Institutions, and Markets. *Public Administration Review*, 66(3), 323–331. <https://doi.org/10.1111/j.1540-6210.2006.00590.x>
- Brown, T. L., Potoski, M., & Van Slyke, D. M. (2018). Complex Contracting: Management Challenges and Solutions. *Public Administration Review*, 78(5), 739–747. <https://doi.org/10.1111/puar.12959>
- Brunjes, B. M. (2020). Competition and Federal Contractor Performance. *Journal of Public Administration Research and Theory*, 30(2), 202–219. <https://doi.org/10.1093/jopart/muz027>
- Cabinet Office. (2019). *The Outsourcing Playbook: Central Government Guidance on Outsourcing Decisions and Contracting*. [Government Commercial Function]. Cabinet Office.
- Cabinet Office. (2021). *Guidance: Moving to step 4 of the roadmap*. Cabinet Office. GOV.UK. <https://www.gov.uk/government/publications/covid-19-response-summer-2021-roadmap/moving-to-step-4-of-the-roadmap>
- Carter, E., FitzGerald, C., Dixon, R., Economy, C., Hameed, T., & Airoidi, M. (2018). *Building the tools for public services to secure better outcomes: Collaboration, Prevention,*

Innovation. Government Outcomes Lab, University of Oxford, Blavatnik School of Government.

https://scholar.google.com/scholar_lookup?title=Building+the+Tools+for+Public+Services+to+Secure&author=Carter+E.&author=FitzGerald+C.&author=Dixon+R.&publication+year=2018

Carter, E., Gibson, M., Heinrich, C. J., & Malatesta, D. (2026a). A tyranny of metrics vs. a tool for learning: Exploring performance information use on the spectrum of formal to relational in outcomes-based contracts. *Public Management Review*, 0(0), 1–26. <https://doi.org/10.1080/14719037.2026.2630985>

Carter, E., Gibson, M., Heinrich, C. J., & Malatesta, D. (2026b). A tyranny of metrics vs. a tool for learning: Exploring performance information use on the spectrum of formal to relational in outcomes-based contracts. *Public Management Review*, 1–26. <https://doi.org/10.1080/14719037.2026.2630985>

Carter, E., Rosenbach, F., Domingos, F., & Van Lier, F.-A. (2025). Contracting ‘person-centred’ working by results: Street-level managers and frontline experiences in an outcomes-based contract. *Public Management Review*, 27(8), 1893–1911. <https://doi.org/10.1080/14719037.2024.2342398>

Carter, E., & Whitworth, A. (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>

Coase, R. H. (1937). The Nature of the Firm. *Economica*, 4(16), 386–405. <https://doi.org/10.1111/j.1468-0335.1937.tb00002.x>

- Considine, M., O'Sullivan, S., McGann, M., & Nguyen, P. (2020). Contracting personalization by results: Comparing marketization reforms in the UK and Australia. *Public Administration*, 98(4), 873–890. <https://doi.org/10.1111/padm.12662>
- Cox, B. R. (2011). *Financing Homelessness Prevention Programs with Social Impact Bonds*. Review of Banking and Financial Law. <https://www.bu.edu/rbfl/files/2013/09/FinancingHomelessnessPreventionPrograms.pdf>
- Cuñat, V., & Guadalupe, M. (2005). How Does Product Market Competition Shape Incentive Contracts? *Journal of the European Economic Association*, 3(5), 1058–1082. <https://doi.org/10.1162/1542476054729437>
- De Schepper, S., Haezendonck, E., & Dooms, M. (2015). Understanding pre-contractual transaction costs for Public–Private Partnership infrastructure projects. *International Journal of Project Management*, 33(4), 932–946. <https://doi.org/10.1016/j.ijproman.2014.10.015>
- DeHoog, R. H. (1990). Competition, Negotiation, or Cooperation: Three Models for Service Contracting. *Administration & Society*, 22(3), 317–340. <https://doi.org/10.1177/009539979002200303>
- Dias, J. J., & Maynard-Moody, S. (2006). For-Profit Welfare: Contracts, Conflicts, and the Performance Paradox. *Journal of Public Administration Research and Theory*, 17(2), 189–211. <https://doi.org/10.1093/jopart/mul002>
- Domberger, S., & Jensen, P. (1997). Contracting out by the public sector: Theory, evidence, prospects. *Oxford Review of Economic Policy*, 13(4), 67–78. <https://doi.org/10.1093/oxrep/13.4.67>

- Dudkin, G., & Väilä, T. (2006). Transaction Costs in Public-Private Partnerships: A First Look at the Evidence. *Competition and Regulation in Network Industries*, 1(2), 307–330. <https://doi.org/10.1177/178359170600100209>
- Eisenhardt, K. M. (1989). Agency Theory: An Assessment and Review. *The Academy of Management Review*, 14(1), 57. <https://doi.org/10.2307/258191>
- Esteve, M., & Reyes-Gonzalez, J. A. (2020). Key mechanisms of the accountability process in public-private partnerships. *Revista de Contabilidad*, 23(2), 210–223. <https://doi.org/10.6018/rcsar.369621>
- Favero, N., Meier, K. J., & O'Toole, L. J. (2016). Goals, Trust, Participation, and Feedback: Linking Internal Management With Performance Outcomes. *Journal of Public Administration Research and Theory*, 26(2), 327–343. <https://doi.org/10.1093/jopart/muu044>
- Fernandez, S. (2007). What Works Best When Contracting for Services? An Analysis of Contracting Performance at the Local Level in the US. *Public Administration*, 85(4), 1119–1141. <https://doi.org/10.1111/j.1467-9299.2007.00688.x>
- Fernandez, S. (2009). Understanding Contracting Performance: An Empirical Analysis. *Administration & Society*, 41(1), 67–100. <https://doi.org/10.1177/0095399708330257>
- Finkenstadt, D., & Cummins, T. (2026). *Outcome-Based Contracting in U.S. Government* | IBM Center for The Business of Government. Commerce & Contract Management Institute. <https://doi.org/https://www.businessofgovernment.org/report/outcome-based-contracting-us-government>

- FitzGerald, C., Carter, E., Dixon, R., & Airoidi, M. (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., Fraser, A., & Kimmitt, J. (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., Hameed, T., Rosenbach, F., James Ruairi Macdonald, Outes Velarde, J., & Dixon, R. (2021). *An Introduction to Life Chances Fund projects and their early adaptations to COVID-19: Life Chances Fund Evaluation Interim Report*. UK Department for Digital, Culture, Media & Sport.
- FitzGerald, C., Moynan, E., Lavelle, C., O'Neill, C., Robinson, K., Boland, P., Meskell, P., & Galvin, R. (2024). Exploring the Impact of COVID-19 on Home Care Workers: A Qualitative Study. *Gerontology and Geriatric Medicine*, 10, 23337214231222114. <https://doi.org/10.1177/23337214231222114>
- FitzGerald, C., Tan, S., Carter, E., & Airoidi, M. (2023). Contractual acrobatics: A configurational analysis of outcome specifications and payment in outcome-based contracts. *Public Management Review*, 25(9), 1796–1814. <https://doi.org/10.1080/14719037.2023.2244501>
- Frydlinger, D., & Hart, O. (2024). Overcoming contractual incompleteness: The role of guiding principles. *The Journal of Law, Economics, and Organization*, 40(3), 625–647. <https://doi.org/10.1093/jleo/ewac027>

- Gerrish, E. (2016). The Impact of Performance Management on Performance in Public Organizations: A Meta-Analysis. *Public Administration Review*, 76(1), 48–66. <https://doi.org/10.1111/puar.12433>
- Girth, A. M., Hefetz, A., Johnston, J. M., & Warner, M. E. (2012). Outsourcing Public Service Delivery: Management Responses in Noncompetitive Markets. *Public Administration Review*, 72(6), 887–900. <https://doi.org/10.1111/j.1540-6210.2012.02596.x>
- Grossman, S. J., & Hart, O. D. (1986). The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy*, 94(4), 691–719. <https://doi.org/10.1086/261404>
- Gustafsson-Wright, E., & Osborne, S. (2020). *Do the benefits outweigh the costs of impact bonds?* Global Economy and Development Program at Brookings.
- Hart, O., & Moore, J. (1988). Incomplete Contracts and Renegotiation. *Econometrica*, 56(4), 755–785. <https://doi.org/10.2307/1912698>
- Hart, O., Shleifer, A., & Vishny, R. W. (1997). The Proper Scope of Government: Theory and an Application to Prisons. *The Quarterly Journal of Economics*, 112(4), 1127–1161. <https://doi.org/10.1162/003355300555448>
- Hasenfeld, Y. (1983). *Human service organizations*. Prentice-Hall.
- Hefetz, A., & Warner, M. E. (2012). Contracting or Public Delivery? The Importance of Service, Market, and Management Characteristics. *Journal of Public Administration Research and Theory*, 22(2), 289–317. <https://doi.org/10.1093/jopart/mur006>

- Heinrich, C. J. (2007). False or fitting recognition? The use of high performance bonuses in motivating organizational achievements. *Journal of Policy Analysis and Management*, 26(2), 281–304. <https://doi.org/10.1002/pam.20244>
- Heinrich, C. J., & Choi, Y. (2007). Performance-Based Contracting in Social Welfare Programs. *The American Review of Public Administration*, 37(4), 409–435. <https://doi.org/10.1177/0275074006297553>
- Hevenstone, D., Fraser, A., Hobi, L., Przepiorka, W., & Geuke, G. G. M. (2023). The impact of social impact bond financing. *Public Administration Review*, 83(4), 930–946. <https://doi.org/10.1111/puar.13631>
- Hodgkin, D., Garnick, D. W., Horgan, C. M., Busch, A. B., Stewart, M. T., & Reif, S. (2020). Is it feasible to pay specialty substance use disorder treatment programs based on patient outcomes? *Drug and Alcohol Dependence*, 206, 107735. <https://doi.org/10.1016/j.drugalcdep.2019.107735>
- Jensen, M. C., & Meckling, W. H. (1979). Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure. In K. Brunner (Ed.), *Economics Social Institutions* (Vol. 1, pp. 163–231). Springer Netherlands. https://doi.org/10.1007/978-94-009-9257-3_8
- Joy, M., & Shields, J. (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>
- Kettl, D. (1993). *Sharing Power: Public Governance and Private Markets*.

- Kettl, D. (2002). Managing indirect government. In *The tools of government: A guide to the new governance*. Oxford University Press.
- Koning, P., & Heinrich, C. J. (2013). Cream-Skimming, Parking and Other Intended and Unintended Effects of High-Powered, Performance-Based Contracts: Cream-Skimming, Parking and Other Intended and Unintended Effects. *Journal of Policy Analysis and Management*, 32(3), 461–483. <https://doi.org/10.1002/pam.21695>
- Larsen, F., & Wright, S. (2014). Interpreting the marketization of employment services in Great Britain and Denmark. *Journal of European Social Policy*, 24(5), 455–469. <https://doi.org/10.1177/0958928714543903>
- Lazzarini, S. G. (2022). *The Right Privatization: Why Private Firms in Public Initiatives Need Capable Governments*. Cambridge University Press. <https://doi.org/10.1017/9781009024167>
- Lazzarini, S. G., Cabral, S., Firpo, S., & Teodorovicz, T. (2022). Why Are Counterfactual Assessment Methods Not Widespread in Outcome-Based Contracts? A Formal Model Approach. *Journal of Public Administration Research and Theory*, 32(3), 509–523. <https://doi.org/10.1093/jopart/muab052>
- Levin, J., & Tadelis, S. (2010). CONTRACTING FOR GOVERNMENT SERVICES: THEORY AND EVIDENCE FROM U.S. CITIES*. *The Journal of Industrial Economics*, 58(3), 507–541. <https://doi.org/10.1111/j.1467-6451.2010.00430.x>
- Lipsky, M. (1980). *Street-level bureaucracy: Dilemmas of the individual in public services*. Russell Sage.

- Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Prentice Hall.
- Lowe, T., & Wilson, R. (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>
- Lu, J. (2016). The Performance of Performance-Based Contracting in Human Services: A Quasi-Experiment. *Journal of Public Administration Research and Theory*, 26(2), 277–293. <https://doi.org/10.1093/jopart/muv002>
- Mackley, A., Hobson, F., Harker, R., & Kennedy, S. (2026). *Managed migration: Completing Universal Credit rollout* (No. 9984). House of Commons Library, UK Parliament. <https://commonslibrary.parliament.uk/research-briefings/cbp-9984/>
- Malatesta, D., & Smith, C. (2014). Designing Contracts for Complex Services. *Public Administration*, 92(3), 531–548. <https://doi.org/10.1111/padm.12004>
- Malatesta, D., & Smith, C. R. (2011). Resource Dependence, Alternative Supply Sources, and the Design of Formal Contracts. *Public Administration Review*, 71(4), 608–617. <https://doi.org/10.1111/j.1540-6210.2011.02392.x>
- Marlin, J. T. (1984). *Contracting Municipal Services: A Guide for Purchase from the Private Sector*. Wiley.
- Miller, G. J., & Whitford, A. B. (2006). The Principal's Moral Hazard: Constraints on the Use of Incentives in Hierarchy. *Journal of Public Administration Research and Theory*, 17(2), 213–233. <https://doi.org/10.1093/jopart/mul004>

- OECD. (2025). Size of public procurement: Government at a Glance 2025. In *Government at a Glance 2025*. OECD Publishing. <https://doi.org/https://doi.org/10.1787/0efd0bcd-en>
- Office of Management and Budget. (2013). *Next Steps in the Evidence and Innovation Agenda* (M-13–17). Office of Management and Budget, Executive Office of the President. https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2013/m-13-17.pdf
- Olson, H., Painter, G., Albertson, K., Fox, C., & O’Leary, C. (2024). Are Social Impact Bonds an Innovation in Finance or Do They Help Finance Social Innovation? *Journal of Social Policy*, 53(2), 407–431. <https://doi.org/10.1017/S0047279422000356>
- Pandey, S., Cordes, J. J., Pandey, S. K., & Winfrey, W. F. (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>
- Patouna, M., De Gruyter, E., & Carter, E. (2025). *Labour Market Evaluation Pilot Fund: Kirklees Better Outcomes Partnership: Understanding the effects of person-centred service reforms in housing support*. Government Outcomes Lab, Blavatnik School of Government, University of Oxford.
- Perry, J. L., Hondeghem, A., & Wise, L. R. (2010). Revisiting the Motivational Bases of Public Service: Twenty Years of Research and an Agenda for the Future. *Public Administration Review*, 70(5), 681–690. <https://doi.org/10.1111/j.1540-6210.2010.02196.x>
- Petersen, O. H. (2019). EVALUATING THE COSTS, QUALITY, AND VALUE FOR MONEY OF INFRASTRUCTURE PUBLIC-PRIVATE PARTNERSHIPS: A

SYSTEMATIC LITERATURE REVIEW. *Annals of Public and Cooperative Economics*, 90(2), 227–244. <https://doi.org/10.1111/apce.12243>

Petersen, O. H., Baekkeskov, E., Potoski, M., & Brown, T. L. (2019). Measuring and Managing Ex Ante Transaction Costs in Public Sector Contracting. *Public Administration Review*, 79(5), 641–650. <https://doi.org/10.1111/puar.13048>

Petersen, O. H., Houlberg, K., & Christensen, L. R. (2015). Contracting Out Local Services: A Tale of Technical and Social Services. *Public Administration Review*, 75(4), 560–570. <https://doi.org/10.1111/puar.12367>

Piatak, J., & Jensen, C. (2024). Public values and sector service delivery preferences: Public preferences on contracting from simple to complex human services. *Public Administration Review*, 84(5), 948–965. <https://doi.org/10.1111/puar.13794>

Potoski, M., Lund-Sørensen, B., & Petersen, O. H. (2025). Measuring Transaction Costs in Public Sector Contracting Through Machine Learning and Contract Text. *Public Administration Review*, puar.13947. <https://doi.org/10.1111/puar.13947>

Radin, B. A. (2006). *Challenging the performance movement: Accountability, complexity, and democratic values*.

Romzek, B. S., & Johnston, J. M. (2005). State Social Services Contracting: Exploring the Determinants of Effective Contract Accountability. *Public Administration Review*, 65(4), 436–449. <https://doi.org/10.1111/j.1540-6210.2005.00470.x>

Rosenbach, F., & Carter, E. (2020). *Kirklees Integrated Support Service And Better Outcomes Partnership: The first report from a longitudinal evaluation of a Life Chances Fund Impact Bond*. Government Outcomes Lab, Blavatnik School of Government,

University of Oxford.

https://assets.publishing.service.gov.uk/media/6294e1e0e90e0703a07c04e6/Kirklees_Integrated_Support_Service_and_Better_Outcomes_Partnership_Report.pdf

Rosenbach, F., van Lier, F., Domingos, F., & Carter, E. (2023). *The Kirklees Better Outcomes Partnership: The second report of a longitudinal evaluation of a Life Chances Fund Impact Bond*. Government Outcomes Lab, Blavatnik School of Government, University of Oxford. <https://www.gov.uk/government/publications/life-chances-fund-second-stage-evaluation-report-kirklees/life-chances-fund-evaluation-second-report-on-kirklees-better-outcomes-partnership>

Rosenbach, F., Van Lier, F.-A., Patouna, M., & Carter, E. (2025). *The Kirklees Better Outcomes Partnership: The final report of a longitudinal evaluation of a Life Chances Fund outcomes partnership*. Government Outcomes Lab.

Sappington, D. E. M. (1991). Incentives in Principal-Agent Relationships. *Journal of Economic Perspectives*, 5(2), 45–66. <https://doi.org/10.1257/jep.5.2.45>

Savas, E. S. (1987). *Privatization: The Key to Better Government*. Chatham House.

Savas, E. S. (2000). *Privatization and Public-Private Partnerships*. Chatham House.

Schepker, D. J., Oh, W.-Y., Martynov, A., & Poppo, L. (2014). The Many Futures of Contracts: Moving Beyond Structure and Safeguarding to Coordination and Adaptation. *Journal of Management*, 40(1), 193–225. <https://doi.org/10.1177/0149206313491289>

Sclar, E. (2000). *You don't always get what you pay for: The economics of privatization*. Ithaca, NY: Cornell University Press.

- Segal, I. (1999). Complexity and Renegotiation: A Foundation for Incomplete Contracts. *Review of Economic Studies*, 66(1), 57–82. <https://doi.org/10.1111/1467-937X.00078>
- Shiva, M., FitzGerald, C., Carter, E., & Airolidi, M. (2024). Beyond “make” or “buy”: Evaluating value-for-money in public service delivery. *Annals of Public and Cooperative Economics*, 95(4), 1171–1196. <https://doi.org/10.1111/apce.12468>
- Simon, H. A. (1957). *Models of Man: Social and Rational; Mathematical Essays on Rational Human Behavior in a Social Setting*. Wiley.
- Sklivas, S. D. (1987). The Strategic Choice of Managerial Incentives. *The RAND Journal of Economics*, 18(3), 452. <https://doi.org/10.2307/2555609>
- Thiel, S. V., & Leeuw, F. L. (2002). The Performance Paradox in the Public Sector. *Public Performance & Management Review*, 25(3), 267. <https://doi.org/10.2307/3381236>
- Tomkinson, E. (2016). Outcome-based contracting for human services. *Evidence Base*, 2016. <https://doi.org/10.21307/eb-2016-002>
- Wallis, J. J., & North, D. (1986). Measuring the Transaction Sector in the American Economy, 1870-1970. In *Long-Term Factors in American Economic Growth* (pp. 95–162). University of Chicago Press. <https://www.nber.org/books-and-chapters/long-term-factors-american-economic-growth/measuring-transaction-sector-american-economy-1870-1970>
- Warner, M. E., & Bel, G. (2008). Competition or Monopoly? Comparing Privatization of Local Public Services in the Us and Spain. *Public Administration*, 86(3), 723–735. <https://doi.org/10.1111/j.1467-9299.2008.00700.x>

- Warsen, R., Klijn, E. H., & Koppenjan, J. (2019). Mix and Match: How Contractual and Relational Conditions Are Combined in Successful Public–Private Partnerships. *Journal of Public Administration Research and Theory*, 29(3), 375–393. <https://doi.org/10.1093/jopart/muy082>
- Whittington, J. (2012). When to Partner for Public Infrastructure?: Transaction Cost Evaluation of Design-Build Delivery. *Journal of the American Planning Association*, 78(3), 269–285. <https://doi.org/10.1080/01944363.2012.715510>
- Williamson, O. (1981). The Economics of Organization: The Transaction Cost Approach. *American Journal of Sociology*, 87(3), 548–577. <https://doi.org/10.1086/227496>
- Williamson, O. (1985). *The Economic Institutions of Capitalism*. The Free Press.
- Williamson, O. (1996). Economics and Organization: A Primer. *California Management Review*, 38(2), 131–146. <https://doi.org/10.2307/41165836>
- Wulczyn, F. (2005). *Performance-based contracting: The basics*. Foster Family-based Treatment Association.
- Zheng, J., Roehrich, J. K., & Lewis, M. A. (2008). The dynamics of contractual and relational governance: Evidence from long-term public–private procurement arrangements. *Journal of Purchasing and Supply Management*, 14(1), 43–54. <https://doi.org/10.1016/j.pursup.2008.01.004>

APPENDIX A: SEMI-STRUCTURED INTERVIEWS TO ESTIMATE TRANSACTION COSTS

Pre-interview overview

Definition of transaction costs (Petersen et al 2019)

Costs involved in finding/selecting vendors, negotiating contract terms, monitoring performance and ensuring that the delivery of final outcomes meets specifications. It comprises:

1. Ex ante costs (before commissioning)
 - Pre-contract signing search and information costs
 - Contract set-up and negotiation
2. Ex post costs (after commissioning)
 - Monitoring and enforcement

FFS (pre-KBOP) and OBC (KBOP)

A comparison of the key contract features of the FFS contract and OBC is provided below.

Table A1: Comparison of key contract features (based on Second Stage Evaluation Report)

Contract features	FFS/Pre-KBOP	OBC/KBOP
Contract parties	Kirklees Council and provider organisations	Kirklees Council and KBOP social prime
Contract management responsibility	Kirklees Council	Kirklees Council
Payment mechanism	Monthly advance block payment	Monthly outcomes payment
Key performance indicators	Service utilisation; throughput; independent living	Accommodation; education, training and employment; health and wellbeing; financial resilience
Auditing	No pre-defined evidence requirements; spot checks of qualitative evidence (e.g. workbooks)	Pre-defined evidence requirements; Kirklees Council audits every outcome
Contract duration	Max 2 years	5 years

Table A2: Definition of transaction costs for the FFS/pre-KBOP and OBC/KBOP context

Transaction cost category	Example costs	Information needed to estimate costs
Ex ante / before commissioning costs		
Pre-contract set-up	• Development work • Feasibility studies • Business case • Early phase legal costs • Market engagement costs	• Time period when work occurred • Consultant contracts or invoices for advisory work • No. of employees undertaking function • No. of FTE • Employee costs
Contract set-up and negotiation	• Legal advice • Financial advice • Procurement costs	
Ex post / after commissioning costs		
Monitoring and enforcement	• Monitoring costs • Evaluation costs • Governance costs	• Time period work occurred • No. of employees undertaking function • No. of FTE • Employee costs

Interview questions: FFS/pre-KBOP transaction costs

We would like you to think back to the pre-KBOP period (before 2019) where the provision for adults with housing-related support needs in Kirklees was commissioned as a housing floating support service. The Floating Support service provided flexible packages of support for a specified number of hours per week. Kirklees Council managed the delivery of these services, which was delivered by nine voluntary sector provider organisations. Kirklees Council initially managed 27 individual contracts with providers until the merging of contracts in 2017/18, where the Council managed 15 individual contracts. Auditing included spot checks of qualitative evidence (e.g. workbooks).

Ex ante costs

1. *Pre-contract set-up*

We would like you think about the period prior to contract set-up where you were trying to find and select suitable service providers.

- a) Can you provide a brief overview of the major tasks undertaken to search and identify these nine voluntary sector provider organisations? Or if too difficult to recall, could you provide a brief overview of your typical procurement process?
- b) We would like to understand the nature of pre-contract set-up activities and whether you would say they are predominantly “fixed” (i.e. activities remain the same no matter how many providers are sought) or “variable” (i.e. activities change based on how many providers are sought)?
- c) What time period did pre-contract set-up occur?
- d) Can you provide details on the team involved in this:
 - i. How many employees undertook this function?
 - ii. Units of time: size of FTE / approximately how much time across (b) did this occur?
- e) In order to estimate costs, we require annual salaries of the employees undertaking the functions in (c).

Are you able to provide us with the grade of these employees so we can estimate their salaries from the Local Government Association:

(<https://www.local.gov.uk/about/who-we-are-and-what-we-do/what-we-spend-and-how-we-spend-it/organisational-information/lga>)?

2. Contract set-up and negotiation

We would like you to think about the period where you were setting up the contract with the selected nine service providers.

- a) Can you provide a brief overview of the major tasks that were involved in setting up the 27 contracts? Or if too difficult to recall, could you provide an overview of your typical contract set-up/negotiation process?
- b) We would like to understand the nature of contract set-up and negotiation activities and whether you would say it is predominantly “fixed” (i.e. activities remain the same no matter how many contracts are set-up) or “variable” (i.e. activities change based on how many contracts are set-up)?
- c) What time period did contract set-up and negotiation occur?
- d) Can you provide details on the team involved in this:
 - a. How many employees undertook this function?
 - b. Units of time: size of FTE / approximately how much time across (b) did this occur?
- e) In order to estimate costs, we require annual salaries of the employees undertaking the functions in (c).

Are you able to provide us with the grade of these employees so we can estimate their salaries from the Local Government Association:

(<https://www.local.gov.uk/about/who-we-are-and-what-we-do/what-we-spend-and-how-we-spend-it/organisational-information/lga>)?

Ex post costs

1. Monitoring and enforcement

We would like you to think about the period after contract set-up with the nine service providers.

- a) Can you provide a brief overview of the major tasks that were involved in monitoring and enforcing contracts? Or if too difficult to recall, could you provide an overview of your typical process post contract set-up?
- b) We would like to understand the nature of monitoring and enforcement activities and whether you would say they are predominantly “fixed” (i.e. activities remain the same no matter how many contracts there are) or “variable” (i.e. activities change based on how many contracts there are)?
- c) What time period did contract monitoring and enforcement occur?

- d) Can you provide details on the team involved in this:
- a. How many employees undertook this function?
 - b. Units of time: size of FTE / approximately how much time across (b) did this occur?
- e) In order to estimate costs, we require annual salaries of the employees undertaking the functions in (c).
- Are you able to provide us with the grade of these employees so we can estimate their salaries from the Local Government Association:
(<https://www.local.gov.uk/about/who-we-are-and-what-we-do/what-we-spend-and-how-we-spend-it/organisational-information/lga>)?
- f) Was there a reduction in time required for monitoring and enforcement following the merging of contracts in 2017/18?

Interview questions: OBC/KBOP transaction costs

We would like you to think about the period during which there was a shift towards the adoption of the SIB (launched in September 2019) and beyond. The KBOP OBC was commissioned by Kirklees Council who initially defined the outcome measures that feature as payment triggers. The upfront capital for service provision is sourced as social investment and is managed by Bridges Fund Management. Ongoing funding is generated from the outcome payments. The investor-owned social prime (KBOP social prime) is a newly constituted organisation responsible for overall coordination of the delivery arrangement. Kirklees Council holds a contract with the KBOP social prime and this contract defines the conditions and outcome measures that direct payment to the social prime – there are pre-defined evidence requirements where Kirklees Council audits every outcome. The KBOP social prime holds bilateral contracts with eight provider organisations.

Ex ante costs

1. *Pre-contract set-up*

We would like you think about the period prior to contract set-up where you were trying to find and select suitable vendors.

- a) Can you provide an overview of the major tasks undertaken to search and identify the KBOP social prime? Or if too difficult to recall, could you provide a brief overview of your typical procurement process?
- b) What time period did pre-contract set-up occur?
- c) Can you provide details on the team involved in this:
 - i. How many employees undertook this function?
 - ii. Units of time: size of FTE / approximately how much time across (b) did this occur?
- d) In order to estimate costs, we require annual salaries of the employees undertaking the functions in (c).

Are you able to provide us with the grade of these employees so we can estimate their salaries from the Local Government Association:

(<https://www.local.gov.uk/about/who-we-are-and-what-we-do/what-we-spend-and-how-we-spend-it/organisational-information/lga>)?

2. *Contract set-up and negotiation*

We would like you to think about the period where you were setting up the contract with the KBOP social prime.

- a) Can you provide an overview of the major tasks that were involved in setting up the contract? Or if too difficult to recall, could you provide an overview of your typical contract set-up/negotiation process?
- b) What time period did contract set-up and negotiation occur?
- c) Can you provide details on the team involved in this:
 - a. How many employees undertook this function?
 - b. Units of time: size of FTE / approximately how much time across (b) did this occur?
- d) In order to estimate costs, we require annual salaries of the employees undertaking the functions in (c).

Are you able to provide us with the grade of these employees so we can estimate their salaries from the Local Government Association:

(<https://www.local.gov.uk/about/who-we-are-and-what-we-do/what-we-spend-and-how-we-spend-it/organisational-information/lga>)?

Ex post costs

1. *Monitoring and enforcement*

We would like you to think about the period after contract set-up with the KBOP social prime.

- a) Can you provide an overview of the major tasks that were involved in monitoring and enforcing contracts? Or if too difficult to recall, could you provide an overview of your typical process post contract set-up?
- b) What time period did contract monitoring and enforcement occur?
- c) Can you provide details on the team involved in this:
 - a. How many employees undertook this function?
 - b. Units of time: size of FTE / approximately how much time across (b) did this occur?
- d) In order to estimate costs, we require annual salaries of the employees undertaking the functions in (c).

Are you able to provide us with the grade of these employees so we can estimate their salaries from the Local Government Association:

(<https://www.local.gov.uk/about/who-we-are-and-what-we-do/what-we-spend-and-how-we-spend-it/organisational-information/lga>)?

APPENDIX B: MATCHING VARIABLES

Table B1: Matching variables

Variable	Description	Timeframe
Demographics		
Age	Age at the intervention start date, in years. Originally sourced from DWP Customer Information System (CIS) extracts	Intervention start date
Age squared	Derived from age at the intervention start date, in years. Originally sourced from DWP Customer Information System (CIS) extracts	Intervention start date
Sex of the individual; 1 is male and 0 is female)	Originally sourced from DWP Customer Information System (CIS) extracts	Intervention start date
Indicator of whether an individual has dependent children	Based on whether the individual receives Child Tax Credit, Child Benefit – Parent, Maternity Allowance or Child Disability Payment at the first month of their intervention start	Intervention start month
History of benefit receipt		
Flag indicating if there is history of Housing Benefit receipt in the two years prior to intervention start	Originally sourced from Single Housing Benefit (SHBE) extracts	2 years pre-intervention
Flag indicating if there is history of Child Benefit – Parent receipt in the two years prior to intervention start	Originally sourced from HMRC Child Benefit (CHB) extracts	2 years pre-intervention
Flag indicating if there is history of Working Tax Credit receipt in the two years prior to intervention start	Originally sourced from HMRC Working Tax Credits (WTC) - Work and Pension Longitudinal Study (WPLS) extracts	2 years pre-intervention
Flag indicating if there is history of Child Tax Credit receipt in the two years prior to intervention start	Originally sourced from HMRC Child Tax Credits (CTC) - Work and Pension Longitudinal Study (WPLS) extracts	2 years pre-intervention
Flag indicating if there is history of Employment and Support Allowance receipt in the two years prior to intervention start	Originally sourced from DWP Generalised Matching Service Benefit Data (GMS) extracts	2 years pre-intervention
Flag indicating if there is history of Jobseeker's Allowance receipt in the two years prior to intervention start	Originally sourced from DWP Generalised Matching Service Benefit Data (GMS) extracts	2 years pre-intervention
Flag indicating if there is history of Incapacity Benefit receipt in the two years prior to intervention start	Originally sourced from DWP Generalised Matching Service Benefit Data (GMS) extracts	2 years pre-intervention
Flag indicating if there is history of Invalid Care Support receipt in the two years prior to intervention start	Originally sourced from DWP Generalised Matching Service Benefit Data (GMS) extracts	2 years pre-intervention
Flag indicating if there is history of Income Support receipt in the two years prior to intervention start	Originally sourced from DWP Generalised Matching Service Benefit Data (GMS) extracts	2 years pre-intervention
Flag indicating if there is history of Passported Incapacity Benefit receipt in the two years prior to intervention start	Originally sourced from DWP Generalised Matching Service Benefit Data (GMS) extracts	2 years pre-intervention

Variable	Description	Timeframe
Flag indicating if there is history of Bereavement Benefit receipt in the two years prior to intervention start	Originally sourced from DWP Generalised Matching Service Benefit Data (GMS) extracts	2 years pre-intervention
Flag indicating if there is history of Bereavement Support Payment receipt in the two years prior to intervention start	Originally sourced from DWP Generalised Matching Service Benefit Data (GMS) extracts	2 years pre-intervention
Flag indicating if there is history of Severe Disablement Allowance receipt in the two years prior to intervention start	Originally sourced from DWP Generalised Matching Service Benefit Data (GMS) extracts	2 years pre-intervention
Flag indicating if there is history of Widow's Benefit receipt in the two years prior to intervention start	Originally sourced from DWP Generalised Matching Service Benefit Data (GMS) extracts	2 years pre-intervention
Flag indicating if there is history of Universal Credit receipt in the two years prior to intervention start	Originally sourced from DWP Universal Credit (UC) extracts	2 years pre-intervention
Flag indicating if there is history of employment in the two years prior to intervention start	Originally sourced from the Pay As You Earn (PAYE) - HMRC 56NZ P14 extract	2 years pre-intervention
Flag indicating if there is history of Disability Living Allowance receipt in the two years prior to intervention start	Originally sourced from DWP Generalised Matching Service Benefit Data (GMS) extracts	2 years pre-intervention
Flag indicating if there is history of Personal Independence Payment receipt in the two years prior to intervention start	Originally sourced from DWP Personal Independence Payment extracts	2 years pre-intervention
Flag indicating if there is a recorded sanction in the two years prior to intervention start	Derived from DWP and HMRC benefit extracts (see above)	2 years pre-intervention
Previous involvement in Department for Work & Pensions interventions		
Participation in a DWP intervention or programme in the two years prior to intervention start	Originally sourced from DWP interventions datasets (e.g., European Social Fund dataset)	2 years pre-intervention
Participation in a DWP intervention or programme at the intervention start	Originally sourced from DWP interventions datasets (e.g., European Social Fund dataset)	2 years pre-intervention
Referral to, but not participation, in a DWP intervention in the two years prior to intervention start	Originally sourced from DWP interventions datasets (e.g., European Social Fund dataset)	2 years pre-intervention
Labour market status history		
Monthly indicator of employment or self-employment status during the 24-month pre-intervention period	Flagged if an individual has an active employment or self-employment record. Determined by During The Month (DTM) conditionality in Universal Credit (UC)	2 years pre-intervention
Monthly indicator of 'looking for work' labour market status during the 24-month pre-intervention period	Flagged if an individual has an active Jobseeker's Allowance (JSA) record or an active Universal Credit (UC) record where their conditionality regime is either searching for work or working – with requirements	2 years pre-intervention

Variable	Description	Timeframe
Monthly indicator of inactive labour market status during the 24-month pre-intervention period	Flagged if an individual has active an Employment Support Allowance (ESA) record or an active UC record where their conditionality group is either no work requirements, preparing for work, or planning for work	2 years pre-intervention
Monthly indicator of other labour market status during the 24-month pre-intervention period	Flagged if an individual does not fall under any of the above labour market categories	2 years pre-intervention

Based on UK Government administrative data from Central Analysis and Science Strategy Unit (2024).

APPENDIX C: BALANCE STATISTICS

Table C1: Post matching balance diagnostics – 100 Nearest Neighbours with 0.01 Caliper

Balance diagnostics	
Matching estimator	100 NNM, 0.01 caliper
Mean Bias Reduction	2.22
Median Bias Reduction	1.78
Rubin's B	15.68
Rubin's R	1.00
Variance	2.84
Skewness	0.61
Kurtosis	2.28
Off support OBC group participants	17
Off support FFS contract group participants	0
Total N	2,728

APPENDIX D: VARYING THE CHOICE OF MATCHING ESTIMATOR

To test the sensitivity of the impact estimates, balance diagnostics were evaluated across a range of alternative matching estimators (Table D1). The results consistently indicate that covariate balance was maintained across all specifications, reinforcing the robustness of the matching procedure. Mean bias reduction ranged from 2.12% to 3.43%, while median bias reduction varied between 1.55% and 2.99%, reflecting minimal residual imbalance in all cases. Rubin's B values remained within accepted thresholds, though slightly elevated for kernel and radius matching with a 0.001 caliper. Rubin's R values remained close to 1 throughout, suggesting stable variance ratios between groups.

Although the more restrictive estimators (kernel 0.001 and radius 0.001) delivered marginally higher bias reduction, this came at the cost of significant sample loss. In particular, radius 0.001 excluded 136 treated individuals due to lack of suitable matches. Distributional statistics, including variance (ranging from 2.74 to 5.65), skewness (0.07 to 0.76) and kurtosis (2.10 to 2.69), remained stable and within reasonable bounds for all specifications, indicating no substantial distortions.

These diagnostics highlight that the primary matching approach (100 NN with 0.01 caliper) provides the most favourable trade-off between covariate balance and sample retention. While alternative estimators achieved broadly similar balance, none demonstrated improved overall performance in this context. Consequently, 100 NN with a 0.01 caliper was retained as the most robust and effective matching strategy for this evaluation.

Table D1: Comparison of balance diagnostics across alternative matching estimators

Matching estimator	Mean Bias Reduction Post- Matching	Median Bias Reduction Post- Matching	Rubin's B	Rubin's R	Variance	Skewness	Kurtosis	Off support treatment group individuals	Off support control group individuals	Total N
100 NN 0.01 caliper	2.22	1.78	15.6805	1.00	2.84	0.61	2.28	17	0	2,728
20 NN 0.01 caliper	2.12	1.80	16.6663	1.01	2.74	0.76	2.69	17	0	2,728
Kernel 0.01 caliper	2.15	1.55	16.4491	1.07	3.12	0.60	2.17	1	0	2,760
Kernel 0.001 caliper	3.28	2.99	21.8959	1.03	5.65	0.45	2.19	1	0	2,760
Radius 0.01 caliper	2.20	1.83	15.7281	1.01	2.89	0.61	2.26	17	0	2,728
Radius 0.001 caliper	3.43	2.97	18.9220	0.77	3.71	0.07	2.10	136	0	2,490