Mental Health and Employment Partnership:

The final evaluation of the Life Chances Fund outcomes partnerships

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

Department for Culture, Media & Sport

Final evaluation report July 2025

Emily Hulse, Eve Grennan, Dr Mara Airoldi, Dr Eleanor Carter

Acknowledgements

The authors gratefully acknowledge the support of colleagues at the Department for Culture, Media and Sport. The authors also acknowledge the National Lottery Community Fund, which administered the Life Chances Fund. We extend thanks to Social Finance for their support in compiling relevant MHEP documentation required for our research activities over the multi-year evaluation period. We are very grateful to Isaac Grennan for supporting the quantitative analysis.

We are particularly grateful to the staff at each MHEP delivery site who participated in interviews and shared relevant data. The research wouldn't have been possible without the participation of all MHEP delivery sites.

Contribution statement

Emily Hulse authored this report. She contributed to the research design, developed the conceptual report design and led the data collection and analysis. Emily reviewed and co-edited the report.

Eve Grennan co-authored this report. She contributed to data analysis and drafting of the report.

Dr Mara Airoldi co-authored this report. She reviewed and co-edited the report. Mara also provided final quality assurance for the report.

Dr Eleanor Carter co-authored this report. She designed the overall research strategy, supported data collection and analysis. Eleanor is responsible for overall research quality.

Jessica Reedy authored the Executive Summary and co-edited the report. She also supported final reviews, formatting, and quality assurance processes.

Michael Gibson co-edited the report.

About the Government Outcomes Lab

The Government Outcomes Lab (GO Lab) is a global centre of expertise based at the Blavatnik School of Government, University of Oxford. Our mission is to enable governments across the world to foster effective partnerships with the non-profit and private sectors for better outcomes.

We are an international team of multi-disciplinary researchers, data specialists & policy experts. We generate actionable knowledge, offering a comprehensive and evidence-based approach to the study of cross-sector partnerships through the three main strands of our work: research, data and engagement.

You can find out more about our work at golab.bsg.ox.ac.uk.

About the Department for Culture, Media and Sport

The Department for Culture, Media and Sport supports culture, arts, media, sport, tourism and civil society across every part of England — recognising the UK's world-leading position in these areas and the importance of these sectors in contributing so much to our economy, way of life and our reputation around the world. The department champions sport for all at every level, supports our world-leading cultural and creative industries, and enhances the cohesiveness of our communities.

DCMS delivered the Life Chances Fund (LCF) between 2016-2025. The LCF aimed to help those people in society who face the most significant barriers to leading happy and productive lives. The £70m Social Outcome Partnership fund contributed to outcome payments for locally commissioned social outcomes contracts which involve socially-minded investors. Projects have helped support tens of thousands of beneficiaries in areas like youth unemployment, mental health and homelessness.

Recommended citation for this report

Hulse, E., Grennan, E., Airoldi, M., Carter, E. (2025) The Mental Health and Employment Partnership: the final evaluation of the Life Chances Fund outcomes partnerships. Government Outcomes Lab, Blavatnik School of Government, University of Oxford.

Table of Contents

EXECUTIVE SUMMARY

What are the Life Chances Fund & the Mental Health Employment Partnership?

The Life Chances Fund (LCF) was a £70 million programme funded by the Department for Culture, Media and Sport (DCMS). It ran between 2016–2025 and is the largest outcomes fund launched to date in the UK. The LCF was designed to tackle complex social problems across policy areas including child and family welfare, homelessness, health and wellbeing, employment and training, and more. You can learn more about the LCF on our website.

The LCF was delivered through 29 locally-commissioned social outcomes partnerships (SOPs – also known as social impact bonds). The Mental Health and Employment Partnership (MHEP) SOPs made up five of these 29 SOPs and were delivered across London boroughs and North England. MHEP supported the delivery of an intervention known as 'Individual Placement and Support' (IPS) to help people experiencing mental health issues or learning disabilities to find and remain in competitive, paid work. Established in 2015 by the organisation Social Finance, MHEP was backed by social investment from Big Issue Invest for a total of £1.2m across the SOPs of Haringey and Barnet, Shropshire, Enfield, Tower Hamlets (Mental Health), and Tower Hamlets (Learning Disabilities). MHEP features within a wider set of evaluations the Government Outcomes Lab undertook for DCMS as their knowledge and learning partner for the Life Chances Fund. You can read more about MHEP and SOPs on the Government Outcomes Lab website.

Individual Placement and Support (IPS)
IPS is a rigorously tested employment support intervention that follows a 'place, then train' model, where employment specialists support service users to secure employment quickly before providing them with ongoing support to ensure sustainment.

This report

This is the third and final report of a five-year research study investigating the effectiveness of social outcomes partnerships as a commissioning tool to improve social outcomes for citizens. The report asks two primary questions: whether the MHEP SOP made a difference to the social outcomes achieved, compared with alternative commissioning approaches, and through which mechanisms it contributed to improved services and positive social outcomes.

This inquiry is framed through four questions in the report:

- Were the MHEP SOPs effective in achieving their outcomes targets and how does their effectiveness compare with that of traditionally commissioned IPS contracts?
- Did the SOP have higher costs than expected and if so, why?
- Was the SOPs' outcomes achievement (the SOP effect) related to the intensity of the performance management or performance incentive?
- How were different actors incentivised for performance?

It is worth noting some limitations to this report. It was not possible to directly compare IPS delivered through a SOP with IPS delivered through a traditional grant taking place at the same time and subject to the same environmental conditions (eg delivering during the COVID-19 disruptions). The lack of direct comparator data from non-SOP sites made it impossible to answer the original research question quantitatively. Instead, the evaluation relies heavily on qualitative methods and a *dose-response analysis*, examining whether different levels (doses) of performance incentives and management affect outcomes. Therefore, this final report examines the effects of varying SOP intensity within MHEP on: 1) outcomes, 2) perceived costs, and 3) incentives (and performance management). This approach is useful given the variation in SOP structures and lack of a standard definition.

Other limitations included further data collection challenges including limited participation from commissioners and providers for interviews, likely due to research fatigue and staff turnover, resulting in an interview sample mostly of Social Finance staff. The fact that only four SOP sites were included in the dose analysis limits project-level conclusions, though a large service-user dataset strengthens analysis at the individual level. Finally, it is worth noting that the report relies on provider-supplied data (with potential unobserved biases) about service delivery which was complicated by the impact of COVID-19 on service delivery and the difficulty of isolating the 'SOP effect' from the broader influence of MHEP and IPS.

Despite these challenges, the MHEP SOPs remain a strong case study due to:

1. The use of a **proven intervention (IPS)** with an established fidelity model
2. Their **large scale**, serving over 10,000 users since 2015
3. Access to **individual-level data** through a partnership with Social Finance, allowing for robust evaluation design.

Were the MHEP SOPs effective in achieving their outcomes targets and how does their effectiveness compare with that of traditionally commissioned IPS contracts?

The MHEP SOPs achieved engagement from 68% of service users; 55% of service users started jobs, and 55% sustained jobs at over 16 hours per week. Conversion

rates and IPS fidelity scores were largely in line with national expectations for IPS, *despite COVID-19 disruptions*. Compared with traditionally commissioned IPS contracts, MHEP SOPs benefited from stronger performance management, closer commissioner-provider relationships, and a clearer focus on outcomes – although some stakeholders felt SOPs imposed greater administrative complexity.

Did the SOP have higher costs than expected and if so, why?

While some costs of the SOP were higher than expected, particularly for management and setup, others such as investment and delivery were lower, helping to balance the overall spend. Higher management costs were largely driven by the time and care taken in negotiating complex multi-stakeholder contracts, managing performance, and navigating external challenges such as COVID-19 and staffing changes. Setting up SOPs required extensive modelling and coordination, even for projects that did not ultimately go ahead. However, these investments helped build strong foundations, and efficiencies were achieved through standardised tools, a clear service specification and a centralised performance management structure.

Was the SOPs' outcomes achievement (the SOP effect) related to the intensity of the performance management or performance incentive?

The analysis found a significant link between performance incentives and job outcomes, particularly in sites where more of the contract value depended on results. Our modelling suggests that for every £1,000 increase in outcome payments, job starts were 21% more likely and happened 17% faster. Enfield, where 30% of provider funding was tied to performance, saw especially strong results. However, the data suggests other unmeasured factors (eg the intermediary's increased support and attention) also influenced results. No clear quantitative results were found between performance management exposure and outcomes, highlighting the need for further research.

How were different actors incentivised for performance?

In the MHEP SOPs, the strongest performance incentives sat with the intermediary, Social Finance, whose payment was entirely outcomes-based and who also faced investor scrutiny and reputational risk. While investors recovered capital only after operational costs were met, and most providers were shielded through block payments (except in some cases like Enfield, where up to 30% of funding was outcome-based), commissioners and central funders bore minimal financial risk. The intermediary's incentive was most effective when supported by clear role separation from investors, transparent governance, and shared goals with commissioners – though practical barriers like contract complexity and resource constraints could weaken this effect.

Policy recommendations

1. Recognise the significant relationship between financial incentives and outcomes achievement.
The research found that for every £1,000 increase in incentives, the likelihood of outcomes achievement increased by about 20%. The reverse also applies. However, this finding does not simply mean that greater financial input automatically results in more outcomes from service providers. Although the effect was statistically significant for the MHEP SOPs, the analysis also indicated that more data is needed before it can be used to make predictions.

The qualitative analysis revealed that the important factor was the intermediary, who translated the hard incentives (more outcome payments) into increased softer incentives (eg intrinsic motivation, personal bests) for service providers by means such as more relational and operational support. Thus, the incentives work less through hard-edge financial pressures on service providers and more through the motivation and accountability coming from the intermediaries. These incentives' effectiveness at the intermediary level relied on clear role separation between the investor and the performance managers; transparency; a governance board that was capable of holding people accountable; and a robust reporting and contractual framework.

***Recommendation**: Design financial incentives with a clear understanding of how different actors respond: intermediaries are more likely to be influenced by financial incentives, while service providers may be more motivated by softer, relational or mission-driven factors.*

2. Anticipate setup and wind-down effort as a foundation for adaptive delivery
The MHEP SOPs required appreciably more time and effort during setup and completion phases than initially anticipated due to complex negotiations, outcomes modelling, and end-of-grant reconciliation. However, these investments laid the groundwork for more adaptive, data-driven delivery during the contract. Stakeholders noted that while the transaction costs were high up front, they enabled robust structures, trust and shared understanding, which ultimately supported better performance management and problem-solving throughout delivery.

***Recommendation**: Build in adequate time and resourcing for SOP setup and closure phases, recognising them as critical foundations for continuous improvement and collaborative service delivery.*

3. Enable responsive problem-solving via continuous monitoring, frequent engagement, and bespoke data analytics
A key strength of the MHEP SOP model was its ability to enable responsive

problem-solving through continuous monitoring, frequent engagement and bespoke data analytics. The intermediary played a central role in identifying underperformance early and working closely with providers and commissioners to adapt strategies in real time. Regular data reviews, site visits and tailored performance improvement plans allowed challenges – such as staffing gaps, referral delays or outcome dips – to be addressed proactively. This dynamic, data-informed approach contrasted with more static traditional contracts and was widely credited by stakeholders with improving both service quality and outcomes achievement across the SOPs.

***Recommendation**: Programmes should incorporate continuous monitoring, frequent engagement and tailored data analytics to enable responsive problem-solving and improve service outcomes.*

Previous report findings

Report 1: The first report found that:

MHEP SOPs provided additional value compared with traditional commissioning via:

- A dedicated performance management function that was seen to drive additional focus on achieving outcomes
- More effective working culture within each local partnership
- Identifying and successfully unlocking the LCF funding. This was understood to bring additional financial and human resources to projects.

Report 2: The second report found that:

- The SOPs improved accountability and commissioning practice compared to traditional contracts of IPS
- Social Finance's MHEP contributed to the national scaling of IPS in the NHS
- MHEP SOPs' contractual and payment structures were seen as unnecessarily complex and could be simplified via earlier buy-in for design principles, annual caps, and more realistic expectations in forecasting outcomes performance
- MHEP brought enhanced capacity to providers by building an IT data system, talent pipelines for staff, and efficient data routines. It brought enhanced capacity to commissioners via experience in partnership working and the creation of a new baseline for expected IPS outcomes
- MHEP SOPs' incentives on providers were more muted than expected.

Introduction: The evaluation and research context

The Department for Culture, Media and Sport (DCMS) commissioned the LCF evaluation to understand how SOPs add value when compared with more conventional public service commissioning arrangements.[1]

The evaluation was structured across three strands:

- **Strand 1 – Fund-level programme evaluation**
  Sought to evaluate the whole LCF fund as a tool for growing the social outcomes partnership market. It evaluated the process involved and lessons learnt from fund administration[2]

- **Strand 2 - SOP mechanism evaluation**
  Sought to evaluate SOPs as a commissioning tool. It evaluated the impact, process and value for money of selected LCF SOPs and will compare the SOP model with alternative commissioning approaches

- **Strand 3 - SOP project-led evaluation**
  Sought to evaluate innovative interventions and specific aspects of the SOP model to inform local implementation[3]

The Government Outcomes Lab was responsible for the SOP mechanism evaluation (Strand 2). The LCF evaluation and GO Lab's accompanying research on social outcomes partnerships aim to respond to current evidence gaps by focusing specifically on SOPs as a tool for public service delivery and reform rather than centring only on the intervention effect. The ambition is to assess 'the SOP effect' – that is, the influence of this commissioning model on social outcomes.

Previous evaluations of SOP programmes have primarily focused on the implementation or efficacy of specific interventions (ie the particular service funded by the SOP), often without robust quantitative impact evaluation (Carter et al., 2018; see also Fox & Morris, 2019). Impact evaluations are important to help us understand how SOPs differ from typical government commissioning mechanisms in terms of the social 'impact' they deliver against objectives. As the largest outcome fund in the UK, the LCF provided an opportunity to undertake both process and impact evaluations to help improve future policy and practice.

This is the third and final evaluation of the Life Chances Fund's Mental Health and Employment Partnership (MHEP). This report covers five MHEP social outcomes partnerships (SOPs, also known as social impact bonds) contracted under the Life Chances Fund: Haringey and Barnet, Shropshire, Enfield, Tower Hamlets Mental

Health, and Tower Hamlets Learning Disabilities. SOPs are defined as contractual arrangements that have two key characteristics:

- A contractual relationship that includes payment for social outcomes achieved (ie an outcomes contract)
- Up-front repayable finance provided by a third party, the repayment of which is (at least partially) conditional on achieving specified outcomes.

1. Background to the Life Chances Fund

The Life Chances Fund (LCF) was a £70 million fund supporting the growth and development of 29 locally-commissioned social outcomes partnerships (SOPs), also known as social impact bonds (SIBs), in England. These outcomes-based projects were co-commissioned by central government and a range of local public sector organisations.

LCF projects aimed to tackle complex social problems across policy areas like child and family welfare, homelessness, health and wellbeing, employment and training, criminal justice, and education and early years. Following three application rounds, funding was made available for multi-year SOP projects to run within the LCF's nine-year lifespan from July 2016 to March 2025. The first LCF projects began service delivery in 2018, with the bulk of projects launching between 2019 and 2020. LCF Projects were only able to claim payments for outcomes achieved up to the end of September 2024. Some projects continued to deliver services beyond the lifetime of the LCF.

The Fund had the following objectives[1]:

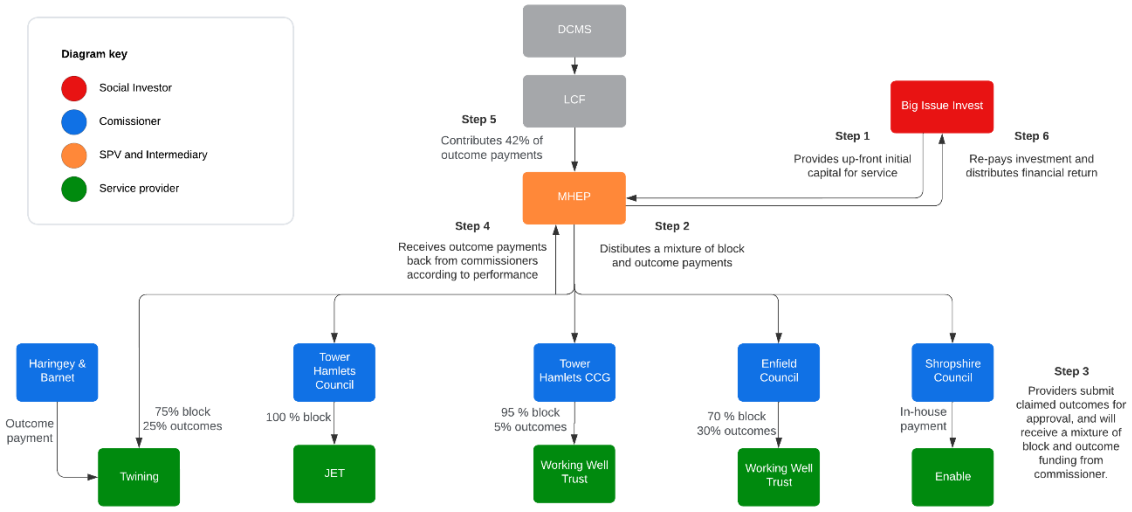
- Increasing the number and scale of SOPs in England
- Making it easier and quicker to set up a SOP
- Generating public sector efficiencies by delivering better outcomes and using these successes to understand how and whether cashable savings could be achieved
- Increasing social innovation and building a clear evidence base of 'what works'
- Increasing the amount of capital available to a wider range of voluntary, community and social enterprise (VCSE) sector providers to enable them to compete for public sector contracts
- Providing better evidence of the effectiveness of the SOP mechanism and the savings that are being accrued
- Growing the scale of the social investment market.

The LCF was administered by The National Lottery Community Fund (TNLCF, formerly known as the Big Lottery Fund) on behalf of the Department for Culture, Media and Sport (DCMS).

2. Background to the Mental Health and Employment Partnership

The Mental Health and Employment Partnership (MHEP) was established in 2015 to drive the expansion of a high-quality employment support intervention known as Individual Placement and Support (IPS) for people with severe mental illness, addictions, and/or learning disabilities.

Figure 1: MHEP's SOPs supported by the LCF



Note: MHEP as a SPV was run by Social Finance, which acted as the intermediary in the SOP working with commissioners in local authorities or NHS. Social Finance played a central role coordinating the payment flows and managing the relationship with the investor.

MHEP was set up as a special purpose vehicle (SPV), ie a separate legal entity created and managed by Social Finance that acted as the contractual counterpart originally for nine SOPs. MHEP was run through an intermediary SOP model in which Social Finance was the intermediary managing the performance and the contract. With the exception of Haringey & Barnet SOP, each MHEP SOP had an outcome-based contract sitting between MHEP and the local commissioner (typically a city council). To manage the MHEP SPV and the SOPs within it, Social Finance allocated to the project implementation a core staff consisting of: 1-2 analysts, 1 manager, 1 operational expertise director and 1 oversight director, as well as a governance board with 4-5 volunteers, including the impact investor.

MHEP was distinct in several aspects. Firstly, it began the world's first SOPs aimed at helping people with mental health issues into paid employment. Secondly, to date, it is the longest running SOPs project globally, having delivered services for

nine years. Thirdly, unlike other SOPs which allow providers flexibility on the service intervention funded as long as the outcomes are achieved, MHEP mandated for its SOPs that the service delivery intervention was 'IPS.' Fourthly, MHEP acted as an outcome payer on the outcome contracts, meaning that it partnered with the traditional commissioner from the local authority or NHS for the outcome-based contract (as opposed to being a traditional intermediary for the LCF's payments as in most other SOPs of the LCF). Lastly, it structured the contractual relationships of the SOPs through the SPV, whilst in several SOPs there is a direct relationship between the outcome payer and service providers. We would hence need to be cautious in interpreting the wider generalisability of the result.

Throughout the period of its operation, several central government outcomes funds supported MHEP's activities (see timeline in Figure 2). MHEP initially partnered with three areas in 2016 (Haringey, Tower Hamlets, and Staffordshire) to secure £1.3m of 'top-up' funding from the Commissioning Better Outcomes Fund[1] and the government's Social Outcomes Fund[2]. MHEP successfully applied to the Commissioning Better Outcomes Fund[3] in 2017 for second top-up funding to extend to three more services: Enfield, Camden, and Barnet under the same CBO award. A third top-up outcome funding was awarded to West London Alliance of £1.08m, which selected MHEP as the social investment partner in January 2018 to fund its Addictions IPS Social Impact Bond. Lastly, a fourth top-up outcome funding from the Life Chances Fund[4] extended the platform to another five services (Haringey and Barnet, Shropshire, Enfield, Tower Hamlets Mental Health, Tower Hamlets Learning Disabilities).

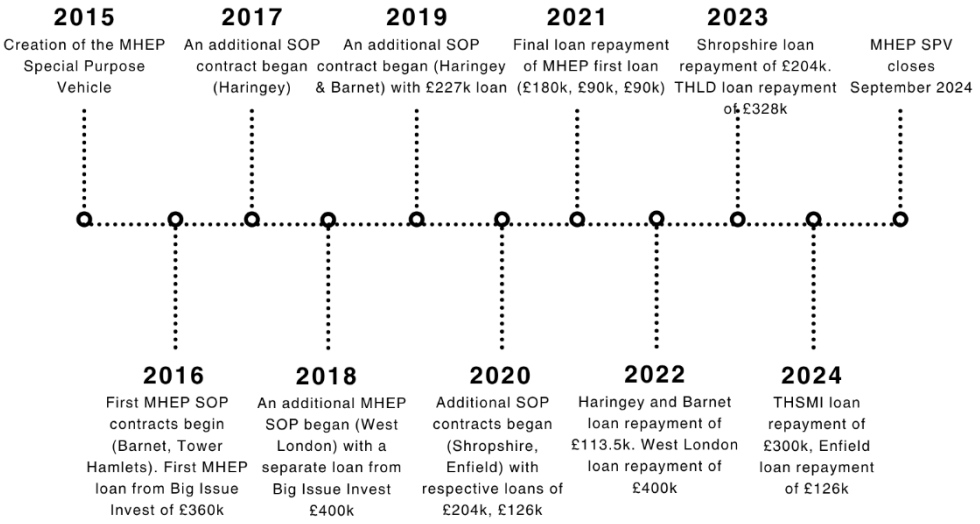
[1] See the INDIGO Outcomes Fund Directory for more information on this fund: https://golab.bsg.ox.ac.uk/knowledge-bank/indigo/fund-directory/INDIGO-FUND-0003/

[2] See the INDIGO Outcomes Fund Directory for more information on this fund: https://golab.bsg.ox.ac.uk/knowledge-bank/indigo/fund-directory/INDIGO-FUND-0004/.

[3] See the INDIGO Outcomes Fund Directory for more information on this fund: https://golab.bsg.ox.ac.uk/knowledge-bank/indigo/fund-directory/INDIGO-FUND-0003/

[4] See the INDIGO Outcomes Fund Directory for more information on this fund: https://golab.bsg.ox.ac.uk/knowledge-bank/indigo/fund-directory/INDIGO-FUND-0012/

Figure 2: MHEP SOP Timeline



Note: SOP= social outcomes partnership, THLD= Tower Hamlet's Learning Disabilities, THSMI=Tower Hamlet Severe Mental Illness, SPV=special purpose vehicle. These are unsecured loans explained below.

Throughout its lifetime, MHEP was backed by Big Issue Invest, a socially motivated impact investor, and Health and Employment Partnerships, a social purpose organisation within Social Finance[5].

The total investment made by Big Issue Invest was £1.945m, through 7 unsecured 'loans' for £360,000 for first round of MHEP SOPs (Haringey, Tower Hamlets, Staffordshire, Enfield, Camden, and Barnet), £400,000 for Addictions SOP, £227,000 for Haringey & Barnet, 328,000 for Tower Hamlets Learning Disabilities, £126,000 for Enfield, £300,000 for Tower Hamlets Severe Mental Illness, and £204,000 for Shropshire.

Big Issue Invest took 100% of the financial risk of MHEP, which managed the contracts with providers. Big Issue Invest had a significant equity investment in MHEP(£40,000 for 40,000 Class A shares). However, the key resources to finance the projects were unsecured loans. This meant that, unless the outcomes contracts produced a surplus (ie outcome payments greater than costs), they did not claim any payments. This was a high risk: they did not claim any payment while outcome payments started to be released. It was only after sufficient outcomes payments had been paid to cover all the operational costs that they started making claims

[5] Social Finance is a not-for-profit consultancy aiming to find better ways of tackling social problems in the UK and globally. In 2010, they pioneered the first social outcomes partnership (social impact bond) in the world in a contract for services for prison leavers in Peterborough. Social outcomes partnerships have since expanded globally with an estimated over 700 million USD raised in social investment, 41 countries, and 308 total projects.

for loan repayment. This is capital at risk, regardless of the label of 'loan.' However, surpluses on one contract could be used to offset losses on a different contract managed by MHEP (including contracts not supported by the LCF), allowing BII to better manage the risk by pooling it. If the surplus minus the negative cash balance was a positive cash balance, 49% of that would go to the investor, 51% to MHEP.

MHEP varied the basic SOP structure, taking a platform approach to support multiple SOPs via a special purpose vehicle[6]. Over time, MHEP developed different co-commissioning structures and explored the most appropriate outcomes for commissioners to make payments for in IPS services. All of MHEP's performance management and contract management was provided by Social Finance, a not-for-profit founded in 2007 to provide consultancy support in finance, strategy, design and data to build partnerships. The SPV structure of MHEP is not unique, with the majority of Life Chances Fund projects (61%) having this structure (19 out of the original 31) rather than a direct contract between providers and outcome payers.

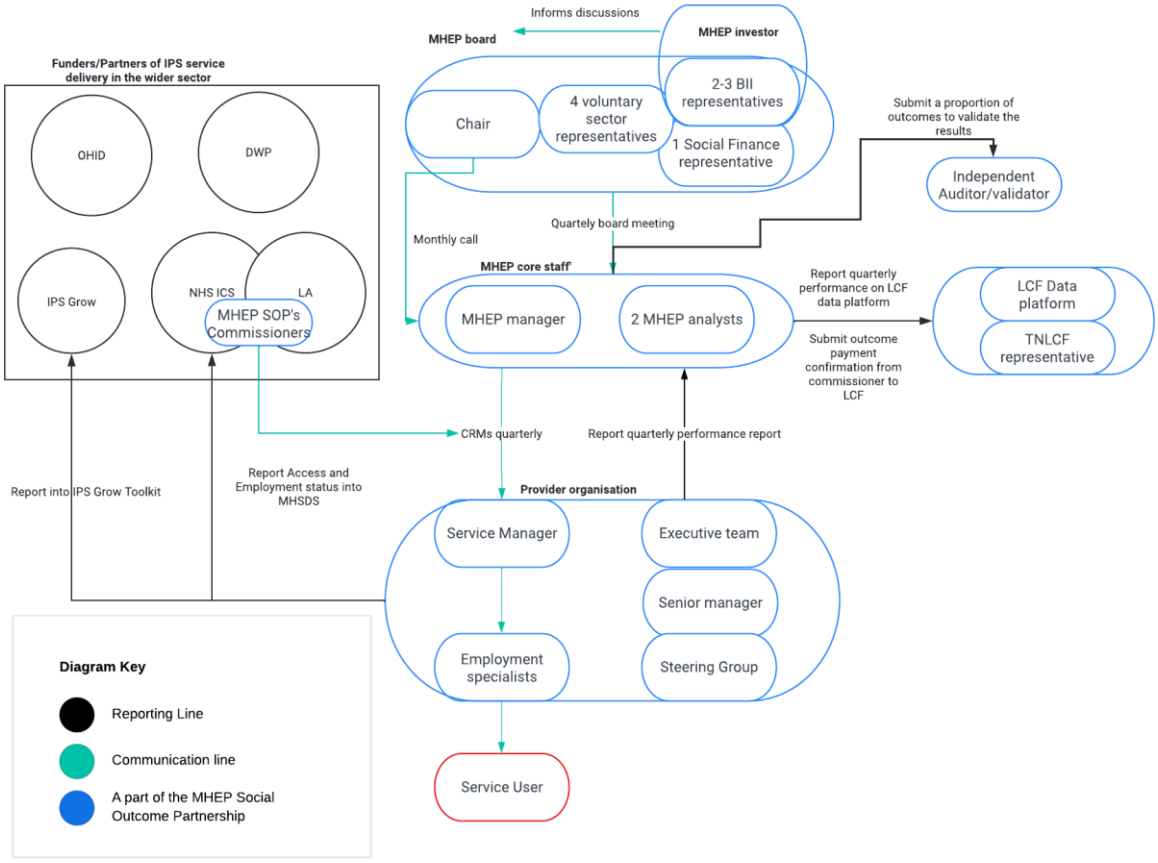
Figure 1 above describes the basic structure of the SOPs within MHEP[7]. This structure is as follows:

1. Upfront social investment from Big Issue Invest is channelled through MHEP as a special purpose vehicle/company.
2. MHEP provides block payments every quarter to commissioners (to pay to providers).
3. Providers' quarterly outcomes claims are submitted to commissioners, and outcomes funding (in addition to block payments) is paid following approval of claims. There are three outcomes tied to payment: engagement, job start and job sustainment.
4. Commissioner pays MHEP quarterly payments based on the achievement of these outcomes.
5. DCMS provides 'top-up' funding to outcomes payments through the LCF.
6. MHEP leads the reporting of outcomes and manages funding flows to Big Issue Invest.

[6] A legal entity (usually a limited company) that is created solely for a financial transaction or to fulfil a specific contractual objective. Special purpose vehicles have been sometimes used in the structuring of impact bonds. SPVs are typically a way to isolate the financial and legal risks of specific contracts and to protect parent companies from exposing their entire balance sheet to the liabilities of those specific contracts.

[7] The specific SOP structure under MHEP is described in previous reports (Hulse et al (2023), Hulse et al (2024)).

Figure 3: Communication and Reporting Lines of the MHEP SOPs



Note: OHID=The Office for Health Improvement and Disparities (a government unit within the Department of Health and Social Care); DWP=The Department for Work and Pensions; ICS=Integrated Care Systems; LA=local authorities; SOP=Social Outcome Partnership; LCF=Life Chances Fund; BII=Big Issue Invest; MHEP=Mental Health and Employment Partnership; IPS=Individual Placement and Support; TNLCF=The National Lottery Community Fund

3. Background to Individual Placement and Support

Most SOPs adopt a 'black box' model, allowing significant discretion in defining the service and how outcomes are achieved. In comparison, MHEP mandated the delivery of Individual Placement and Support (IPS) as a prescribed, manualised intervention, supported by a well-defined operating framework. IPS involves the integration of employment specialists in mental health teams to support the return to work of people experiencing mental health (and addiction) issues. It is based on 'place then train' principles, and evidence suggests it is more effective than traditional approaches such as vocational training and sheltered work (Modini et al. 2016).

This evidence has been bolstered by rigorous trials across a range of different cohorts, including the IPS Trial of Homelessness (2024) by Centre for Homelessness Impact, The Health Led employment trials (HLT) by DWP (2018), and Individual Placement and Support for Alcohol and Drug Dependence (IPS-AD) randomised controlled trial (2022). One recent meta-analysis concluded that IPS consistently outperformed traditional vocational programmes and that the evidence for efficacy of IPS is 'very strong' and can be generalised between countries (Brinchmann et al 2019). While the intervention was originally developed in the United States for those with severe mental illness, it has demonstrated effective success in helping other people with complex needs to remain in employment.

The performance of IPS programmes is enhanced by adherence to IPS principles, integration with mental health services, skilled and dedicated staff, effective management, and contextual adaptation (Priest & Lockett, 2020; Browne & Waghorn, 2010; Porteous & Waghorn, 2007; Waghorn et al., 2011; van Erp et al., 2007). These factors work synergistically to create an environment conducive to achieving positive employment outcomes for individuals with mental illnesses.

For over 10 years, IPS has been a part of the UK's national strategies for transforming community mental health services. It was formally recognised in the NHS Long Term Plan in 2019[8] and the Five-Year Forward View for Mental Health in 2015[9].

4. The interaction of IPS and MHEP: what is the potential value of a SOP for delivering IPS?

Prior to MHEP's SOPs, IPS delivery in the UK was reportedly patchy and fragmented (Hutchinson, 2022). But since 2018, Social Finance has been supporting DWP and NHS England to build the infrastructure for the world's largest scale-up of the IPS model (Social Finance, 2023) via IPS Grow (for more information read: MHEP's legacy in scaling up IPS in Hulse et al, 2024). Supported by Social Finance's leadership, advocacy, and adaptability, the SOP model that was utilised by MHEP also played an important role:

'SOPs can be used to demonstrate the utility of evidence-informed interventions that previously have not been implemented at scale within the NHS. They can help provide qualitative and quantitative evidence to enable policy champions to convince commissioners of the value of the respective interventions.' – Hulse & Fraser (2024)

[8] For more information: https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf

[9] For more information: https://www.england.nhs.uk/wp-content/uploads/2016/02/Mental-Health-Taskforce-FYFV-final.pdf

The expansion of IPS in England has been 'world-leading', yet challenges have included: poor data quality and flows from the Mental Health Services Data Set (MHSDS), limited investment, contracting arrangements that limit the ability of services, workforce recruitment issues, and insufficient referrals (NHS England, 2023). SOPs could represent a way to improve investment through the use of central outcomes funds and social impact investing and improve data quality through outcomes-focused data analytics and tracking. The SOP's contracting arrangement has been observed to feature supportive collaboration between commissioners and providers. However, evidence generation around the SOP's impact is still ongoing.

Commissioning for services at the intersection of health and employment faces significant challenges, especially in overcoming fragmentation in service provision. There is a keen interest in understanding which forms of commissioning enable effective and efficient services.

This report aims to evaluate the impact of commissioning using a SOP to fund IPS delivery and evaluate whether better employment outcomes can be achieved for those with severe mental illness.

Box 1: Summary of the key findings in the last two reports of MHEP's longitudinal evaluation

The first report focused on theories of change, explored the distinctive contribution of MHEP SOPs, and analysed performance data on the key outcomes metrics through time and across different sites and providers. The first report found that MHEP SOPs provided additional value compared to traditional commissioning via:

- a dedicated performance management function that was seen to drive additional focus on achieving outcomes
- more effective working culture within each local partnership
- identifying and successfully unlocking the LCF funding. This was understood to bring additional financial and human resources to projects.

The second report focused on the implementation experience of the MHEP SOP, including whether the MHEP SOPs affected service quality, provider incentives and legacy for providers and commissioners. The second report found that:

- There was improved accountability and commissioning practice under SOPs;
- MHEP contributed to the national scaling of IPS via Social Finance's advocacy;
- MHEP SOPs' contractual and payment structures were unnecessarily complex and could be simplified via earlier buy-in for design principles,

annual caps, and more realistic expectations on forecasting outcomes performance;

- MHEP brought enhanced capacity to providers by building an IT data system, talent pipelines for staff, and efficient data routines. It brought enhanced capacity to commissioners via experience in partnership working and the creation of a new baseline for expected IPS outcomes; and
- MHEP's incentives on providers were more muted than expected.

Aims and Research Method

5. Aims for the third report

This report is the third and final MHEP report within a longitudinal series (Hulse et al, 2023; Hulse et al, 2024), which is a part of a larger evaluation of the Life Chances Fund (LCF). The aim of the overall evaluation is to develop evidence on the effectiveness of social outcomes partnerships as a commissioning tool to improve social outcomes for vulnerable citizens.

The primary research questions for the three MHEP evaluation reports are:

1. *Did the MHEP Social Outcomes Partnerships make a difference to the social outcomes achieved, compared with alternative commissioning approaches?*
2. *Through which mechanisms did specific aspects of the MHEP Social Outcomes Partnership arrangement contribute to these impacts?*

This study has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee (Reference number: SSH/BSG_C1A-21-1).

6. Data and methods for the third report

Due to data limitations, it was not possible to make a direct comparison of IPS provided through SOPs and IPS provided through traditional grants, although we were able to draw comparison with the IPS literature whenever appropriate. As a second best, we exploited the variation in the typical SOP features' implementation or perception across the different MHEP sites, which could be thought as having different 'intensities' of the SOP approach. This final report tested the intensity of the SOP approach within MHEP and its effect on: 1) outcomes, 2) perceived costs and 3) incentives (and performance management). The intensity of MHEP SOP approach has been investigated by means of:

- Quantitative analysis through a dose-response analysis (see definition on page 22),
- Comparison of actual costs of SOP compared to forecasts
- Comparative analysis based on in-depth interviews and a survey.

Table 1 outlines the data included in this report with regards to 1) dose-response analysis and performance achievement, 2) transaction costs and 3) comparative interviews.

Table 1: Data collected for Report 3

Research Questions	Aims	Data	Sources	Analytical procedures
RQ1	To analyse the performance of outcomes achievement across the MHEP SOPs	Individual-level service user data and employment outcomes (n= 4,176)	Supplied by MHEP SOPs' service providers	Descriptive statistics of success rates, conversion rates and fidelity scores.
RQ1	To test the effect of performance management and incentives on outcomes achievement (in other words, to analyse the dose of a SOP utilised by MHEP).	Individual-level service user data and employment outcomes (n=4,176)	Supplied by MHEP SOPs' service providers	Dose-response analysis using logistic regressions, survival analysis and two-proportion z-tests.
RQ2	To understand the additional time and resources required of a social outcomes partnership aside from the cost of the service.	10 structured interviews + survey	1 provider (4 declined), 1 commissioner (2 declined), 8 Social Finance staff (1 manager, 2 analysts, 2 developers, 1 operational director, 2 operational directors) (3 declined).	Narrative synthesis, cross-referencing the survey responses with interviewee data
RQ2	To analyse the financial resources,	5 qualitative semi-	IPS traditional commissioned	Thematic analysis

	performance management and collaboration in **traditional IPS commissioning** in contrast to social outcomes partnerships.	structured interviews	providers (3) and commissioners (2)	

7. Limitations and considerations

It is important to note the limitations of our evaluation.

First, because of barriers to accessing data for counterfactual or comparator sites (ie sites that delivered IPS through traditional grants at the same time as the MHEP SOPs), we were unable to answer the original research questions through quantitative methods. Our causal investigation of the SOP effect (the difference between SOPs and non-SOPs) is explored predominantly through our qualitative research. We encourage improved access to data to support investigation of alternative commissioning approaches and therefore more robust impact evaluations.

Due to the unavailability of non-SOP project data, it is not possible to determine how the presence or absence of an SOP payment structure (compared with traditional commissioning) impacts outcome performance. However, our previous qualitative research identified hypotheses about how a SOP may generate impact, particularly around the role of performance management and performance incentives. We can hence explore the effect of different SOP dosage levels – that is, whether a higher level of performance incentives (through outcomes-based payments) or stronger performance management increases the likelihood of achieving job outcomes through IPS services.

In particular, in the absence of a counterfactual or non-SOP comparator, one motivation for a dose-response analysis is that a 'dose-response' relationship between exposure and outcomes can support a causal interpretation or test a theoretical prediction (Callaway et al., 2024). In simple words, we used this method to answer the question: 'Do performance management and outcome price impact outcome performance in MHEP SOPs?'

Another motivation for this design is practical: variation in a dose (or exposure) permits the evaluation of treatments for which binary difference-in-difference is either infeasible or undesirable. The data in this evaluation across MHEP SOPs are

‘an attractive source of variation.’ The dose-response is also beneficial as it accommodates for the lack of an agreed conceptualisation of SOP. The approach allows each MHEP SOP to be treated differently, recognising the variation in their structures.

As seen in Table 1, there were 10 structured interviewees from Social Finance and 5 semi-structured interviewees from comparative providers and commissioners. Despite our best attempts to secure a balanced sample for this final round of interviews, many MHEP SOP stakeholders declined to participate in this round, which may indicate research fatigue. Specifically, 1 out of 3 commissioners eligible and 1 out of 4 MHEP provider managers eligible gave perspectives on transaction costs. At least 5 previous commissioners had moved on and didn't have a contactable email. All providers had stopped service delivery as part of this contract and were time-poor for additional evaluation beyond the contract co-commissioned with MHEP. Therefore, per stakeholder group, most interviews were from Social Finance (ex)staff (n=8), providers (n=4; 1= MHEP, 3= non IPS SOPs), commissioners (n=3). Most Social Finance staff were still active on the contract so had time to be interviewed compared to commissioners and providers. Additionally, in terms of team size, there were more Social Finance team members (due to turnover) than provider manager teams and commissioners attached to the MHEP SOPs.

A second challenge limiting conclusions at the project level is that the sample size of dosages is limited to four sites as the Tower Hamlets Learning Disabilities SOP was not comparable to the other four, which focused on mental health. However, there are more substantial sample sizes for service users, which will aid the robustness of the service user-level social outcomes partnership dosage analysis (n= 4,176). This evaluation is based on MHEP stakeholders, so generalisability to other SOPs is limited.

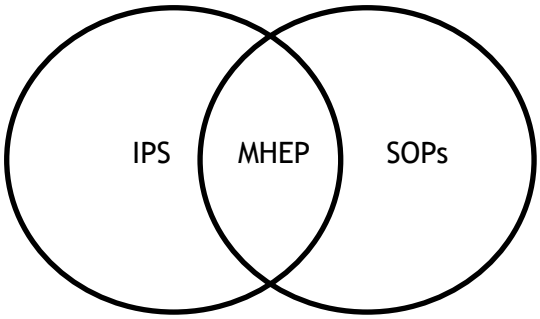
A third limitation is that our analysis relies on the data provided by each of the service providers, who are all VCSE third-sector charities. As in other evaluations, there is a potential risk that unobserved characteristics not captured in the data could influence the outcomes of IPS.

The fact that all services included in the analysis were affected by COVID-19 is a fourth challenge. Service delivery operated from April 2019 (Haringey & Barnet) or April 2020 (Shropshire, Enfield, Tower Hamlets) and experienced shocks to the labour market as a result of COVID-19 policy ie lockdowns. Therefore, recorded performance may not be an accurate representation of total possible performance.

Lastly, we face a difficulty in that our original research questions were to investigate the ‘SOP effect.’ As seen in Figure 4, in order to evaluate the MHEP’s

SOPs, we are interested in the IPS delivery that falls within the SOPs managed by MHEP via Social Finance staff. However, the added value of SOPs that underpin the MHEP projects are difficult to analyse without considering the effect of MHEP as an intermediary/co-commissioner/special purpose vehicle and IPS as an intervention. What interviewees consider distinctive aspects of the MHEP projects may be an effect of the MHEP team members rather than the SOP functions, or vice-versa.

Figure 4: The intersectionality of IPS, MHEP, and the SOP/OBC



Note: IPS=individual placement and support; OBC=outcome-based contract; SOP=social outcomes partnership; MHEP=Mental Health and Employment Partnership. This figure indicates that IPS exists both within the MHEP SOPs and in the NHS, local authorities, OHID and DWP. There are also IPS national-level trials. While MHEP supports 5 place-based SOPs in this evaluation through its SPV, there are 314 SOPs globally as of 31/03/25 and even more outcome-based contracts.

Despite the above limitations, there are three main reasons why MHEP's SOPs remain a robust case for evaluation. Firstly, they deliver an internationally established evidence-based intervention via IPS, which has a well-defined fidelity scale[10]. This differs from other SOPs that test new or 'black box' interventions. Secondly, all of MHEP's SOPs are large projects which have a high number of service users (>10,000 people since 2015) compared with other LCF projects and SOPs pilots internationally. Finally, due to a memorandum of understanding and a collaboration with Social Finance, individual service user-level data was available for analysis across all sites. We were able to discuss with providers what data they held and the quality of that data. This allowed us to construct a comprehensive conceptual model for the evaluation, which increases our confidence in isolating the SOP effect by controlling for confounders and variables that affect IPS performance.

[10] The IPS Fidelity Scale is a prominent part of implementing IPS services. It is a translation of the eight IPS principles into 25 items that can be scored. The IPS fidelity scale is sometimes used in performance management, especially amongst service providers. Formal fidelity reviews are not mandated as part of the MHEP contract, but some fidelity elements are included in the meetings between Social Finance, the commissioner, and the provider.

Findings

8. Were the MHEP SOPs effective in achieving their outcomes targets and how does their effectiveness compare with that of traditionally commissioned contracts?

This section describes the final outcomes achieved by the SOPs' service providers between 2019-2024, which is the period supported by the LCF. Achievement is captured by 3 main metrics:

1. success rates (the percentage of outcomes that were achieved by the end of the project compared to best case scenario)
2. conversion rates (the rate of progress from one outcome to the next on the causal logic)
3. IPS fidelity score comparisons (this score measures the quality of IPS services).

Outcome claims made for sites' metric achievements will also be detailed. For consistency, this section only includes MHEP SOPs supporting service users with mental health disorders (MHEP Haringey and Barnet, MHEP Shropshire, MHEP Enfield, MHEP Tower Hamlets Mental Health). An analysis of the MHEP Tower Hamlets Learning Disabilities' performance against best-case scenario expectations can be found in Appendix C.

Box 2: Key findings of Chapter 1

Key Findings for MHEP SOPs serving users with mental health issues:

- On average, the SOPs achieved 68% of the best-case scenario target figure for service user engagements, 55% for job starts, 30% for job sustainment (13 weeks) working less than 16 hours per week and 55% for job sustainment (13 weeks) working more than 16 hours per week.
- Of engaged service users, 33% went on to start a job, which is in line with the literature on IPS or NHS expectations, despite MHEP operating in part during the COVID-19 disruptions.
- Of service users who started a job, 55% achieved a sustained job outcome (ie full or part-time employment for at least 13 weeks).
- All but one of the MHEP SOPs experienced 'good' fidelity score rankings; Haringey & Barnet achieved 'fair' fidelity.
- The qualitative interviews revealed a difference in the perceptions of providers between traditionally commissioned IPS and MHEP in terms of outcomes, commissioner engagement, reporting of outcomes, flexibility,

pressure to perform and accountability, while the commitment to fidelity is the same for SOP and non-SOPs.

8.1. Success rates

The success rate is the percentage of outcomes that were achieved by the end of the project divided by best case scenario outcomes expectations. These targets indicate the number of outcomes that a project could potentially achieve in a best-case scenario. Social Finance (via MHEP's SOPs) set these projections when they were completing their Grant Baseline Form[11] on receipt of top-up outcome funding from the Life Chances Fund. It should be noted that these targets were set before the COVID-19 disruptions and they aimed to help commissioners set a budget envelope under best-case scenario conditions to pay for the outcomes. Unlike several other projects supported by the LCF, MHEP did not make any formal request to revise these figures.

Table 2 below reports expected best-case scenarios and actual achievements by outcomes metrics for each project. The mean success rate over the 4 MHEP mental health SOPs was:

- 68% for 'Engagement with IPS service' (the individual attends at least three appointments with an IPS employment specialist and a vocation profile is completed)
- 55% for 'Job start individual gains competitive employment' (a service user gets a job)
- 30% for 'Individual sustains job for less than 16 hours per week for 13 weeks'
- 55% for 'Individual sustains job for more than 16 hours per week for 13 weeks.'

Only Shropshire achieved success rates over 100%, achieving 105% for 'Engagement with IPS service.' A success rate exceeding 100% may occur if the service overperforms their targets, individuals are referred to the service multiple times, or more people are referred to the service than initially expected in the best-case scenario. Overall, Shropshire achieved closest to its best-case scenario expectations, achieving 89% (654/738) of its targeted outcomes. The other sites' achievements were: 70% (975/1384) for Haringey and Barnet, 54% (1723/3218) for Tower Hamlets Mental Health, and 48% (400/837) for Enfield. As these success rates may be biased by varying capabilities in making predictions on achievement,

[11] After the award of the contract, LCF projects were required to complete a Grant Baseline Form. In this form, projects reported a set of initial figures on investment commitment, outcome achievement expectations, outcome funding contribution from each outcome payer, initial expectation around costs and savings for commissioners, among other figures. This form was stored in the DCMS Data Portal.

the subsequent site comparison analysis (see Section 3) will investigate the effect of site on job start probability irrespective of best-case scenario expectations.

The reasons for projects not hitting best-case scenario expectations were identified by Social Finance in their LCF end-of-grant forms as including labour market disruptions due to COVID-19, NHS transformation, and staff retention issues which made it difficult to maintain a fully resourced team. Additionally, two providers suffered from referral source issues as new services. One service provider found evidencing job sustainment challenging, which affected their ability to achieve this outcome.

Table 2 – Actual and expected best-case scenario outcomes metric achievement and percentage success rate for mental health cohort sites.

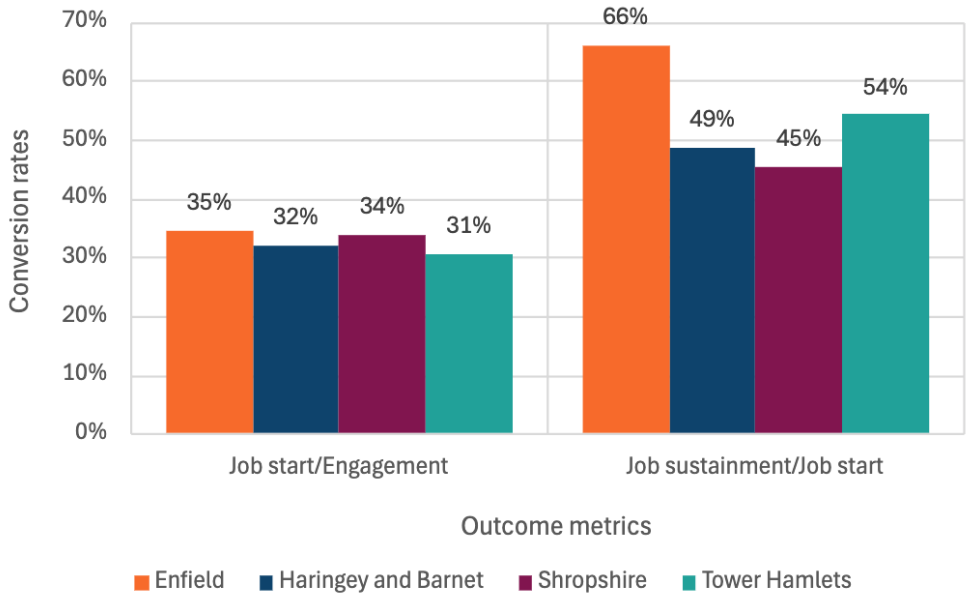
		Engagement	Job start	Job sustainment (<16 hours)	Job sustainment (>16 hours)
Enfield	Best-case scenario expectations	546	181	55	55
	Actual	254	88	16	42
	% Success rate	47%	49%	29%	76%
Haringey and Barnet	Best-case scenario expectations	799	379	113	93
	Actual	660	212	43	60
	% Success rate	83%	56%	38%	65%
Shropshire	Best-case scenario expectations	419	197	66	56
	Actual	439	148	16	51
	% Success rate	105%	75%	24%	91%
Tower Hamlets Mental Health	Best-case scenario expectations	1950	718	248	302
	Actual	1171	358	68	126
	% Success rate	60%	50%	27%	42%
Mean (for mental health related projects)	Best-case scenario expectations	929	369	121	127
	Actual	631	202	36	70
	% Success rate	68%	55%	30%	55%

8.2. Conversion rates

The outcomes conversion rate is the rate at which one type of outcome transitions into the next successive outcome in a causal chain, eg engagement to job start.

Figure 5 depicts the overall conversion rates across the sites.

Figure 5: Outcomes metrics conversion rates



This is calculated by dividing the total site achievement of the successive metric by the total site achievement of the prior metric, and then the rate is expressed as a percentage. The average conversion rate in MHEP's SOPs, which serve clients with mental health disorders, equals 33% for Engagement into Job starts and 53% for Job start into Job sustainment (for Tower Hamlets Learning Disabilities conversion rates, see Appendix C).

This result is broadly in line with IPS literature and NHS expectations. The rate at which engagements convert to job starts (the job outcome rate) is widely used in the employment support literature and therefore facilitates comparison between MHEP's SOPs and existing evidence on IPS effectiveness. Two systematic reviews find employment rates above 40% for IPS programmes targeting severe mental illness (Richter and Hoffmann, 2019; Bond et al., 2012). Given the COVID restrictions and substantial disruptions to the job market, it is not possible to make direct comparison to trial IPS interventions which did not operate under such restrictions. More recently, NHS England (2023) suggests that a new IPS service should be achieving a minimum of 30%-40% of clients into employment/helping them to retain their existing employment. All MHEPs have a total job outcome conversion rate over 30%.

Projects' final conversion rates are slightly higher than the mid-project conversion rates assessed in 2021, which sat at 29% for job outcome rate for severe mental illness (engagements to job start), 30% for Haringey & Barnet, 29% for Tower Hamlets (SMI), 36% for Enfield, and 27% for Shropshire. The highest job outcome

conversions for the severe mental illness cohort were Enfield in 2021 and 2024 (at the end of their contract). Therefore, conversion rates remained steady across the contract.

Enfield had the highest conversion rates on each of the transitions, compared to other sites. This contrasts with the above success rate descriptives, from best-case scenario expectations set prior to the contract, but is in line with Section 3's site comparison analysis.

8.3. Fidelity score

IPS fidelity scores are measurements of the service delivery quality. They allow projects to be evaluated on 25 measurable items (Becker et al., 2019). The IPS fidelity scale is a validated scale that has been used to benchmark other IPS services (Becker et al., 2019). The higher the fidelity score (out of 125 points), the greater the quality of the IPS service, the more closely delivery adheres to the IPS model, and the higher the expected job outcomes. There is no expectation for the fidelity to be formally measured annually; however, a recent quality mark is seen to help with contract tenders. As seen in Table 3, all MHEP SOPs with recorded scores are rated as having at least 'fair' fidelity.

Table 1: Fidelity score classifications

Fidelity score	Fidelity classification
115-125	Exemplary Fidelity
100-114	Good Fidelity
74-99	Fair Fidelity
73 and below	Not Supported Employment

MHEP SOPs' service providers achieved the following fidelity scores:

- 82/125 (66%) in Haringey (2023)
- 97/125 (78%) in Barnet (2023)
- 100/125 (80%) in Enfield (2022)
- 110/125 (88%) in Shropshire (2023)
- 101/125 (81%) in Tower Hamlets Mental Health (2023).

In other words, all but the Haringey & Barnet service experienced 'good' fidelity. We did not receive comparative data from IPS Grow to be able to compare the fidelity scores from MHEP's SOP service delivery with those of other IPS services.

The items that the MHEP SOP service providers had room for improvement on were: zero exclusion (item 9), disclosure of information (item 13), ongoing vocational assessment (item 14), and having a community-based service (item 24).

However, the latest literature states the mean score for fidelity on the IPS-25 item scale reported in the UK is **102** (Waghorn et al 2018). Other average scores found in the literature fall within the range of 92–110[12]. Therefore, the fidelity of the MHEP SOP service delivery is similar to and/or higher than the UK average reported in the literature, with the exception of Haringey.

Sites that had IPS services prior to MHEP (Shropshire and Tower Hamlets) had higher fidelity scores in their previous IPS iterations. Several reasons may explain this. Firstly, fidelity scores were measured by a different professional body (Centre for Mental Health up to 2019 and IPS Grow from 2020) (IPS Grow, 2015). Additionally, COVID-19 caused a drop in fidelity as some key fidelity items could not be fulfilled due to the COVID-19 government restrictions. For example, item 24, 'being in the community', and item 17, 'face-to-face employer engagement', were not possible due to lockdowns.

Moreover, we were only able to source one fidelity score during each of the projects' delivery of the SOP. Thus, the scores are not reflective of the service delivery quality throughout the SOP. Due to the high degree of missing data in the available fidelity scores, this measure will not be included in the subsequent dose-response analysis.

According to the qualitative interviews conducted for the previous report, high service quality under the MHEP SOPs was observed in four factors: 1) more rigorous caseload management, 2) more emphasis on integration with clinical teams, 3) greater attention to a wider range of outcomes, and 4) continuous discussions on fidelity and service quality (Hulse et al, 2024). But the MHEP SOPs' commitment to fidelity and their focus on outcomes sometimes came into conflict because formal assessments of fidelity occupy so much of the staff's time, reducing providers' ability to achieve outcomes. During a fidelity review period, it is not possible for a service to operate at full capacity (Hulse et al, 2024). Although this is an issue for all IPS services, in MHEP this time pressure had a knock-on effect for outcome-based payment levels and income for providers. For instance, providers would prepare for a loss of that outcome-based payments during that fidelity period (Hulse et al, 2024).

[12] An average of 92.2 in selected outpatient programmes in New York State (Margolies et al, 2018), 90.43 in mental health in Netherlands (Roeg et al 2021), 98.16 in a RCT in Norway (Fyhn et al 2020), and 100-110 in Canada (Erickson 2021) (Poresmski, 2017).

Table 2: Available fidelity score percentages achieved by sites for the year in which the fidelity score was measured

	Previous IPS service delivery	LCF MHEP service delivery
Enfield	n/a (no service delivery)	80% (2020-2023)
Haringey	66% (2019)	66% (2020-2023)
Barnet	78% (2019)	78% (2020-2023)
Shropshire	94% (2017-2019)	88% (2020-2023)
Tower Hamlets (MH)	78% (2017); 92% (2018-2019)	81% (2020-2023)

8.4. Outcomes claimed

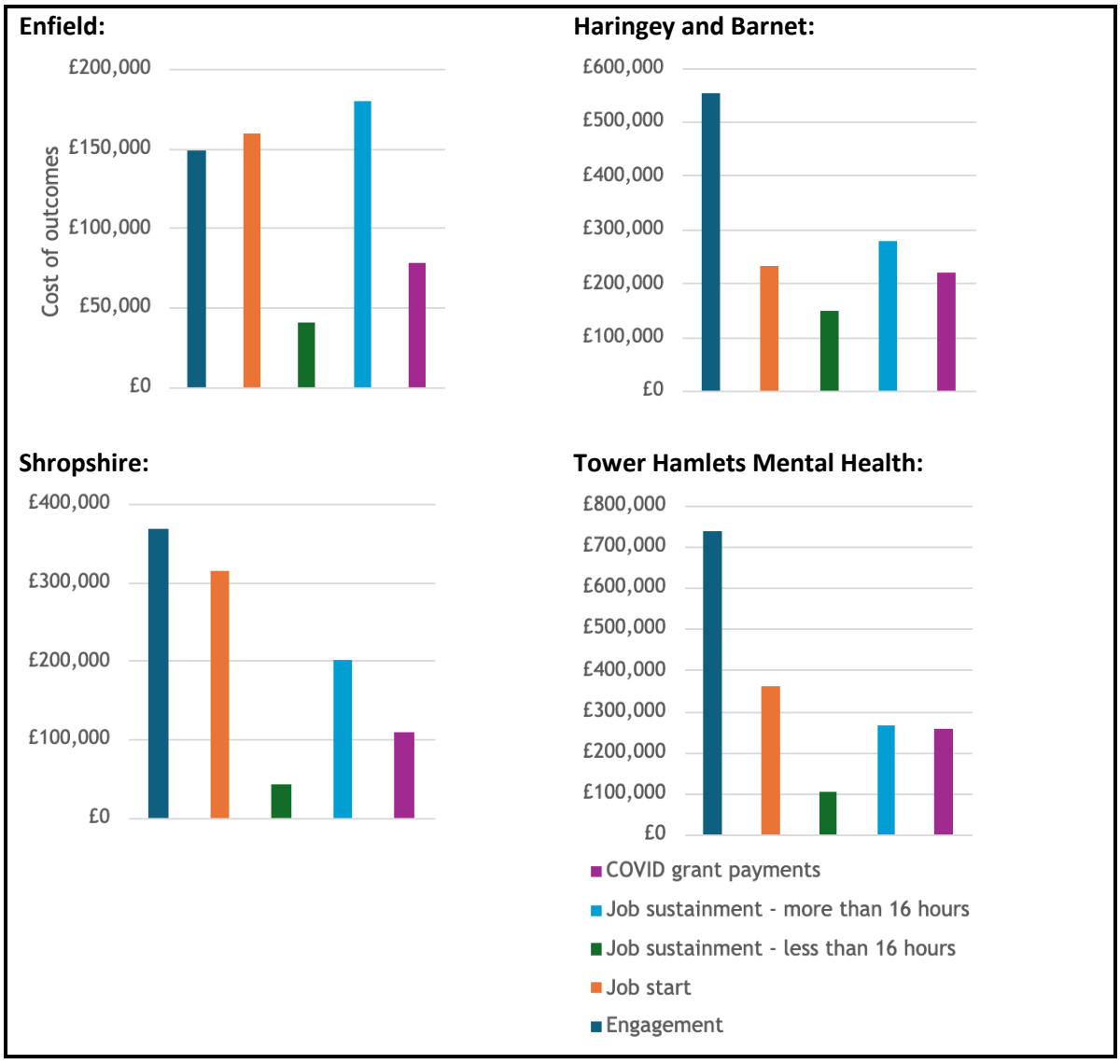
To provide further insight into the performance of the MHEP SOPs, we illustrate outcomes claims. Specifically, Figure 6 depicts outcomes claimed by fiscal year whilst Figure 7 and Figure 8 illustrate total outcomes payments by outcome metric and by fiscal year, respectively.

Figure 6 demonstrates that the second fiscal year of service delivery was the highest for outcomes claimed across all sites except for Enfield, which had its greatest outcomes claims in its last fiscal year. This is despite the substantial staff turnover issues that Enfield reported in its last fiscal year of delivery.

Figure 6: Number of outcomes claimed by fiscal year

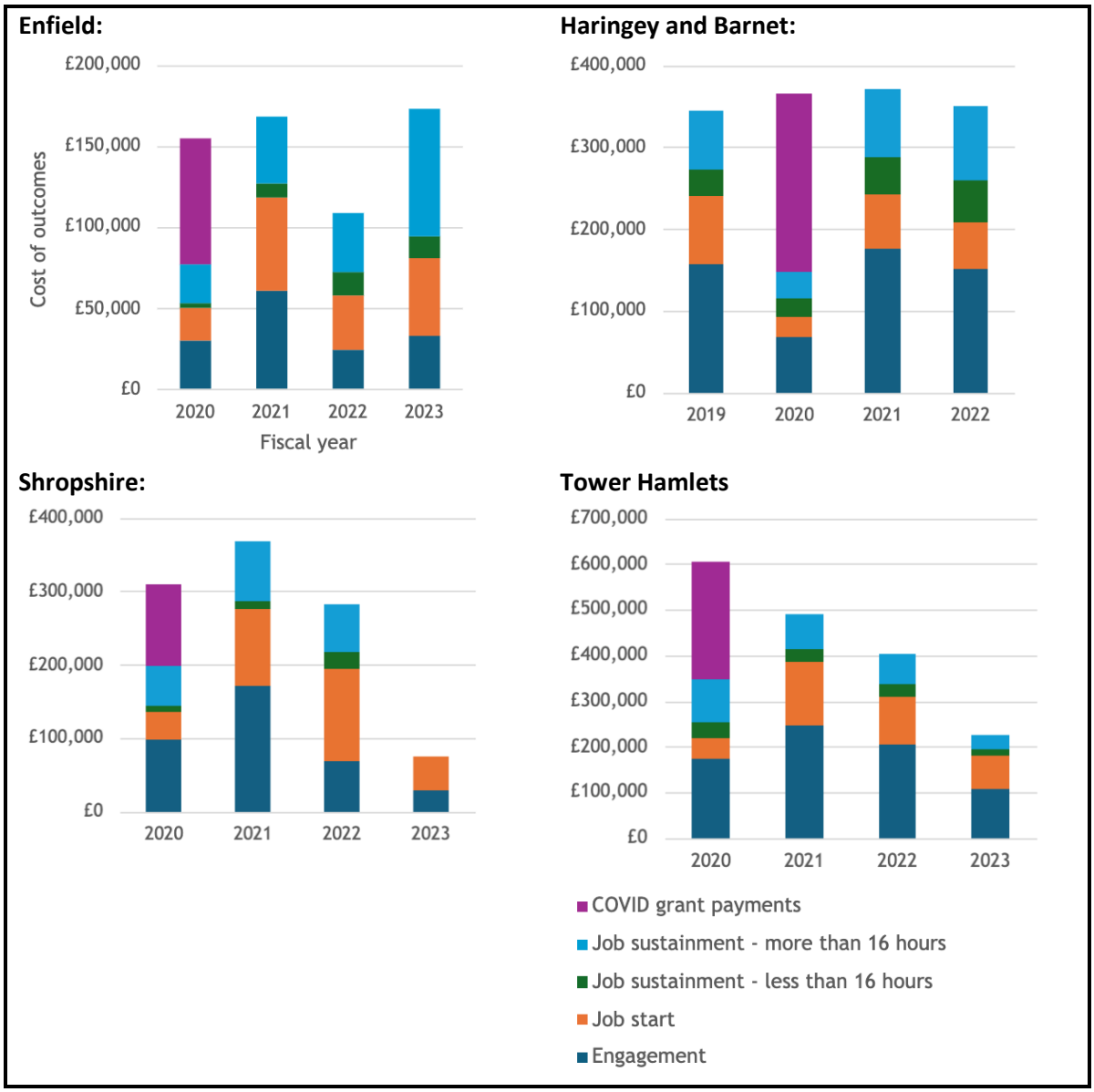
Figure 7 shows that the greatest outcomes payments claimed for an outcome metric were for 'Engagement' in all sites except for Enfield, which had the highest outcomes payments for 'Job sustainment – more than 16 hours.' On the other hand, 'Job sustainment – less than 16 hours' was consistently the outcome with the smallest total outcome payments across sites. For all sites, this was also the metric with the lowest success rate (%) compared to all other metrics – as seen in Figure 7.

Figure 7: Outcomes payments



As depicted in Figure 8, total outcomes payments varied by fiscal year and sites. The figure indicates that COVID grant payments were only given over the fiscal year of 2020. These pandemic grants were paid by LCF and local commissioners and were based on medium-case scenario plans for site outcome achievements. Performance incentive variation and its potential effects on outcome achievement will be explored further in the subsequent dose-response analysis.

Figure 8: Cost of outcomes by fiscal year



8.5. How does this compare with traditional contracts?

In our comparative interviews for this report, commissioners and providers from traditional IPS contracts revealed their perceptions of fidelity, outcomes, reporting, and accountability for the success of IPS.

Outcomes and commissioner engagement are different

IPS providers in traditional contracts are typically working towards fidelity scores and access targets for their main KPIs (alongside job start outcomes); in contrast, MHEP SOP providers worked towards fidelity scores, service engagement outcomes, job start outcomes and job sustainment outcomes.

Providers in traditional commissioned contracts felt little performance management beyond the service managers. Despite having communication with IPS Grow and local Integrated Care Boards, interviewees felt pressure because *'the management of IPS effectively stops at us.'* Service managers under the SOP felt that MHEP, when added as a co-commissioner, served as a provider's advocate and a bridge between the commissioner and themselves. This collaboration between the provider and MHEP performance management team meant that providers did not feel isolated as they typically do with traditional commissioners. This perception was also emphasised in a provider interview from Phase 2 of MHEP's SOPs: *'[T]here is the recognition that you want, but also again having the regular catch ups puts further rigour into what you do and you're not feeling like you're saying "We're over here." You are part of a team [with MHEP], you are part of something. You're not that Secret Service I described earlier on really [in traditional commissioning].'*

All IPS providers, both SOP and non-SOP, commented that traditional commissioners have other priorities that compete with IPS delivery. Providers lamented the drop off in commissioner engagement after IPS's first introduction in the UK: *'[P]erhaps when it was first bought in because it was new and it was getting this new investment, there was a focus on it. But there's something else which takes priority and actually, IPS then just becomes another performance KPI on another dashboard.'* This is similar to a previous interview from a MHEP SOP provider: *'[A] commissioning manager might have loads of contracts on their books, might not have as much time to particularly focus in on the performance of a particular provider. They might not meet with them as frequently as we do [with MHEP].'*

While MHEP became the middleperson between the commissioner and providers within the SOP, prior to 2019 no actor would provide that role for traditional contracts. A traditionally commissioned provider stated that prior to IPS expansion, the relationship between commissioner and provider was a lot closer. However, as IPS expanded, commissioners' and providers' relationships and their frequency of engagement decreased. Interviewees described MHEP as stepping into that gap. Interviewees predicted that the provider-commissioner relationship would not improve under a SOP within a three-to-four-year contract but agreed that they liked the type of closer relationship observed under a SOP.

However, after 2019, as a result of the experience of MHEP, Social Finance brought the learnings from the SOPs into a separate enterprise, IPS Grow, which would serve all IPS providers. The comparative interviewees had never heard of MHEP, but did know of IPS Grow. In fact, the traditionally commissioned providers stated that IPS Grow was the middleperson between them and their commissioner: *'I don't directly liaise with NHS England. It's IPS [G]row as the middle-person that tells us what our access target is, what our funding is going to be and that's who we report to on a monthly basis.'*

While this brought in additional funding for Social Finance in the IPS sector (won a tender with NHS), the effectiveness of IPS Grow suggests that the data derived from a SOP as a result of outcome-reporting and close engagement can produce learnings. This innovation from the SOP experience under MHEP suggests that all IPS providers can experience support and accountability similar to those experienced under a SOP.

Commitment to fidelity is the same

According to the traditionally commissioned provider interviewed for this report, high fidelity is also a priority for them. All of the providers (regardless of contracting type) interviewed believed that job outcomes would follow high fidelity. Traditionally commissioned providers and IPS Grow (previously interviewed) stated that achieving a quality mark was a strong incentive and motivator. A quality mark is achieved through a minimum conversion rate of 30% from accessing a service to the start of employment. Recently, this threshold has now increased to a 40% conversion rate. One traditional commissioned provider stated the value of fidelity in keeping accountable to service quality: *'[T]he fidelity reviews have been extremely valuable from IPS [G]row. People don't listen to you as much as an external body being independent. At the end of the day, having someone else to be accountable to put us under pressure. We don't sit there thinking it's all OK and that we're not judged or we're not being measured. We are and it gives us accountability.'*

Reporting of outcomes and flexibility is different

Instead of submitting KPIs and outcomes to MHEP, traditionally commissioned providers report directly and regularly to IPS Grow. Providers stated that they did not report regularly to their commissioner; their only contact was through data reporting on the RIO platform to the NHS dashboard via the Mental Health Services Data Set (MHSDS). Similar to MHEP SOP providers, traditionally commissioned providers use this reporting as a tool to manage expectations and to performance manage their frontline staff (employment specialists).

Nevertheless, what differs between providers who were SOP-commissioned under MHEP and traditionally commissioned is flexibility with outcomes. Unlike providers under traditional contracts, MHEP SOP providers review their outcomes achieved biquarterly with their commissioner and MHEP. This greatly differs in traditional commissioning. One traditionally commissioned provider explains: *'[W]ith the outcomes that [are] with us looking at those, but it's going to be [a] yearly target instead of per month because we know that sometimes some months are great for outcomes, some months are not. Sometimes it depends on their client, on the economy, or what the jobs are out there. So we give that little bit of flexibility to our staff members.'* Another traditional commissioned provider explained his

aversion to results-based financing due to this perceived inflexibility: *'I worked under RBF before (Welfare to Work and DWP). Longer-term targets are more beneficial, and focus on access over job outcomes has allowed them more flexibility in this traditional commissioned service.'*

All IPS providers (regardless of their contracting) have had to make some changes to their IT systems so they can capture IPS specific information. The perception of the service managers is that the frontline staff should be able to self-monitor. Traditional commissioners' and MHEP's ways of supporting that are obviously different. However, what is needed for any contract, according to interviewees, are technical assistance (fidelity reviews, IPS specialist), resources for maintaining good staff, support for performance management (IT systems, process to regular review) and expectation management. This view is supported by both the interviewees and the academic literature.

Pressure to perform and accountability are different

According to our interviewees, traditional commissioned IPS providers still feel the pressure to perform in order to win future contracts. However, this pressure is substantially less than that felt by those in a SOP. One service manager described it as pressure that operates over five years rather than quarterly pressure from outcome payments in a SOP. For instance: *'If we don't achieve, we still get the money but it's not grant funding. In that sense, obviously you have to perform because that's what's then going to dictate the future of the service. That adds a degree of pressure that you almost feel we've got the money, we need to achieve it but if we missed the target for this year, we're not going to not exist. But in a five-year period if every service misses it, then potentially the funding will reduce or fall back. And I think that what's interesting.'*

Traditionally commissioned providers stated that they feel more pressure from IPS Grow but feel more accountable to their commissioner since they 'pay the bills.' This was acknowledged that there was more pressure from IPS Grow since they are external and have more vested interest if IPS works. In comparison, if IPS does not work, commissioners can just take it off their list of procured services, so it is suggested they feel less accountable for the success of IPS. For instance: *'[T]o be honest, the commissioners might only email us like once a year. You normally get something from them in like March or April, from end of the financial year to beginning of the financial year. As IPS [G]row shows interest all year round of our performance.'*

Unlike providers delivering IPS under MHEP SOPs, traditional IPS providers stated that there was a lack of consequences for underperformance. Interviewees were concerned that in traditional commissioning, poorly performing services that are not held accountable can put the whole IPS sector at risk. This was described as:

'[H]opefully I think having the funders to have the balls to pull the plug on services that aren't doing well. The fact that services seem to be able to go quite a long time not doing well before they're reviewed or held accountable. Some services might have some context, maybe they haven't been given the money in time and it's unrealistic and but they're still working hard and practicing IPS; services like that should be supported. But when you've got services that are just not doing IPS for six months a year, resisting and looking for excuses rather than solutions, that's where it puts the whole IPS community at risk.' This greatly differed from the previous reports' finding that SOPs emboldened accountability of IPS delivery due to more levers for underperformance, clarity of monitoring, high involvement and problem solving (see 'Is there greater accountability in SIB-funded IPS compared to traditional commissioning' in Hulse et al. 2024).

	Non-SOP IPS delivery from **Phase 3 interviews**	**SOP IPS Delivery in MHEP from previous interviews**
Commissioner engagement	Not meeting commissioners frequently.	More meetings, meeting at least biquarterly.
Payment	Commissioners set up the purchase order and send the provider money regardless of performance.	In some ways, being in the SOP creates extra work for commissioners because the amount of money that they're paying out each period is different since it depends on performance.
Reporting of Outcomes	Submitting KPIs and outcomes to IPS Grow and the Mental Health Services Dataset regularly. Commissioners typically do not flag underperformance until the end of a contract and do not have levers to hold providers accountable.	Submitting KPIs and outcomes to MHEP, IPS Grow and the Mental Health Services Dataset regularly. Commissioners are able to flag underperformance more quickly and have levers to hold providers accountable.
Outcome engagement	Still have targets, but fewer than a SOP has.	Have more outcomes to track; this increases administrative burden but less for the provider and more substantially for the intermediary (in chasing different invoices from performance, extra reporting for board, effort to validate outcomes with external body).

The evolution of IPS commissioning in the UK, and why do MHEP SOPs differ?

The early implementation of IPS in the UK, led by Miles Rinaldi and Rachel Perkins, was characterised by close, collaborative relationships with commissioners, often from local authorities. These relationships included regular face-to-face interactions, monthly reporting and strategic discussions that provided both context and support. Data reporting was manual and paper-based, but commissioners were actively involved and invested in the programme's success, as IPS had yet to establish its evidence base in the UK.

However, the rapid and large-scale expansion of IPS – particularly from 2019 onward – has significantly altered the commissioning landscape. The scale-up brought operational challenges and introduced distance between service providers and commissioners. Today, providers interviewed report only minimal engagement with commissioners, often reduced to an annual email, while relationships are now largely mediated through IPS Grow, a national support and fidelity body. IPS Grow is seen as having a more sustained and vested interest in IPS outcomes, acting as an external source of accountability, especially in the absence of regular commissioner oversight.

This shift has contributed to fragmentation: service delivery is positioned between IPS Grow and commissioners, both of whom may operate at different speeds and with differing priorities. Providers describe feeling more pressure from IPS Grow, yet they perceive greater financial accountability to commissioners, who ultimately 'pay the bills.' The lack of clarity and consistency from traditional non-SOP commissioners regarding performance expectations and consequences for underperformance creates uncertainty, which can negatively affect service culture, staff morale and strategic planning.

The current IPS commissioning contrasts with MHEP, where contractual clarity, long-term funding and stable reporting frameworks contributed to improved provider's perceptions of service performance and morale. According to interviewees, today's block contracts often lack outcome-linked incentives or penalties, which can result in underperforming services persisting without sufficient scrutiny. Conversely, excessive reliance on payment-by-results models (and more risk against the provider) risks promoting superficial performance over service quality.

Overall, the IPS provider interviewees called for a more integrated, transparent commissioning approach that balances supportive accountability with operational flexibility. A renewed emphasis on stable, long-term planning and psychologically safe environments is seen as vital to sustaining quality and workforce engagement amidst ongoing IPS expansion.

9. Did the SOP have higher costs than expected and if so, why?

This section explores the costs associated with the MHEP SOP across all five sites.

SOPs differ from standard contracting and grants processes because they have more intensive setup, monitoring and evaluation costs. These are 'transaction costs': the costs incurred in delivering the service beyond the cost of the service itself.

This analysis aims to understand the monetisable and non-monetisable transaction costs associated with the MHEP SOPs, and whether these were higher or lower than expected. There was no data available on comparable non-SOP IPS service delivery to directly compare against MHEP projects by costs or outcomes. Therefore, this analysis instead looks at whether transaction costs for the MHEP projects were higher or lower than originally expected, and why, based on quantitative and qualitative data collected over the lifetime of these five projects.

Box 3: Key findings: Did the five MHEP SOPs overspend/underspend compared to forecasts? What were the key drivers of overspend or underspend?

For all MHEP SOPs, management costs turned out higher than anticipated and investment costs lower than anticipated.

MHEP directors interviewed stated that setup costs and time were equal for SOPs that eventually signed contracts and for SOPs that fell through at the last minute, resulting in no contract signed. It was estimated that at least 5 additional potential SOPs fell through in the final stages.

Searching and negotiating with commissioners was described as a 'circular' process rather than a straightforward back-and-forth. It took about **nine months**, involving searching, agreeing on structures and co-developing specifications. The modelling of outcomes values to ensure financial viability took up about 80% of the search and negotiation stage. Once this was resolved, the contracting proceeded more quickly. Delays often occurred due to approval processes, procurement timelines, and staff turnover. For instance, external approvals lags (eg LCF approval processes) and commissioner role changes meant that several commissioners rescinded their interest. MHEP developers expressed this frustration, saying that the '*centrally offered top-up funding model doesn't really sink in with local Commissioner contract timing*' and suggesting that addressing this issue could reduce costs.

The key factors leading to costs being higher than anticipated were: COVID-19, higher need for managing underperformance, service underperformance to targets, and delayed start of delivery.

The key factors leading to costs being lower than anticipated were early repayments of loans and thus reductions in total interest. Higher engagement and sustainment compared with predictions also led to financial gains from outcome payments in Shropshire and Tower Hamlets Learning Disabilities.

Three phases in the MHEP SOPs' contract turned out to be associated with the highest transaction costs: the setup, moments when performance differed from expectations, and completion. Social Finance interviewees from this report revealed that factors which increased setup effort were upskilling commissioners and providers on SOPs and misalignment in contract timings between actors or priorities, causing delays in negotiation. Factors which increased all other transaction costs were the complexity and unfamiliarity of SOPs.

Efficiency, standardisation, the use of an SPV, manualised service specification, did decrease time and effort in monitoring the SOP, but the procurement of the SOP was not perceived to be more burdensome than that of a traditional contract. A perceived administrative burden in monitoring requirements of the SOP was mixed according to those interviewed.

Social Finance highlighted that the end-of-grant reporting for an outcomes fund requires reconciling original projections of outcomes and finances of a project. This involved returning to five-year-old documentation often produced by a member of staff who was no longer at Social Finance. This challenge may reflect the importance of appropriately managing documentation of baseline expectations of a SOP when applying to an outcomes fund like LCF. These reconciliations required significant effort, exacerbated by the presence of caps and LCF engagement requirements.

Intermediaries (and other investment fund managers in the social investment market) often expressed criticism of rigidity over caps and argued for caps to be raised or to be considered over the duration of the projects rather than on an annual basis.

9.1. Monetisable costs: forecast vs. actuals

Monetisable costs are costs which have a direct monetary value attached to them. This section assesses whether the monetisable costs for each MHEP SOP were higher or lower than originally forecast. MHEP SOPs' actual costs were estimated to be £6,160,973, inclusive of costs for investment, management, evaluation and learning, and delivery as defined above. These costs were lower than expected for investment, delivery, evaluation and learning, but higher for management than expected by 25.8%.

Box 4: Key monetisable costs in a SOP

Monetisable costs

Management costs included the cost of items such as coordination and oversight personnel, performance management systems, financial management systems, resources spent on governance discussions and partnership building, etc.

Investment costs included the cost of items such as setting up and maintaining a Special Purpose Vehicle (SPV), expenses related with setting up the investment, the return to social investors, etc.

Investment return is a ratio that measures the profitability of an investment. This is typically the ratio of the income from the investment over the cost of the investment. A positive return indicates that a social investor has made a profit, while a negative return indicates that the social investor lost (some of) their capital.

Evaluation & Learning costs included the cost of contracting the services of an external evaluator and any other learning cost associated with internal learning activities.

Delivery costs included the cost of all items related to the implementation of the intervention, such as cost of frontline personnel, special material and licenses to deliver the programme, training costs, etc.

Table 3: Difference in actual versus original baseline of key monetisable cost and savings, and final actual costs in the MHEP SOPs.

	Enfield	Haringey & Barnet	Shropshire	THSMI	THLD	TOTAL
LCF funding: baseline	£260,689	£596,918	£434,484	£868,966	£589,673	£2,750,730
LCF funding: actual	£254,072	£596,914	£434,483	£720,452	£522,220	£2,528,141
Difference	-£6,617	-£4	-£1	-£148,514	-£67,453	-£222,589
Percentage difference	-3%	0.0%	0.0%	17.1%	-11.4%	-8.1%
Investment: baseline	£126,000	£227,000	£204,000	£300,000	£414,000	£1,271,000
Investment: actual	£126,000	£227,000	£204,000	£300,000	£328,000	£1,185,000
Difference	£0	£0	£0	£0	-£86,000	-£86,000
Percentage difference	0%	0.0%	0.0%	0.0%	-20.8%	-6.8%
Investment cost: baseline	£58,046	£250,909	£102,000	£162,099	£183,261	£756,315
Investment cost: actual	£55,306	£201,208	£60,995	£113,559	£76,293	£507,361
Difference	-£2,740	-£49,701	-£41,005	-£48,540	-£106,968	-£248,954

Percentage difference	-5%	-19.8%	-40.2%	-29.9%	-58.4%	-32.9%
Management cost: baseline	£78,084	£98,715	£99,803	£98,763	£137,159	£512,524
Actual	£120,554	£92,115	£131,842	£131,442	£168,557	£644,510
Difference	£42,470	-£6,600	£32,039	£32,679	£31,398	£131,986
Percentage difference	54%	-6.7%	32.1%	33.1%	22.9%	25.8%
Delivery cost: baseline	£475,200	£1,071,611	£810,000	£1,632,000	£1,058,604	£5,047,415
Delivery cost: actual	£450,916	£1,094,907	£810,000	£1,591,200	£985,523	£4,932,546
Difference	-£24,284	£23,296	£0	-£40,800	-£73,081	-£114,869
Percentage difference	-5%	2.2%	0.0%	-2.5%	-6.9%	-2.3%
Evaluation & learning cost: baseline	£9,360	£19,200	£21,600	£20,160	£24,960	£95,280
Evaluation & learning cost: actual	£4,222	£19,584	£15,984	£15,984	£20,782	£76,556
Difference	-£5,138	£384	-£5,616	-£4,176	-£4,178	-£18,724
Percentage difference	-54.9%	2.0%	-26.0%	-20.7%	-16.7%	-19.7%

Note: These statistics have been collected from the DCMS data portal (a platform for all LCF projects) for the original baseline and from End of Grant forms/DCMS data portal for the actuals. This is because the End of Grant forms did not consistently cover the original baseline, whereas the actuals were covered and validated by both sources. Please check the Glossary to find definitions for these terms. THSMI=Tower Hamlets Severe Mental Illness and THLD=Tower Hamlets Learning Disabilities. THSMI and THLD are two different service delivery sites and SOPs both based in Tower Hamlets, London.

Drivers of differences in monetisable costs

- **Higher performance support demands:** two sites (Shropshire and Enfield) were subjected to a formal performance improvement plan and one was subjected to an informal improvement plan, requiring the intermediary to invest more time reporting and modelling the financial implications and invest in more frequent catch-ups – see the previous report Appendix B in Hulse et al., (2024).
- **Investment costs were lower than expected** by 32.9%. Investment cost was lower than expected in every MHEP SOP due to early repayment of loans compared to predictions and a resulting reduction in interest cost. This early repayment was negotiated between the MHEP board, investor and Social Finance to help reduce future admin burden of financial planning when all SOPs were completed and maintaining the investor's preferred rate of return.
- Big Issue Invest, the social investor for MHEP SOPs, invested the same amount of funding as originally planned for Enfield (£126,000), Shropshire (£204,000), Tower Hamlets Mental Health (£300,000), and Haringey & Barnet (£227,000). BII invested less in Tower Hamlets Learning Disabilities SOP than originally planned (£328,000 compared to £414,000 planned) but

this was due to a delayed start. The return on the investment (ROI)[13] was negative for Enfield (-1.30%, compared with a planned 8.30%) and for Tower Hamlets Mental Health (-6.80% compared with a planned 18.10%), while positive for Shropshire (8.80% compared with a planned 9.80%), Tower Hamlets Learning Disabilities (7.40% compared with a planned 11.49%), and Haringey & Barnet (25.51% compared with a planned 18.50%). The agreements with BII allowed MHEP to pool gains and losses across projects (including MHEP projects not supported by the LCF), creating a risk pooling effect at the portfolio level. Overall, the social investor incurred a gain in line with expectations.

Table 6: Investment costs and return

	Enfield	Haringey & Barnet	Shropshire	THSMI	THLD	Total
Forecast loan from investor	£126,000	£227,000	£204,000	£300,000	£414,000	£1,271,000
Actual loan from investor	£126,000	£227,000	£204,000	£300,000	£328,000	£1,185,000
Investment cost	£55,306	£201,208	£60,995	113,559	76,293	£507,361
Forecast ROI	8.30%	18.50%	9.80%	18.10%	11.49%	
Actual ROI	-1.30%	25.51%	8.80%	-6.80%	7.40%	

Note: the 'actual loan from investor total' is from Haringey & Barnet, Shropshire, Enfield, Tower Hamlets Severe Mental Illness (THSMI), and Tower Hamlets Learning Disabilities (THLD) SOPs under the Life Chances Fund outcomes fund. This does not include the additional loan amounts from Big Issue Invest under the West London's Alliance SOP (£400,000) and under CBO/SOF outcomes funds (£360,000). The total investment loan from the investor, Big Issue Invest, was 1.945 million for all nine of the MHEP SOPs. As stated in the Introduction, these loans are unsecured loans, meaning capital is at risk if services underperform. Also, as BII could pool risk across projects, gains in one project could be used to compensate losses on another.

Drivers of differences in monetisable revenue (outcomes payments)

- **COVID-19:** initial revenue loss was seen due to the suspension of outcomes payments during April 2020–October 2020 and job market disruption because

[13] Investment return can be defined as a ratio that measures the profitability of an investment. It is typically the ratio of the income from the investment over the cost of the investment. A positive return indicates that a social investor has made a profit, while a negative return indicates that the social investor lost their capital. For example, a project that reported an initial investment of £1,000,000 and return of £1,200,000 has achieved an investment return of 20%.

of lockdowns. All MHEP SOPs suspended outcome payments during this period. When the first lockdown in the UK began in March 2020, Haringey and Barnet was the only MHEP LCF project to have launched. It had done so under the original outcome payment tariffs (known as 'Type 1 tariffs'). By April, the Haringey and Barnet MHEP project elected to shift to grant payments based on medium performance forecasts (known as 'medium scenario payments'). The other four MHEP projects launched in April 2020, also on medium scenario payments. By October 2020, all five projects shifted back to outcome-linked payment, as planned.

- **Forecasting uncertainty:** in the preliminary forecasts, there was no awareness of the cohort variances between severe mental illness and learning disabilities, which compounded the fact that outcomes were lower than expected in both cohorts due to referral issues resulting from NHS transformation.
- **Delayed start of delivery:** Tower Hamlets delivery was delayed by three months due to delays in contract negotiations.
- **Annual cap[14] challenges:** the rigidity of annual caps made it challenging for the project to recuperate revenue across periods of fluctuating performance.
  On the other hand, financial gains were identified from higher outcomes than expected due to nearly double the expected number of clients sustaining employment (for Tower Hamlet Learning Disabilities) and higher numbers of engagements and sustainments (for Shropshire).

Uncertainty and forecasting

It is innately difficult to forecast IPS project outcome achievements for both SOP and non-SOP projects. The original forecast for MHEP projects costs and outcomes faced several challenges. Firstly, predicting outcomes in social care and for a severe mental illness and learning disabilities cohort remains incredibly challenging. Secondly, ongoing national issues with retaining staff in IPS services continues to be a risk for IPS projects. This is a national issue regardless of the IPS contracting type. Thirdly, over the course of the MHEP projects, The Five Year Forward view set out a vision of the NHS moving from fragmented individual institutions to more collaborative, place-based, local health and care systems. This meant moving from a CCG model of NHS to Integrated Care Systems. This change greatly affected referral pathways, impacting the local care pathways and the IPS

[14] SOP commissioners in the UK often put a cap on the maximum amount of outcomes payments they are willing to make, which can also cap the investor's maximum return. Cap attempt to reduce uncertainties and risks in terms of financial commitments; however, their use (and consequences) is often debated. MHEP SOPs also have a cap.

services. Lastly, forecasts for all Life Chances Projects had not factored in the COVID-19 pandemic.

9.2. Non-monetisable costs

Many of the transaction costs associated with the MHEP SOPs did not have a monetary value, or it was difficult to explicitly link the transaction cost to a monetary cost. Non-monetisable costs are costs which do not have a direct monetary value attached to them, such as time spent setting up a service or complexity of a SOP.

Drivers of higher non-monetisable transaction costs

Three phases in the MHEP SOPs' contracts turned out to be associated with the highest non-monetisable transaction costs: the setup, the end of the SOP, and when performance differed from expectations.

Contractual setup

According to Social Finance, there was an enormous amount of iteration that occurred in understanding the space of health and employment before Social Finance landed on the exact project of a SOP delivering IPS. The setup required an initial landscape review of the key challenges faced by people with health issues and employment in order to understand the needs of the sector and the value the SOP could provide. The fragmentation of IPS funding and lower IPS outcomes achieved in the UK compared with trial evidence promoted further investigation.

Once the intervention was chosen, MHEP staff spent a large amount of time engaging/searching for commissioners before being able to discuss SOP contracting. Engagement often involved educating commissioners on outcomes-based contracting and aligning their budget and procurement cycles with SOP requirements. Many initial conversations did not lead to contracts, resulting in high upfront costs. In fact, MHEP directors interviewed stated that setup time and therefore costs were equal for SOPs that eventually signed contracts and for the SOPs where contracts were not signed because they fell through at the last minute. It was estimated that at least 5 additional SOPs fell through in the late stages of contracting.

Some commissioners found it difficult to accept that a portion of their limited funds would be allocated to investor returns rather than directly to frontline services, despite the promise of better outcomes and more effective services. A MHEP director stated that locally led SOPs are rare in the social investment market, partly due to this financial concern.

Searching and negotiating with commissioners was described as a 'circular' process rather than a straightforward back-and-forth. It took about **nine months**, involving

searching, agreeing on structures and co-developing specifications. The modelling of outcomes values to ensure financial viability took up about 80% of the search and negotiation stage. Once this was resolved, the contracting proceeded more quickly. The MHEP development director explains: '*[S]o balancing and wrangling the modelling to determine what the outcomes value could be, and agreeing with Commissioners in order for top-ups to be approved by LCF, which would then be financially viable to take through the investment committee was by far and large the biggest time and effort.*'

Some contracts faced significant hurdles, such as commissioners refusing to link payments to specific outcomes or requesting unconventional measures (eg A&E spending reductions). Delays often occurred due to approval processes, procurement timelines and staff turnover. For instance, external approvals lags (eg LCF approval processes) and Commissioner role changes meant that several commissioners rescinded their interest. MHEP developers expressed this frustration, saying that the '*centrally offered top-up funding model doesn't really sink in with local Commissioner contract timing*' and suggesting that addressing this could reduce costs.

However, according to multiple Social Finance respondents from the survey, setup costs of finding a commissioner were reportedly substantially lower in LCF than in the earlier Commissioning Better Outcome Fund due to the estimated search time reduction: in LCF, there was no need to spend time learning the ideal characteristics of a commissioner under a SOP and methods to engage them. A large part of this was that SOPs were a more established model in 2018/2019 compared to 2015/2016, with more people interested due to the number of SOPs operating in the UK and the commitment to social innovation in local areas.

The first full year of MHEP SOP delivery revealed gaps in how outcomes and processes (eg financial reconciliations, outcomes payments) were designed and implemented, requiring additional effort to improve design and implementation. Examples include mismatches in financial expectations and miscommunications with stakeholders.

Box 1: Recommendations for Outcome Funders of SOPs

Increasing Setup Effort:

- **Misalignment:** local Commissioner contract timings and centrally procured top-up funding timelines may defer and delay negotiations. Delays in approvals caused commissioners to drop out or change priorities. Aligning diverse stakeholder requirements (Commissioners, investors, and funders) added to this complexity.
- **Upskilling:** Scanning the market, engaging potential commissioners and upskilling them consumed significant time and resources, due to the novelty of SOPs. Those familiar with outcomes-based payments required

less negotiation time. On the other hand, some Commissioners lacked clear budgets or were unwilling to commit to specific outcomes.

Decreasing Setup Effort:

- **Standardisation through learning:** Over time, lessons learned and improved processes (eg standardised contracts, specifications) reduced the time and effort for subsequent setups. By later phases (eg LCF), SOP programme launches were quicker, taking a few months rather than the year-long setup in earlier phases.
- **Supportive stakeholders:** The facilitator of setting up the MHEP SOPs was a good relationship with the social impact investor. Investors, like Big Issue Invest, were supportive once initial agreements were established, though initial negotiations required detailed modelling and structuring.

Managing external uncertainties

Because of ongoing service changes, staff turnover and the need for continuous adaptation, managing a SOP was not a straightforward process. MHEP's team (Social Finance) expressed that uncertainty did impact management efforts and time. These uncertainties arose from unforeseen disruptions in service delivery due to commissioning context, difficulty in predicting the causes of underperformance, site-specific performance variability and the need for sustained capacity in Social Finance to address these issues. One analyst described the skill of addressing uncertainty sensitively: *'Being confident but also being sensitive and understanding that most of these services are operating in a pretty challenging service environment. So the success of the service is not down to one thing usually, it's down to a huge number of things and it's difficult sometimes looking in, especially at the beginning with the relationship to understand/interpret what's happening and where the blockers to achieving outcomes. But I think the longer you stay on, you pick up tropes.'*

Unforeseen disruptions included frequent restructuring of referral pathways due to the NHS transformation[15], which involved moving clinical commissioning groups into integrated care boards, extending access to mental health services into broader NHS care pathways and digitalising health and social care.

Managing underperformance

The challenge of investigating causes of underperformance in different MHEP SOPs' service providers over time meant that performance management efforts were 'lumpy' for most sites, with disproportionate resources spent during periods of underperformance. Identifying underperformance sometimes relied on expert judgment from Social Finance analysts rather than clear, objective indicators. This

[15] For more information about the NHS transformation, see: https://transform.england.nhs.uk/

uncertainty necessitated deeper analysis (both quantitative and qualitative) and increased effort to validate findings and communicate them effectively to stakeholders, especially the governance board/investors. Efforts to address underperformance included stress testing, additional reporting and qualitative assessments, all of which increased time and effort.

Relative to the MHEP SOPs, traditional commissioning very rarely investigates the reasons for underperformance on a regular basis. When underperformance occurred, Social Finance often overcommitted resources beyond its funding to manage performance, conduct analyses, and support services. This required renegotiations with the board for additional funding to address gaps, thus further slowing processes. One director explains: *'The trade-off was that we as Social Finance as the executor of MHEP, the operational team around MHEP, we weren't fully funded really to do the level of performance management that we were involved in.'*

Evidencing outcomes

Providers had to validate their outcomes with their service users (for the independent auditor to verify). For the MHEP SOPs' outcomes achievement to be verified, and hence payment to be released, 'evidence' in the form of a job slip was required. While Social Finance emphasised to providers the importance of providing evidence of service users' outcome attainment, the difficulty of this task for providers may have been underestimated, especially due to the population group of the service. Providers emphasised that some people with severe mental illness will not provide evidence of outcomes achievement (ie their start of employment via a job slip) due to the paranoia of not wanting someone else to know where they work. While the MHEP management team made suggestions, the realities of dealing with 'human services' means that applying a consultancy logic may not always work. In other words, poor performance data can result partly from the specific health condition being treated and the patient's experience of consenting to data sharing rather than entirely from the service's underperformance. This was a learning for Social Finance.

Closing the SOP

MHEP managers in Social Finance revealed that the completion of the SOP took longer than expected. They initially assumed that one quarter would have been sufficient for closing the SOP after service delivery ended (31 March 2024). Contractual completion activities included:

- aligning contract legalities for shareholder agreement and commissioner contracts
- financial repayments to the SOP investor (Big Issue Invest)
- end-of-grant reporting for the LCF, including actual spending versus original projections and reconciling any discrepancies

- terminating governance for the board, including deeper financial analysis. Social Finance negotiated an additional extension until September 2024.

The finalisation of the investor repayment requires understanding of the final cash balance per site according to performance, plus the surplus on previous loans (Camden, Addictions and Haringey & Barnet). However, there was uncertainty surrounding these amounts due to the timings of payments with Social Finance dealing with a preferred rate of return, interest, and repayment of initial equity to the investors. Clarifying these issues required work from MHEP's management team. Ultimately, some loans had a negative cash balance. However, it is important to note that financial performance does not entirely equal outcomes performance due to the presence of caps.

Social Finance highlighted the fact that the end-of-grant reporting for an outcomes fund requires reconciling original projections of outcomes and finances of a project. For MHEP, this involved returning to five-year-old documentation often produced by a member of staff who was no longer at Social Finance. This may suggest the importance of appropriate management of the documentation of a SOP's baseline expectations when applying to an outcomes fund like LCF. These reconciliations required significant effort, exacerbated by the presence of caps and LCF engagement requirements.

Financial reconciliation with caps[16]: Caps were negotiated with the commissioners and served the budgetary purpose of preventing unlimited commissioner spend on outcomes. However, intermediaries (and other investment fund managers in the social investment market) often expressed criticism of their rigidity and argued for caps to be raised or to be considered over the duration of the projects rather than on an annual basis[17]. Three main criticisms of caps from Social Finance in the interviews were:

- Providers would 'disconnect' or be unmotivated for performance when they hit the caps. However, since providers reported higher intrinsic motivation, this specific criticism is likely unfounded.
- Since performance can fluctuate in providers (especially in the employment and social care sector), high-performing periods could not help counter-balance low-performing periods in a contract due to an annual cap on payments.

[16] SOP commissioners in the UK often put a cap on the maximum amount of outcomes payments they are willing to make, which can also cap the investors maximum return. A cap attempts to reduce uncertainties and risks in terms of financial commitments; however, their use (and consequence) is often debated. MHEP SOPs also have a cap.
[17] LCF had placed both an annual cap and total cap for MHEP on each MHEP site. These amounts can be found in Table 4 of Report 2 (Hulse et al. 2024). Ranging from £65,172 to £217,241 as an annual cap across the 5 SOPs, while the total cap ranged from £260,690 to £868,966.

- The purpose of a cap to limit the upside of investment returns was argued to be less important if one is generating social value and a 'net benefit' for all.

Financial reconciliation between forecast and actuals: Interviewees at Social Finance emphasised the pressure and additional time spent engaging with LCF, especially concerning financial reconciliation. Workforce management and employer relationships can dramatically affect the outcomes achievements of a provider, but this fact may not have been incorporated into original assumptions. Commissioners may wish to assess referral pathways and potential disruptions when procuring providers and to ensure that SOP stakeholders such as intermediaries (eg Social Finance) understand this risk. One MHEP manager explained that if two or three staff members in a team left (which occurred more than once), this would significantly reduce outcomes achievements for one quarter. Similarly, when a referral pathway began operations after delays due to NHS transformation, or an opportunity opened in the job market, this could create a '300% improvement' in outcomes. This means that the numbers and the outcomes payments which LCF paid could vary significantly from those which were initially modelled, and communicating the reasons for these discrepancies could be challenging. One MHEP manager explained:

'I think there's a constant pressure to document what happened compared to what we said would happen at various times. So we're able quite confidently to say what has happened on MHEP, but we're less confident in what has happened against a benchmark that was set at a particular point in time. Because it's not always clear why that benchmark was set or who set it or what assumptions were in it.'

The difficulties of reconciling what the SOP achieved over the three to four years also led to communication challenges: *'So the thing is that I don't think LCF grasped a lot of the time is that the nature of a SOP, you're never going have an even period, you're never going hit targets evenly across three years or whatever. Sometimes LCF would be very up on our back about that and it's like well they achieved the outcomes so, what can WE do? We're just reporting what they've achieved and they've evidenced the outcomes so. But they were like "Your forecasts are wrong" and it's like, "OK, well, yeah, but the tariff is the tariff and they achieved the outcomes." So that was quite challenging.'*

9.3. Drivers of non-monetisable transaction costs

Transaction costs are defined as the expenses incurred when buying or selling a service in addition to the cost of the service itself. These include additional

monitoring, searching, negotiating and enforcing in a SOP on top of a traditional service delivery contract of IPS. In MHEP SOPs, according to the interviews and survey, transaction costs were perceived to increase, decrease, or be negligible depending on eight key factors:

- efficiency
- Special Purpose Vehicle structure
- standardisation
- Outcomes-based contract (OBC) readiness[18]
- perceived complexity or unfamiliarity
- uniqueness of OBC procurement
- high-performing delivery
- administrative burden

Box 6: Key findings: What are the key factors affecting the costs of the MHEP SOPs compared to IPS in traditional contracts?

The key factors that reduced transaction costs across the MHEP SOPs compared to traditional IPS were:

- The efficiency of the outcomes monitoring process, which was streamlined and in part automated
- The standardisation of processes using a Special Purpose Vehicle which allowed the pooling of resources, smoothed cash flows and facilitated the inclusion of additional commissioners or geographical areas
- The standardisation of the contracts across sites which provided a clear understanding of responsibilities across all projects, allowing teams to focus on execution rather than reinterpreting terms
- The OBC readiness of commissioners and providers.

The key factor that increased the transaction costs of the MHEP SOPs compared to traditional IPS was their perceived complexity, particularly in terms of new contractual clauses and new stakeholders.

Overall, transaction costs were high during setup of the SOP, the completion, and during periods of uncertainty in underperformance.

Table 4 below presents key results from the survey. In this section, we explore each of these factors using quotes from MHEP SOP stakeholders to explain its impact. Lastly, we examine the key phases of the SOP contract/lifecycle, to

[18] OBC readiness can be defined as the range of factors that influence the suitability and feasibility of launching a successful SOP. GO Lab and Social Finance partnered to produce a number of resources to assess service providers and commissioners' readiness to be part of a SOP, including: https://golab.bsg.ox.ac.uk/documents/Methodology.pdf

identify times when transaction costs are reportedly higher and to explore how they can be mitigated.

Table 4: Summary of key factors which affected transaction costs across the MHEP SOP

Key question (comparing to traditional IPS)	**Answer**
Did **efficiency** over time decrease the time needed for the MHEP SOPs?	Yes
Did the **SPV** decrease the time needed for the MHEP SOPs?	Yes
Did the **standard contracts and/or standard manual** for IPS decrease the time needed for the MHEP SOPs?	Yes
Did knowing **OBC readiness** characteristics decrease the time needed for the MHEP SOPs?	Yes
Did **perceived complexity and/or unfamiliarity** increase the time needed for the MHEP SOPs?	Yes
Did the **uniqueness of procurement** for OBCs increase the time needed for the MHEP SOPs?	Negligible
Did **high-performing** sites decrease the time needed for monitoring the MHEP SOPs?	Negligible
Did **perceived administrative burden** increase the time needed for the MHEP SOPs?	Undecided

Efficiency reduced transaction costs

There were several gains in efficiency over the length of MHEP's SOPs history, especially in outcomes monitoring. Efficiency was gained through:

1. the SOP's clarity on outcomes, which meant providers increased their focus on performance (preventing stagnant caseloads)
2. automating data cleaning/analytical tasks from the MHEP management team
3. well-drafted contracts that made performance management steps easy to follow
4. financial models built from MHEP's first year that LCF MHEP SOPs adopted
5. the expertise of MHEP management teams, which was learned over time.

The key driver of the efficiency gains was the MHEP management's choice to prioritise putting structures and processes in place to ensure the SOPs ran smoothly. MHEP used the profit from outcomes payments towards this goal. MHEP made explicit resource and financial investments in improving efficiencies, eg by automating a range of processes related to outcomes monitoring. A MHEP director explains: '*[I]t was basically what we were earning through the outcomes*

payments, coming from the MHEP accounts itself [we would recycle]. It's like spending more to earn more by improving the efficiency of the services.'

For example, analysts ensured repetitive tasks such as data cleaning and analysing that were necessary for monitoring providers' outcomes in the SOP were automated using IT software. The software performed up to *'30 to 40% in terms of repetitive tasks'*, allowing managers to focus more on supporting the providers.

However, staff turnover in Social Finance (four different managers and six different analysts over the LCF contract) risked efficiency gains. Delays in communication from the governance board of MHEP (which occurred only quarterly) and the performance management team (which met more regularly) also posed challenges to efficiency. Interviewees warned that efficiencies in a SOP cannot always be achieved on the service delivery-side due to the nature of human services: *'[T]here would always be service-related things that were getting in the way.'* This view was neatly summarised by a MHEP manager: *'I don't perceive that the efficiency over a time is about our relationship with Commissioners and providers. I think it's more about how to be an effective performance management team in MHEP.'*

Standardisation via the SPV structure

The MHEP management team argued that MHEP's approach to SOPs (hosting several SOPs under one SPV) helped them save time across the contractual lifecycle. Also, the SPV allowed one SOP's loss on the contract due to underperformance to be balanced with the performance revenue of another SOP. The ability to pool resources across the SPV and hedge cash flows against varying performance across contracts added financial resilience and reduced risks of underperformance impacting individual contracts, easing the financial stress for the investor and the management team.

The SPV also simplified the process of adding commissioners and expanding across geographical areas or cohorts, reducing search and setup costs. This is because the service models, investor, standard contract, outcomes, payments, financial models, board and governance were already existing and mostly fixed. One MHEP director makes this argument known:

'The best thing about MHEP is the fact that the funding and investment flows into MHEP as a vehicle. Basically, the top-up payment goes into MHEP and the investment goes into MHEP and then MHEP has been able to contract with different commissioners and in some cases, providers. And the beauty about that is that then becomes a platform. So the thing that was dramatically simpler with MHEP I think we've ever seen…and I think that is a really underplayed story in MHEP.'

Standardisation of IPS and the contracts reduced transaction costs

The use of similar contracts across projects simplified understanding and execution, as teams did not need to familiarise themselves with different contractual terms. One analyst stated, *'[T]he standardisation of the work done for each project, I would say it's saved me time for sure. The fact that I could understand what our responsibilities were for all projects because they were all the same made it easier.'*

Social Finance co-developed a service specification for IPS with commissioners, manualising the IPS intervention (for both SOPs and non-SOPs) and developing a detailed delivery guide based on learning from previous MHEP SOPs. This guide is now incorporated into IPS Grow for non-SOP contracts due to its usefulness. This standardisation made it easier to engage with commissioners, draft contracts, and evaluate provider bids. One MHEP director explained: *'So as it became more manualised, it became a lot easier, procurement became easier, the document like how we needed to go out to the market became more aligned, understanding what good look like when providers did respond…Because when you put out the offering, they knew what they were signing up to at the very beginning.'*

However, it is important to note that standardisation in the MHEP SOPs was achieved through a focused commitment to learning from Social Finance. This was demonstrated by the fact that *'there was a lot of work in the initial phases and yeah, we just got smarter and better in LCF.'*

As a result of standard contracts and IPS service specifications, MHEP SOP commissioners and providers benefited from knowing what to expect from the outset, reducing negotiation and alignment time. Standardisation provided a clear understanding of responsibilities across all projects, allowing teams to focus on execution rather than reinterpreting terms.

OBC readiness decreased transaction costs (compared to other SOPs)

As the Commissioning Better Outcomes Fund (CBO) gave way to the Life Chances Fund, MHEP developers began to identify ideal characteristics of a commissioner and provider (see p. 38-39 in Hulse et al, 2024). Knowing these characteristics during the search phase of finding commissioners for a SOP and procuring a provider substantially reduced the time required for negotiating contract terms, evaluating bids and scanning the market.

According to our qualitative interviews, commissioners and providers often signed contracts without fully grasping outcomes payment structures, leading to inefficiencies and challenges during implementation. Those with prior OBC experience required less upskilling and had a smoother setup. Conversely,

misaligned expectations between funders and commissioners not familiar with outcomes-based models hindered progress in implementing the SOP. For instance:

- Commissioners were often unfamiliar with the nuances of SOPs, such as performance-linked outcomes payments, leading to delays and restructuring especially around annual caps. Contract restructurings were required in two sites to accommodate commissioners' financial systems.
- Providers also underestimated the complexity of SOPs, particularly around achieving certain targets (and their rationale of focusing on one over the other) and managing risks.

However, over time, the MHEP team gained knowledge about which characteristics in a provider and commissioner signified that they were 'ready' for an outcomes-based contract and a SOP. While some characteristics of OBC readiness are mandatory for a SOP, many skills can be developed and taught. Social Finance highlighted that while providers with basic data collection systems can be trained and supported to develop more robust systems, the capacity for learning and a willingness to embrace change cannot be instilled through the implementation of a SOP.

Earlier MHEP SOPs (supported by CBO) demanded extensive effort to align commissioner timelines, budgets, and approvals and providers' systems. Later MHEP SOPs (supported by LCF) required less time as it became easier to spot the interested providers and commissioners.

Perceived complexity increased transaction costs

Despite the advantages of having standard contracts across the portfolio, some complexity remains since *'it's still quite hard to get alignment between different areas [and actors]',* which increased the time required of the management team.

The contractual arrangements inherent in the SOP structure were often complex, involving caps, varying payment models, and different terms for each project. These details required significant effort to understand and reconcile. Each project's unique financial and operational setup meant stakeholders often had to start from scratch to understand the differences between contracts.

Furthermore, the inclusion of extra stakeholders, like the LCF and Social Finance, to a traditional provider-commissioner arrangement added complexity to decision-making and discussions. Therefore, the SOP structure necessitated coordination across multiple parties, thereby increasing the time and effort spent on alignment and communication in all phases of the contract (search, negotiation, monitoring, enforcement). However, the need for these lengthy coordination efforts was reduced over time by holding trust-building exercises, inviting an LCF representative to MHEP board meetings, and upskilling commissioners.

As mentioned earlier in the report, searching and negotiating with commissioners was described as a 'circular' process which took about **nine months**, involving searching, agreeing on structures and co-developing specifications. The MHEP development director explained:

'[S]o balancing and wrangling the modelling to determine what the outcomes value could be, and agreeing with Commissioners in order for top-ups to be approved by LCF, which would then be financially viable to take through the investment committee was by far and large the biggest time and effort.'

As the intermediary, SPV, and co-commissioner in the SOP, MHEP had to absorb much of the complexity, taking on a balancing act between capacity building and additional administrative work. Despite the challenges, the SOP structure offered benefits such as faster response times in addressing underperformance and in financial negotiations during COVID. Stakeholders highlighted that MHEP acted as a facilitator, enabling quicker resolution compared to traditional funding mechanisms.

Uniqueness of procurement had negligible effects on transaction costs

Commissioners ran an open procurement and while MHEP supported this process, they were not directly involved. Negotiations were slightly longer, but procurement of the SOPs run by MHEP was not different from procurement for a traditional contract.

High-performing services had negligible effects on transaction costs

Monitoring activities (eg data collection, report validation) remained consistent across all sites, regardless of performance level. Enforcement however can differ.

High-performing sites typically need minimal enforcement efforts and rarely require extensive problem-solving. In contrast, underperforming sites demand significant time for developing improvement plans and engaging commissioners, which is an additional cost. This additional enforcement cost is negotiated with the governance board for the intermediary to spend more time performance managing low-performance sites. Nonetheless, every site requires a standard package of performance management as specified in the contract, regardless of performance. Therefore, high-performing sites do not and cannot take less time than the standard.

Outcomes reporting and administrative burden had mixed effects on transaction costs

Outcomes reporting was noted as a challenge for both the MHEP management team (which used the LCF data portal) and the VCSE providers. It was consistently an issue that took more time than anticipated. While this time-consuming work was perceived as a result of the novelty of outcomes-based contracting (and SOPs), it was an area highlighted in need of improvement as it raised initial transaction costs (though it decreased them over time).

Reporting to three separate kinds of stakeholders (outcomes fund, impact investor/governance board and provider was perceived as imposing a small administrative burden, but it was viewed as an acceptable task of monitoring, and the actual time spent on reporting wasn't significantly affected. One analyst described reporting this way: *'[A]fter a point of time you really didn't actually think about monitoring as an administrative burden like because when you actually gain your providers' trust at that time you really want to do something for them which basically helps them in the service.'*

As the lead applicant for the LCF funding, MHEP provided its SOPs with crucial support to apply for and unlock additional financial resources through the LCF. The MHEP team led the multi-stage LCF applications for projects. By assuming responsibility for this process and most of the reporting, MHEP insulated projects from the majority of the administrative burden involved in the Outcome Fund.

However, some aspects of performance management were identified as needing improvement. Providers described an excessive level of reporting requirements in IPS delivery that detracted from employment specialists' focus on delivery. These 'layers of reporting' involved multiple submissions to MHEP, IPS Grow spreadsheets, separate reports for Commissioners, the Mental Health Services Data Set (MHSDS) and previous additional spreadsheets for NHS England. This was not a unique issue of the SOP but was suggested to be due to the fragmentation of IPS delivery in the UK.

Moreover, the final stage of invoicing for validated outcomes was seen as a key bottleneck in the performance management process for Social Finance. MHEP team members described confirmation delays from local commissioners, which in turn postponed release of outcomes payments from LCF. Generally, providers described a feeling of reassurance from MHEP's support during the validation stage and in rechecking figures.

Providers noted that the time-consuming task of evidencing SOP outcomes was not required in traditional contracts. It was estimated that each outcome claimed required 1.5-2 hours of 'chasing evidence from clients' in the form of a job slip and explaining why that was needed. Each frontline staff member commented that a lack of evidence made it impossible to claim some job outcomes. These unclaimed outcomes were estimated to constitute at least 20 job starts and 34 job sustainments over the course of the MHEP SOPs' contract duration. The provider's team lead spent the most time and effort – an average of 2 hours each week – to

follow up on the evidence of outcomes achievement. Nevertheless, this time was considered worthwhile according to providers interviewed: *'MHEP lighting that little bit of a fire in terms of…you need to collect evidence for those outcomes and putting that bit of rigour by doing that and satisfying those [MHEP] requirements. I think it's given us that legacy, it's given us that outcome approach. I don't think we would be flying as high as we are now if it wasn't for that really.'*

9.4. Did the MHEP projects have higher monetisable and non-monetisable transaction costs than expected?

As described in this section, several elements were more costly in a SOP:

- Management costs in the MHEP SOPs were £644,550 from the Social Finance intermediary – 25.8% more than expected.
- Higher management costs were attributed to higher performance support demands of the SOPs. The three sites undergoing performance improvement plans (due to missing targets) required more time for reporting and modelling the financial implications and for frequent catchups.
- Setup was a costly phrase, with at least five additional SOPs not signing contracts despite investments in searching for commissioners, creating contracts, aligning budget and procurement cycles, and educating commissioners on outcomes-based contracts.
- The search phase and negotiations with commissioners for MHEP SOPs took an estimated nine months.
- Unexpectedly, management effort and associated cost did not decrease over time thanks to providers' learning and organisational development. Management effort and associated cost was high at the start of the contract, at the end of the contract, and during any period of underperformance and local service uncertainties.
- Contractual wrap-up for the SOPs under the SPV was delayed by two quarters. This was due to additional time required for terminating board governance, deeper financial analysis, and end-of-grant reporting for the Life Chances Fund outcomes fund and investor.
- Actors involved in the partnership agreed that complexity did increase SOP costs. Caps, varying payment models and different payment terms for each SOP required significant effort to understand and reconcile. Additional effort was required to align the LCF, local commissioners and the investor. This coordination was reported to require high effort from the management team, contributing to the unexpected increased costs.
- These costs are specific to the MHEP's approach to SOPs via a SPV structure, so they cannot be directly generalised to other SOPs with different structures.

10. Was the SOPs' outcomes achievement (the SOP effect) related to the intensity of the performance management or performance incentive?

This section describes the two SOP aspects of interest – its performance management and performance incentive – and their relationship with outcomes achievement, specifically achievement of the primary outcomes of 'job start.' For further details of job start achievement, see section '8. Were the MHEP SOPs effective in achieving their outcomes targets and how does their effectiveness compare with that of traditionally commissioned contracts?'

Box 7: Key findings: Are higher performance incentives in the SOPs associated with better job outcomes?

The key findings from the dose-response analysis are:

1. The higher the performance incentive, the higher the probability of job start achievement: a £1,000 increment in performance incentive is associated with a 21% increase in job start achievement.
2. The higher the performance incentive, the faster the job start achievement: a £1,000 increment in performance incentive is associated with approximately 17% increase in the speed of job start achievement.
3. The performance incentive effect, however, is only a partial explanation of the different results in terms of job start. The statistical model indicates there are missing variables, hence it is not sufficient to make predictions.
4. Enfield showed higher probability of job start achievement than Tower Hamlets.
5. This evaluation was unable to find a robust significant relationship between the performance management dose, measured through the duration of the provider's exposure to MHEP performance management, and job start achievement.

Interpretations: The quantitative analysis shows that there was a significant relationship between performance incentives and job outcomes achievement in MHEP SOPs, especially in Enfield. This means that as the price per outcome for engagement and job starts increased, more job start outcomes were achieved.

There are two structural reasons which may explain why Enfield displayed a particularly significant relationship between performance incentive and outcomes achievement:

- It had the highest percentage of provider funding linked to outcomes achievement in the MHEP LCF SOPs, with 30% of providers' contract value contingent on outcomes achievements.
- It rarely hit the outcomes cap, meaning it was still able to receive payments for outcomes achievements.

The intermediary gave significant attention to Enfield due to the financial implications of 30% of its contract value being contingent on performance and of its periods of lower outcomes performance. This focus did contribute to some psychological distress for the provider's service delivery team under their performance improvement plan. Nevertheless, the issue was quickly rectified through more personable performance management, in-person meeting attendance, and transparency.

Limitations: this analysis was constrained by limited data. The results should be interpreted with caution and not used to make generalisable predictions as the statistical model indicates there are other factors at play which we were not able to model with the available data. Furthermore, this evaluation was unable to find a robust statistically significant relationship between the performance management dose (measured through the provider's exposure to MHEP performance management in the SOP) and job start achievement. This is a limitation to our analysis, and we recommend that future studies should explore the potential effect that exposure to different performance management styles has on outcomes achievement.

10.1. Overview of MHEP SOP 'dose'

Table 5 provides an overview of the components of the SOP dose that we have used to construct our analysis.

Table 5: Overview of MHEP SOP doses

	MHEP SOPs' service delivery start date	Performance management range for clients (months)	Previous CBO IPS service delivery prior to LCF MHEP	Mean performance incentive for clients (n=3,538)	% Job start achievement rates (n=4,176)

MHEP Overall Dose	N/A	0 to 96	N/A	£2,270.65	19%
Enfield	04/2020	0 to 47	No	£2,232.89	36%
Haringey & Barnet	Barnet – 01/2016 (previous experience of IPS under CBO) (MHEP under LCF started 04/2019) Haringey – 04/2017 (previous experience of IPS under CBO) (MHEP under LCF started 04/2019)	40 to 83	Yes	£4,632.79	27%
Shropshire	04/2020 (MHEP under LCF)	0 to 47	No	£2,814.42	21%
Tower Hamlets Mental Health	04/2016 (previous experience of IPS under CBO) (MHEP under LCF started 04/2020)	0 to 96	Yes	£1,682.81	17%

Performance management dose

Performance management dose is a proxy for sites' exposure to the intense, data-led performance management routines of IPS and mechanisms of a SOP. More specifically, the variable captures the length of time (in months) between the site's SOP launch (under CBO or LCF) and the date of a service user's referral. So, for a user referred one month after the launch of the MHEP backed SOP, this variable would take a value of '1.' Clients who were referred to the services prior to MHEP SOP's service delivery launch have a performance management dose of 0. This is because at the time of their referral the service had no MHEP performance management experience as part of the SOP.

Figure 9 and Figure 10 show that performance management and average job start outcomes achievement variables do not have a predictable relationship. Rather, Figure 9's distribution of performance management dose (illustrating the number of service users with dose levels) and Figure 10's depiction of performance management dose in relation to average job start achievement rate display a trend that appears random (or stochastic). This is also the case for individual site levels.

Figure 9: MHEP SOPs' performance management dose (n=4,172)

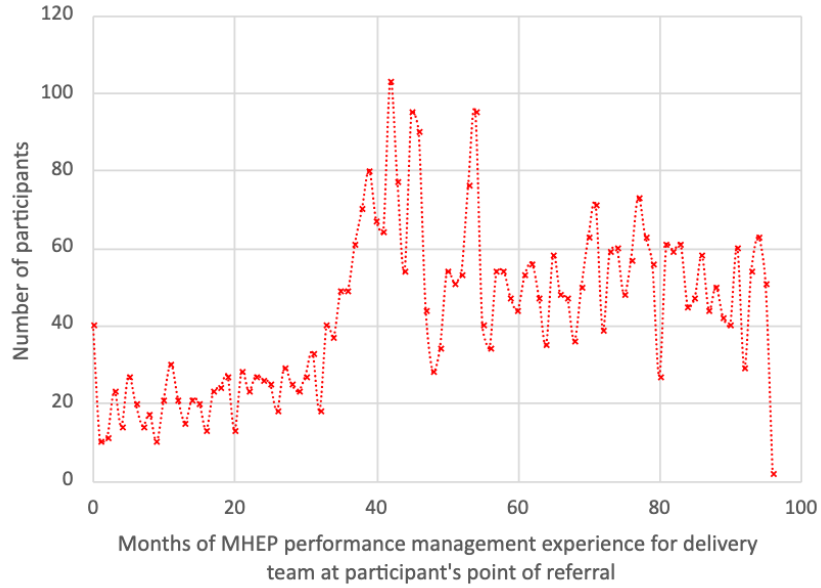
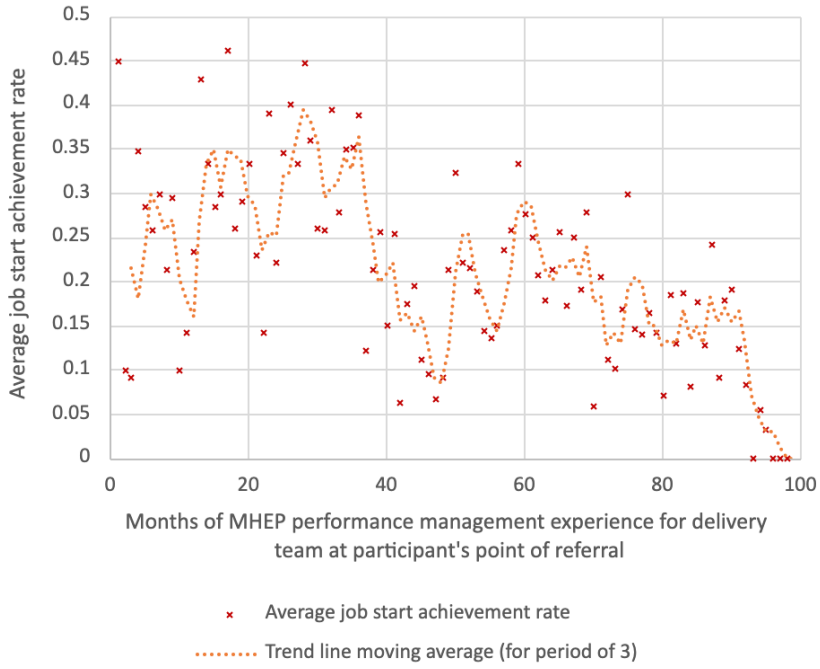


Figure 10: MHEP SOPs' performance management dose in relation to average job start achievement rate (n=4,172)



The dip in the average job start achievement rate at around 36 to 47-months coincides with the wind-down of Enfield and Shropshire's service delivery; both took their last referrals when the performance management dose equalled 47 months. Similarly, the decline of average achievement at 91 months aligns with when Tower Hamlets Mental Health's service delivery wound down. One MHEP manager describes this phenomenon as end of service issues, including problems of

retaining staff due to unpredictability and the risk of the service stopping after the SOP. However, this problem was estimated to not be unique to a SOP, so it could affect any IPS contract. One interviewee said, *'I think the issues we saw towards the end of some of those other services were not actually I would say, related to them being near the end. They just happened to be near the end, but were related to just the difficulties in finding and retaining staff in an IPS service, which is sort of separate from the* SOP *structure.'*

The myth that management of a SOP has a steady decline over time of the contract of a SOP is debunked in MHEP's experience. Nonetheless, wind-down was an observed effect of the MHEP SOPs, so it may need to be anticipated in future SOPs. Providers may experience a drop-off in performance that needs to be included in financial modelling and reporting expectations and mitigated through scaling across the service. Management teams such as Social Finance may experience pressure at the end of a SOP to ensure services continue at similar staffing, to reduce contract insecurity through steady payments (by not hitting caps), and to negotiate with commissioners to extend contracts. This potential wind-down effect could be countered by exit planning being incorporated and prioritised within project management from the outset of SOPs.

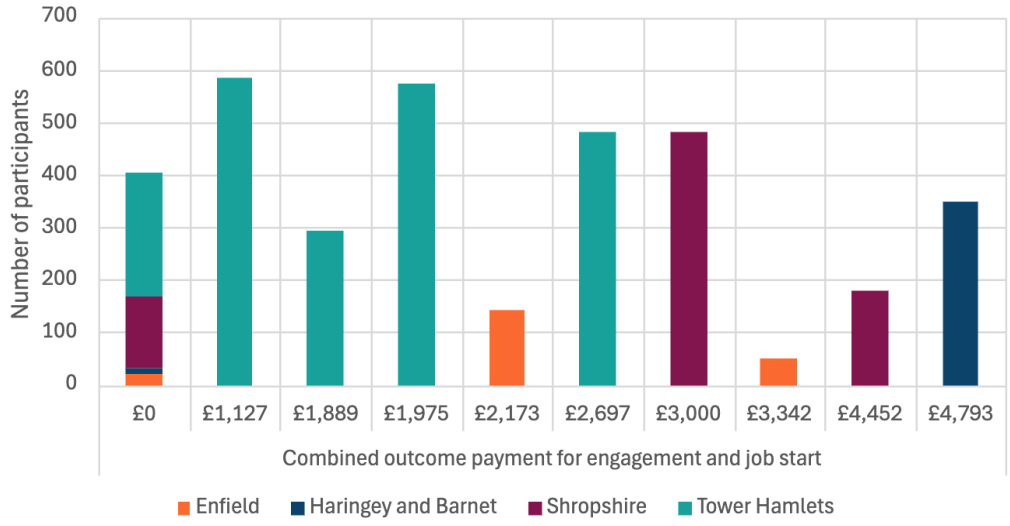
Performance incentive

The distribution of the performance incentive dose, ie the potential fee which can be claimed if clients achieve both engagement and job start outcomes, is depicted in Figure . This total potential fee was calculated by taking the outcomes metric prices when the client was recruited to the project. Only outcomes metric prices for the LCF have been included in this analysis, and no assumptions have been made for the clients referred prior to or after projects' delivery. This combined outcomes metric price varies by site due to contract revisions[19] made through the projects' lifespans. For instance, during the COVID-19 pandemic, the sites had no performance incentive. Instead, they had guaranteed grant payments with fixed values, which were made irrespective of site achievement. Thus, the performance

[19] Outcome payment variations made for MHEP SOPs occurred for 2 main reasons: COVID-19 contract revisions and tariffs. In light of the 'health and employment crisis', MHEP SOPs created contract revisions in May 2020 whereby providers would receive 'medium scenario' payments (grant payments based on medium performance forecasts). All shifted back to outcome-linked payments in October 2020. However, all except the Haringey and Barnet project decided to use modified outcome payment tariffs (known as 'Type 2 tariffs'). These revised outcome tariffs acknowledged the challenges created by COVID-19, particularly for employment support programmes. As a result, they placed a higher payment value on engagements and first jobs in some cases. Additionally, in Tower Hamlets, outcome price to MHEP decreased each year. For example, a successful engagement of a new user would pay £600 in year one and £231 in year 4 of the SOP contract. For more information, see 'Tariff changes from COVID-19 Pay' in Hulse et al. 2024. Since all of these variations in outcome price may have had an effect on performance via incentives, they are included in the analysis to test this relationship.

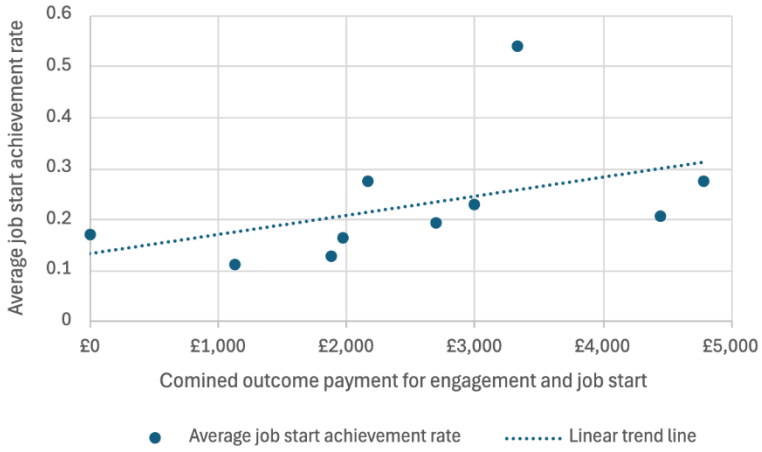
incentive dose has been set to 0 for clients referred when payments were guaranteed (in the COVID-19 pandemic).

Figure 3: Performance incentive dose by site (n=3,538)



The following figure (Figure) depicts the outcomes metric achievement of clients in relation to their performance incentive dose. The performance incentive value with the highest average job start achievement rate is £3,342. This total payment fee was only present for Enfield.

Figure 4: Performance incentive dose in relation to average job start achievement rate (n=4,176)



10.2. Assessing the relationship between performance incentives and job starts

Regression analysis

To test the effect of performance incentives, we initially divided service users with incentive dose of less than £2,000 from those with an incentive dose of more than £2,000. We tested whether there was a statistically significant difference in the proportion of users achieving a job start in the two groups. We found that **the difference is statistically significant**. Therefore, a higher outcomes payment was associated with greater probability for outcomes achievement.

We investigated this trend further through a regression analysis. This statistical technique enables us to predict or explain the variation in one variable based on another variable. Here, we have applied this method to assess the association between the job start outcomes (known as the 'dependent variable') and the performance incentive variable (the 'independent variable') whilst also taking into account, and holding constant, variables like service delivery site, local economic conditions and national unemployment rates. These control variables have been added to the model to control for effects like the changing competitiveness of job attainment through time and different services/geographic locations.

The results of this analysis, and all subsequent analyses with control variables, have been limited as only a constrained selection of control variables could be included. This is because there was a high amount of missing data, specifically for the clients' demographic information. Thus, the inclusion of demographic control variables would have substantially reduced the analysis sample size. For further details on our methodology, please see Appendix A.

Unsurprisingly, all subsequent logistic regressions had low Pseudo R-squared[20] values. This is a measure which suggests that the models do not effectively capture the wide-ranging factors which influence employment obtainment. Therefore, the results need to be interpreted cautiously. However, further regression models with demographic control variables have been trialled, despite their missing data, to test whether results hold regardless of the specifications of the model.

The regression model considers increments/decrements of £1,000 in performance incentives. The model suggests that there is a positive, statistically significant relationship and **for every £1,000 increase in incentives, the probability of a job start outcome increases by 21%. This is equivalent to the observation that, for every increase of incentives by £100, the probability of a job start outcome increases by 2.1%.** Nevertheless, the models' low Pseudo R-squared could reflect that the independent variable has a more nuanced or indirect effect on job start

[20] The regressions have a Pseudo R-squared measure, which is reflective of how the model fits the data. This measure ranges from 0 to 1 and the higher its relative value, the better the model fit. This is a measure which suggests that regressions models effectively capture the wide-ranging factors that influence outcomes achievement.

rate than the results first suggest. Although there is a positive effect, the model is not sufficient to make predictions.

Through our model specification tests, we found that regardless of the inclusion of available demographic control variables in this regression, see Appendix D, Table A-1, the same trend was consistently found and was statistically significant. To offer further validation to our findings, we conducted a statistical analysis that relied on 1,000 simulations of the data, where the payment incentive dose values were randomised on each simulation (all other variable values were not edited and reflect real client data). The regression model was then run for each of the 1,000 datasets to test the effects found with the randomised performance incentive variable. Through this testing, overall, evidence points to there being no relationship between the randomised variable and job start. This supports the view that the relationship between the payment incentive dose and job start outcomes is real and not due to a random effect or the model assumptions. More details of this and further assumption testing can be found in Appendix A.

Figure 5 Proportion of job start for clients with less than and more than £2,000 performance incentives

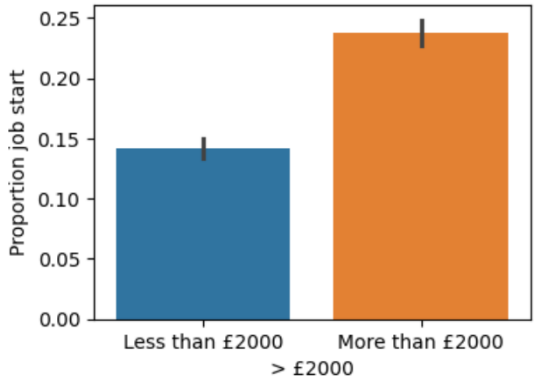


Table 6 Performance incentive dose logistic regression model output, per £1,000 increments/decrements in performance incentives

Number of observations = 3,533
Log likelihood = -1660.3
p-value = 0.00**
Pseudo R-squared= 0.03

Job start outcomes	Coefficient	Standard error	z	P>|z|
Performance incentive	0.191	0.047	4.062	0.000**
Site (Reference: Tower Hamlets)				
Shropshire	0.032	0.154	0.209	0.834
Haringey and Barnet	0.110	0.195	0.567	0.571

Enfield	1.085	0.171	6.349	0.000**
Economic inactivity (%) by local authority district	-0.0278	0.020	-1.358	0.174
Number of unemployed people per vacancy nationwide	0.134	0.055	2.424	0.015*

Significance ratings: ** = <0.01 and * = <0.05.

Survival analysis

To provide a more nuanced understanding of the payment incentive dose, we exploited the availability of data around the length of time between referral to the programme and job start outcomes. Survival analysis is a field of statistical tools used to assess the time until an event occurs. We used this approach to analyse how long it took for service users to achieve job start outcomes and whether the payment incentive influenced this timing (details in Appendix A).

Survival analysis requires information on the duration of time a client is observed until the event of interest occurs. NHS England have stated that there are no fixed time limits to the length of IPS support received, and 'services may choose to review with unemployed clients after 9 months of support… [but] this should be done on a case-by-case basis' (NHS England, 2023). Figure 6 shows that the majority of service users who achieved the job start outcomes obtained their job within the space of a year from their referral date. Thus, a 12-month cut-off was used and clients who started a job after 12 months from their referral were not included in this survival analysis.

Figure 6: Months from referral to job start (n=806)

As we did in the regression analysis, we first ran our analysis dividing users into two groups: those associated with a low payment incentive (less than £2,000) and those associated with a high one (more than £2,000).

We found that service users in the high payment conditions achieved job start outcomes faster as well as at a greater overall rate. Over the time of the programme, the slower achievement in the low payment incentive dose, does not achieve the same proportion of job start outcomes (Figure , left chart). This is as we expected, based on the previous regression.

We then refined the results to account for control variables and to analyse the effect of payment incentive doses in £1,000 increments/decrements (see regression output below, and Figure , right chart). The analysis indicates that a £1,000 increase in payment incentive dose is associated with approximately a 17% increase in *speed* of job start outcomes (see

Table 7).

As before, to further validate the results, we ran a test by simulating the data 1,000 times (see Appendix A for more details). The simulation suggests that the performance incentive variable is responsible for the effect found in the regression output. Additionally, through our model specification tests, we found that the same trends held regardless of the inclusion of demographic control variables in this regression.

Figure 7: Survival analysis outputs. The top charts represent the speed of job outcomes achievement for sites with low (<£2,000) or high (>£2,000) performance incentives (top left) or in £1,000 increments (top right).

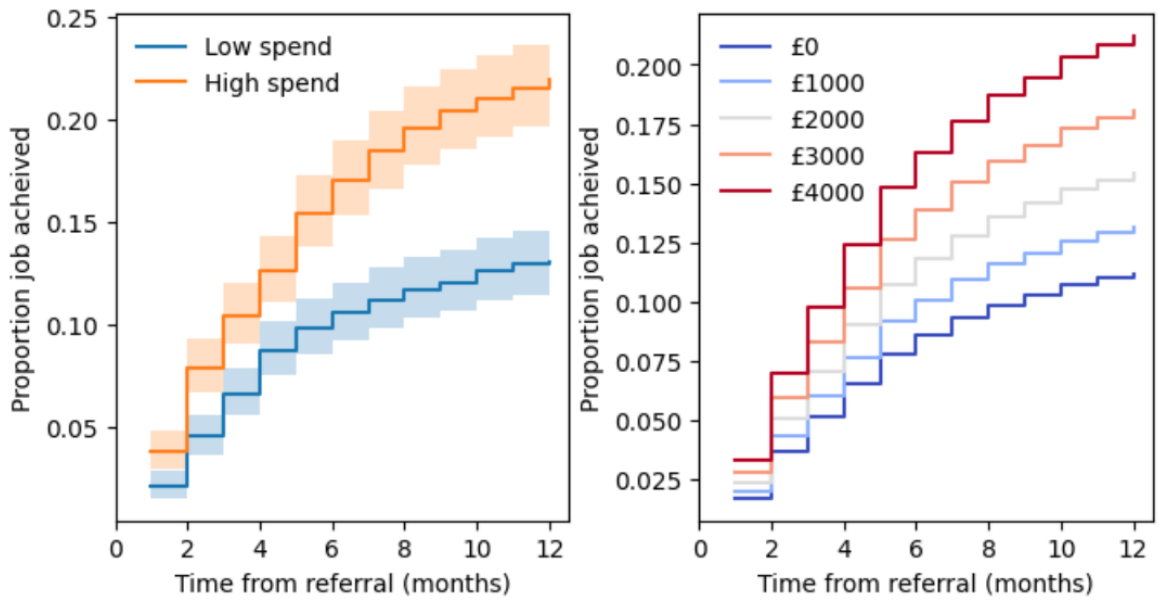


Table 7: Cox proportional hazards regression output (performance incentives in £1,000 increments/decrements)

Job start outcomes	Coefficient	Standard error	z	P>|z|
Performance incentive	0.159	0.044	3.610	0.000**
Site (Reference: Tower Hamlets)				
Shropshire	0.003	0.145	0.022	0.983
Haringey and Barnet	0.075	0.182	0.412	0.680
Enfield	0.936	0.150	6.253	0.000**
Economic inactivity (%) by local authority district	-0.032	0.018	-1.723	0.085
Number of unemployed people per vacancy nationwide	0.067	0.051	1.308	0.191

Significance ratings: ** = <0.01 and * = <0.05.

10.3. Performance management dose analysis

As we explain in the methods section, the qualitative data suggested that the delivery teams needed time for the MHEP way of working to become embedded. We hence measured the months of experience with MHEP at the time of referral as a proxy for performance management dose.

We ran two regression models to assess the association between 'performance management dose' and job start outcomes. The regressions were not able to provide clear results because the relationship between performance management and job start was not consistent or predictable in one direction. This meant that the regression was not reliable as this analysis is not designed to be used with variables with this kind of unpredictable relationship. As a result, performance management testing is self-contained, and the variable has not been applied to any of the other regression analysis. More details of the two regression models and their assumption tests are available in Appendix A.

The relationship between time and dose seems more complex. For instance, in the data we observe a potential wind-down effect: service users referred later in the programme had worse than predicted job start outcomes, possibly due to the wind-down of the programme. This potentially confounding effect has been controlled for in the regression models to isolate the effect of performance management on job obtainment. Nonetheless, further research is required to

assess performance management. It would be benefited by the collection of more data which captures varying service delivery qualities (like granular IPS fidelity scores, amount of staff, and service pressures) and performance management routines (like number of meetings with intermediary organisation and providers and/or commissioners, training frequency, and performance improvement plans). An improved ability to capture performance management, alongside more information on individual characteristics of users, would likely generate models with better fit and predictive power. Thus, the limitations of this analysis support the idea that social welfare initiatives (including SOPs and non-SOPs) need to prioritise data collection/sharing to create an effective knowledge base around the SOP mechanism.

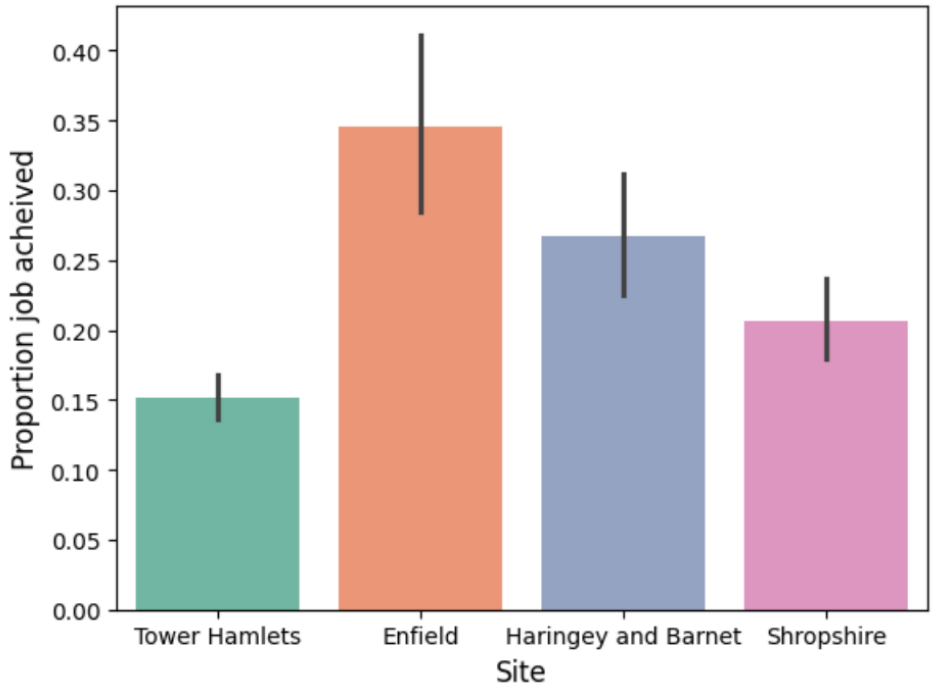
10.4. Site comparison analysis

We ran two regressions models (using the same variables from the performance incentive regression). One model did not include information on the sites, while the other did. We produced indicators to assess the 'fit' of the models. As we expected, the indicators suggest that using site as a variable in the model provides a more accurate explanation of what is happening, ie has a 'better fit.' Furthermore, we also used a Chi-Square test, which suggested that site and job start outcomes variables are associated. A Chi-Square test is a statistical tool to assess whether observed patterns in data are due to real differences or just random chance.

To further unpack this association, we ran a statistical test that compared the proportion of job start outcomes between pairs of sites. See Appendix A for more details on the site comparison analysis. We found that:

- The difference in the proportion of service users with a job start outcome is statistically significant in five of the six site pair comparisons.
- Figure demonstrates that the statistically significant difference in proportion of jobs achieved (suggested through site comparisons) is due to Tower Hamlets Mental Health's relatively poorer performance.
- Enfield had the highest proportion of job start outcomes and Tower Hamlet Mental Health had the lowest. The difference is statistically significant (Figure).

Figure 8: Job start achievement by site



11. How were different actors incentivised for performance?

This section presents the findings on the effectiveness of incentives and accountability within the MHEP SOPs.

Box 8: Key findings: Where is the effective incentive in the SOPs?

MHEP had financial and non-financial incentives within its SOP structure. Those at risk of gaining and losing payments based on outcome performance were the intermediary (Social Finance via the MHEP SPV), the investor and, to a limited degree, providers. In terms of risk by amount, the intermediary was paid 100% on outcome achievement; investors would start to recover the capital provided (and interests) only once the outcome payments covered the operational costs of the contract; and providers were mostly paid through block payments that shielded them from a significant portion of the risk, but in some cases they were in part paid on outcome achievement (for as high as 30% of the contract value in the case of Enfield), bearing the risk of underperformance. Commissioners and LCF were both outcome funders who therefore invoiced according to achievement and were essentially shielded from the risk of underperformance. However, all stakeholders as part of the SOP had varying non-financial incentives for the MHEP SOPs to be a success.

The qualitative data in our evaluation suggest that the most effective financial incentive in MHEP SOPs sat with the performance management team in Social Finance, the intermediary.

This incentive is multifaceted and derived from several sources:

- Outcomes payments incentive
- Investor scrutiny (oversight and pressure)
- Role as a learning tool/system leader
- Aligning stakeholders and relational trust
- Impact-driven mission/reputational success.

The analysis shows that several conditions needed to be met for the incentives on the intermediary for performance to be effective: clear role separation between the social impact investor and the MHEP performance managers, transparency, a governance board that was capable of holding people accountable, a robust reporting and contractual framework, joint goal alignment with commissioners, and centralised oversight.

Furthermore, time constraints and resource limitations, outcomes variability in local pathways, fragmentation/complexity of contracts and transaction cost trade-offs may weaken the strength of the intermediary's incentive.

Table 11: Incentives and risk in the contract from each different stakeholder in MHEP SOPs

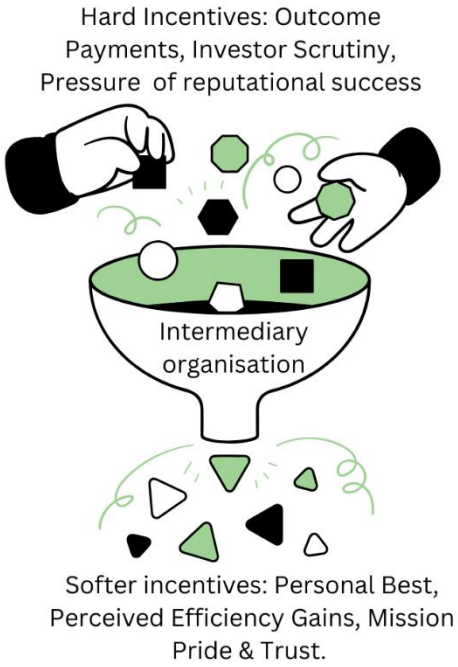
Types of risk	MHEP-Intermediary & SPV manager	Investor-Big Issue Invest	Commissioners-Local Authority/NHS Funders	Providers
Total Risk	High	High	Low	Medium to low
Financial	- Outcome-dependent income - Cash flow risk from upfront payments to providers - Exposure if outcomes not met	- Risk of capital loss if outcomes not achieved (uncapped downside)	- Limited direct financial risk - Risk of post-MHEP 'cliff edge' for ongoing service funding	- Limited exposure due to block payments in most contracts - Some risk where outcome payments applied (ranging from 0–30% of contract value)
Strategic	- Risk of failing to embed systemic change - Legacy depended on uptake by commissioners	- Sought to demonstrate proof of concept for social investment in health/employment	- Incentive to prove IPS effectiveness to secure future funding - Gained strategic influence from success	- Risk of not being sustained post-funding if IPS not embedded into business as usual - Lack of commissioner understanding may affect support
Operational	- High admin and coordination burden - Must maintain credibility with multiple stakeholders	- Limited direct influence on operations - Dependent on MHEP oversight and contract structures	- Limited capacity and expertise to oversee IPS delivery independently - Often reliant on MHEP support	Subject to external scrutiny from MHEP and commissioners - Must manage performance reporting, outcome targets, and data systems, along with Contract Review Meetings

Performance	- Responsible for overseeing provider delivery - Must deliver across diverse geographies & cohorts	- Relied on MHEP and providers to perform well - Delayed returns if performance dips	- Relatively insulated from day-to-day delivery - Held accountable locally for service quality	- Required to meet demanding employment outcomes (eg job starts and sustainments) - Continuous monitoring may feel intense or repetitive - Accountability pressure can cause stress or demoralisation if not well managed - Tension between quality and quantity of outcomes
Reward	- Financial surplus from outcomes if well-managed - Recognition as a national model - Platform to influence policy	- Social impact and financial return if outcomes achieved - Continue to build track record in outcomes-based investing	- Top-up funding via LCF - Value-for-money service model - Enhanced service quality and strategic innovation	- SOP duration over 3-4 years ensured financial stability supporting staff retention and planning - Received support, training, and performance insight from MHEP - Strong performance builds trust with commissioners - Increased accountability created a ‘fire’ to improve - Staff often felt more motivated and valued when regularly engaged with non-financial incentives

Intermediary's & providers' risks and incentives

From qualitative interviews, we found that **incentives were most effective when directed at the intermediary level**. Providers' incentives from the MHEP SOPs' provider-level performance-based payments were more muted than expected (Hulse et al 2023). Providers derived more motivation from mission pride, personal bests, efficiency, and autonomy gains (Hulse et al, 2023). This is because the front-line staff did not internalise the outcomes payments; the intrinsic motivation they felt as providers to IPS clients was more powerful than the aim of achieving payments. So, our evaluation finds that the effective incentive in MHEP SOPs sits with the performance management team in Social Finance, the intermediary.

Figure 9: The incentive flow: between providers and intermediary



As represented in Figure 17, the intermediary translated the financial incentives it received into softer incentives which respected the intrinsic motivation of the providers. With the knowledge that frontline staff are not motivated by financial outcomes payments, the intermediary must find a way to encourage behavioural change among provider staff if and when necessary to produce better outcomes. As we discussed in our previous report, it did this in MHEP by motivating through mission pride
(acknowledging the values and external pressures of doing what's best for an IPS client), personal bests (encouraging teams on personalised positive feedback), and

efficiency and autonomy gains (encouraging teams to buy into the partnership with tangible improvements in data systems and processes)(Hulse et al, 2023).

The intermediary did respond to financial incentive through outcomes payments, investor scrutiny, and pressure of reputational success to drive the achievement of outcomes. Therefore, it can be suggested the success of the MHEP SOP (as defined through effective incentives on performing pre-defined outcomes) was contingent on an intermediary who could balance the act of translating the incentives of different stakeholders.

The translation of 'hard' to 'soft' incentives is a key finding. In fact, the academic literature on financial incentives to drive performance highlights the risk of *crowding out* intrinsic motivation (Ariely et al, 2009). Monetary incentives are seen as demotivating and rewarding a transactional attitude to tasks, creating tensions and damping the intrinsic motivation of actors wishing to 'do good.' In the MHEP SOP, the intermediary responded to the financial incentives but also acted as a buffer between the board and the investor, protecting the muted incentives of the providers and avoiding the 'crowding out.'

Investors' risks and incentives

From the investor's perspective, MHEP SOPs offered an opportunity to generate both social and financial returns. Investors, such as Big Issue Invest for MHEP, are theoretically incentivised by the prospect of contributing to meaningful social change – in MHEP's case, helping individuals with severe mental illness or learning disabilities to gain and sustain employment – while also recouping their investment through outcome payments. These payments may be capped, which limits potential profit even in cases of overperformance. Big Issue Invest did not cap the returns for MHEP, but it had an agreement to share part of any surplus with Social Finance. The downside risk, by contrast, is usually uncapped: if services fail to deliver, investors may lose some or all of their capital. They also face cash flow risk due to the time lag between investment and outcome verification, as well as limited direct control over service delivery. Nevertheless, the use of a pooled special purpose vehicle structure mitigated some of this exposure by distributing risk across multiple contracts and geographies. Investors never met all the IPS service delivery sites, and the only form of contact was the MHEP governance board between Social Finance and themselves. Ultimately, the financial incentive was high for investors, who passed on this pressure to the intermediaries as investors had little control over the IPS service delivery.

Commissioners' risks and incentives

Commissioners' primary incentive lay in their ability to enhance IPS service delivery with reduced upfront costs, primarily through matched funding from the LCF. This financial incentive was complemented by strategic incentives, including alignment with NHS and local authority goals around mental health and

employment outcomes (eg co-location ambitions and increased workforce). The involvement of MHEP as an intermediary provided commissioners with performance oversight, data analytics and technical support that interviewees often expressed they lacked capacity for in-house. However, commissioners also faced several risks. This included the potential 'cliff edge' following the end of LCF funding, which raised concerns about long-term sustainability for the VCSE providers. However, this risk was perceived to be low and mitigated since all services continued IPS delivery in further contracts beyond MHEP SOP contract completion. Moreover, the complexity of the financial model and MHEP's lead role in contract management occasionally led to feelings of reduced autonomy and disengagement among commissioners in the contract review meetings. Since commissioners only paid for the outcomes achieved and all providers were sustained post-MHEP, ultimately the financial incentive and the risk to commissioners in the MHEP SOPs were low.

11.1. The effective incentive sat with the MHEP performance management team

The qualitative data from the interviews suggest that the effective incentive in MHEP SOPs sat with the performance management team in Social Finance, the intermediary. This incentive was multifaceted:

- Outcomes payments incentive
- Investor scrutiny (oversight and pressure)
- Role as system leader
- Aligning stakeholders and relational trust
- Impact-driven mission/reputational success.

Incentive 1: Outcomes payment incentive's effect on intermediary's behaviour

MHEP itself was motivated by a dual financial incentive: to ensure that contracts broke even or yielded a surplus, and to demonstrate the success of its model to stakeholders, including its investors and the board. There was a chain reaction: outcomes payments influenced investors and the board, leading the board to pressure and prioritise MHEP's performance management, which motivated the SOPs' providers through support and performance tracking.

The outcomes payment *'increases incentive because it goes all the way to the board if it doesn't go right. The financial consequence is key here to drive incentive,' and 'the incentivisation comes from the board on the performance management team. So it's another kind of parallel channel of incentivisation, I*

suppose you could call it, which isn't directly related to the provider, but the provider seems to buy into it from what we found anyway.'

Incentive 2: Investor scrutiny

MHEP's accountability to their governance board and investors ensured continuous improvement. Analysts often revealed the pressure they felt leading up to board meetings every quarter. The board had high expectations about the kind of information they wished to receive and the context of providers' performance. Similarly, MHEP managers spent a significant time in the lead-up to the meeting writing board papers, as '*if the SOPs are not doing well, then the board do ask a lot of questions and sometimes they're not that easy to deal with. They take up a lot of work and analysis.*' However, despite the scrutiny and pressure the MHEP staff as the intermediary felt at the time, they all agreed it was '*appropriate since the investor is holding most of the risk'*.

Investors prioritised financial returns and social impact, applying pressure to drive MHEP's focus on outcomes. One MHEP manager described the chain of pressures this way: *'[O]utcomes payment puts pressure on the investors. They put pressure on us as performance management. And we have to find a way to filter that into good performance. I think that DOES work. The lever of just being "I don't think you're performing well" – it didn't have to connect to financial payments – is a genuine lever cause people are human and respond to not performing well. Just [them] being disappointed and we're all Type-As.'*

The relationship between MHEP's management staff and the board was both constructive, with the two collaborating on support solutions for providers depending on performance, and challenging, with the board informally threatening to replace Social Finance with another performance manager, as illustrated by the following quotes by different intermediary staff and the board:

- *'They see us as contractors who are delivering on behalf of this vehicle. But if there's someone else who could do it more cheaply, or effectively, they consider that and they consider whether they should recommission the services. So that's how they hold us to account.' – MHEP director*
- *'Well, their lever ultimately was to replace us. I think what was complicated was they said they would, not actually because of performance but because of money.' – MHEP manager*
- *'Regarding another investment, we've replaced Social Finance because they didn't work flexible and we replaced them with individuals who are now looking after our contract. So we don't have them as administrators of that social outcomes contract anymore, because I have analysed both of them so. The reason why we haven't taken action on MHEP [SOPs] is because*

we've had so many local authorities and there were way more multiple parties involved in actually managing that and we didn't have a runway for a possibility to continue this afterwards.' – investor

While Big Issue Invest did not formally go through with replacing Social Finance on the MHEP contract, they did replace them on a separate SOP titled 'Skill Mill'[21]. Big Issue Invest considered replacing Social Finance in the MHEP project due to a range of frustrations from perceived inflexibility, turnover and concerns over strategic leadership and project continuity. Further explanation can be found below in Section 4.

Incentive 3: Role as a system leader

MHEP used performance data not only to monitor outcomes but also to identify areas for adaptation and improvement in service delivery. The adaptation of services was also seen as a key strength of the MHEP SOP model. One interviewee stated, *'SOPs provide an engine for adaptation…there's an incentive to really drive the data.'* This incentive stemmed both from Social Finance's role as a system leader of SOPs (running the first SOP in the world in Peterborough[22]) and from their role in developing IPS Grow (an organisation supporting the largest scale-up of IPS). It is important to note that Social Finance had a vested interest in seeing both MHEP and IPS Grow succeed. Interviewees described that they felt motivated to pass learnings from the MHEP's SOP delivery of IPS on to the IPS community, the SOP community and the social investment community. In the absence of undertaking a Randomised Control Trial, MHEP directors felt that '*doing it through the SOP mechanism gives like an extra incentive and layer of accountability to actually gather the outcomes and data'*, which then allowed this experimental learning to be delivered to other areas of the IPS sector as needed.

This system leadership was felt especially at Social Finance due to their established role as an intermediary (running for over 10 years) sitting between the providers, commissioners and investors. This role was seen to bring additional incentive and an ability to adapt the data collected from a SOP into meaningful action, not only for the providers under the partnership, but for all providers delivering that intervention in the UK:

'Something what Social Finance has learned over like 12 years of SOPs, is that what these SOPs can do really, really, well is provide an engine for adaptation. Just because we have somebody looking at the data closely and like getting under the hood of it and who has an incentive to really like drive it, which neither the

[21] For more information on The Skill Mill SOP, see: https://golab.bsg.ox.ac.uk/knowledge-bank/indigo/impact-bond-dataset-v2/INDIGO-POJ-0195/
[22] For more information on the first SOP in Peterborough (The One Service), see: https://golab.bsg.ox.ac.uk/knowledge-bank/indigo/impact-bond-dataset-v2/INDIGO-POJ-0153/

provider nor the Commissioner have as a kind of classical incentive built into the structure of their relationship. But a SOP performance manager does have that as their primary incentive.'

Incentive 4: Aligning stakeholders and relational trust

The relationship between MHEP and its providers emerged as another crucial element influencing the SOPs' performance incentive. Relationship-building was seen as a necessary precursor to effective performance management. Interviewees emphasised the importance of trust, noting that it was difficult to manage performance without a strong relationship with providers. One interviewee described, *'It's very difficult to performance manage people you've got no relationship with whatsoever; it's impossible basically.'*

These relational dynamics were especially important during the initial stages of contracts, as one MHEP performance analyst pointed out: *'You need a bit of bedding in period…and to watch how people communicate with each other.'*

Building these relationships is a gradual process, and interviewees underscored the challenge of balancing performance management with the need for trust. The ability to separate MHEP's role as a performance manager from the role of the investor, board, or formal commissioner (as funder) was noted as critical in this regard. MHEP's performance managers often positioned themselves as neutral, supportive agents rather than enforcers of financial or contractual obligations, which helped to maintain positive interactions with providers.

As noted by an MHEP manager, this relational trust was the key to performance incentive: *'the team leader's involvement or engagement with us as a team really depended on the quality of how much she trusted me and how much of our relationship was established.'*

Incentive 5: Impact-driven mission

MHEP's success hinged on balancing the financial sustainability of the SOP with its impact-driven goals. As a not-for-profit consultancy, Social Finance had an impact mission that influenced its interactions with stakeholders in the SOP.

'MHEP has an incentive, as a whole, to make a success story of what MHEP is. So part of that is purely impact, have we facilitated the expansion of these services into councils that wouldn't have otherwise delivered them. And crucially, will those services then continue to procure IPS services after we have like kind of warmed them up to it. So I think we're motivated by the desire to embed/to make

sure these services are successful enough that commissioners will continue to invest in them after we go. So an impact incentive.' – MHEP manager

11.2. Facilitators and barriers to the intermediary's incentive

While the incentive in the MHEP SOP for the intermediary to perform was effective, the qualitative interviews revealed the key factors which contributed to or limited its strength (as summarised in Table below):

Clear role definition and target transparency: Clearly defining roles between commissioners, MHEP and providers fostered collaboration and mitigated confusion about requirements. Some providers interviewed initially described some uncertainty about the role MHEP played, attempting to create distance or defend their results out of fear of losing funding. This was described by one MHEP performance analyst as *'[one provider in particular] were very defensive and didn't want to engage, but that shifted over time.'* However, MHEP worked hard to build trust through transparency, explaining their actual role as being that of a collaborator. This was also helped by holding performance reviews in person, using team-building language, performing data analytical tasks that the provider requested, and facilitating training and development through courses or mentoring with an IPS expert.

Separation of roles: Effective performance incentivisation was facilitated when MHEP managers were independent from the 'funders' or the impact investor. This psychologically shifted the relationship into trust between the providers and MHEP since they would slowly become collaborators on performance; they would both work together and separately to mitigate bottlenecks of outcomes achievement. If performance was going well, MHEP encouraged providers to share learnings across frontline employment specialist staff, within the organisation but also across other MHEP providers. Playing a role separate from the board and investors meant that MHEP could facilitate buy-in with the provider.

This approach is described by both MHEP staff and providers separately below:

- *'You become representative of the board rather than yourself and that helps you with that relationship. Well, this is what…I'm just passing this on. So that was that was quite useful.' – MHEP manager*
- *'Having some separation between the [funder] and our role as the performance manager felt really, really powerful there. Because I think if you're in an ordinary contract, the Commissioner is the funder, like there's no separation so it's harder, I suppose, for you to on a person-to-person level separate those incentives out. Where actually I thought it was very useful to be able to say I'm the messenger of the board, but I'm not the*

board. I'm just here to help you address this concern. So having that slight independence, felt like a really useful lever.' – MHEP manager

- *'Because she said that the funders I suppose of MHEP wanted to see an improvement. So she made it clear where it was coming from, not just herself, in terms of "Yeah. Come on, guys. You can do better."' – Provider*
- *'It was MHEP manager at the time, they did a good job kind of putting this to us, in a very nice way, but in quite an assertive way that, "I'm in the middle now, I've got people above me who say that we've got to start turning things around, I'm here to support, but we are going put performance management in place."' – Provider*

Time constraints and resource limitations: The MHEP team was funded on a performance management fee. MHEP management staff often felt they had to go above and beyond due to their desire to help providers and be accountable to the board/investors. Yet this effort would exceed the time and cost they were allocated to, which would dampen the strength of their incentive. One MHEP director described this barrier of resource limitations on the intermediary's incentive from the board:

'So the relationship is: Big Issue Invest obviously appointed us via the board to deliver services for them. So their role is to hold us to account for performance managing MHEP well. And I think there is sometimes tension around the funding available for us to deliver this work and the expectations of the investor around the amount that we can deliver.'

Table 8: Facilitators of and barriers to the intermediary's incentive

Facilitators of the intermediary's incentive	Barriers to the intermediary's incentive
• Clear role definition and transparency • Separation of roles: independence of MHEP performance managers	• Time constraints and resource limitations

11.3. Potential replacement of the intermediary from concerns over the sustainability of the SOP

According to the investor representatives interviewed, Big Issue Invest recognised that Social Finance excelled at contract administration, such as ensuring adherence to legal terms and producing performance reports. However, they

would have liked Social Finance to play a more proactive role in securing tangible business continuation for MHEP. A second concern was that Big Issue Invest perceived the formal and legalistic approach to contracts as ill-suited to the inherently flexible and experimental nature of SOPs. This situation is not unique, and no SOPs supported under the Life Chances Fund continued SOP funding post-Life Chances Fund. In fact, locally commissioned SOPs without top-up funding from a central government fund are extremely rare in the UK. Therefore, there may not have been many opportunities for Social Finance to pursue continuity of MHEP SPV without top-up funding. The investor's unmet expectations around capacity to operate without top-up funding may simply be due to the constraints of the SOP model.

11.4. Implications of the incentives in MHEP

The role of MHEP extended beyond the management of financial flows. Through MHEP, Social Finance performed contractual management, performance oversight and mediation between investors, commissioners and providers. Since the MHEP's financial health and investor return hinged on outcome achievement, it became an active steward – proposing timely interventions, adjusting targets and working closely with all actors in the SOP to optimise performance. While this performance focus was praised by some providers for improving accountability and rigour, it also introduced potential for concerns or frustrations.

MHEP's incentives were multi-layered: financial return, performance credibility, institutional legacy and system leadership. These incentives have generally worked in favour of good quality, performance-focused contract management of IPS. But the strength of MHEP's incentives was also the cause of potential concerns especially for commissioners and investors, and acknowledging these tensions is crucial to appraising the MHEP's SOP approach.

Due to this vested interest, MHEP, acting as both performance manager and recipient of outcome payments, may have prioritised actions that secured outcome payments – even if they created tension with provider capacity or local commissioner priorities. For example, repeated questioning and pressure to improve may have reflected an internal financial imperative more than balanced system stewardship. Also, MHEP's superior understanding of its own complex contracting and control of information flow placed it in a position of power over smaller VCSE providers and even commissioners – especially those less familiar with outcomes-based contracting. However, despite tensions during the performance improvement plans (Hulse et al. 2024) and investor's frustrations, Social Finance was seen to display a collaborative approach at least for service providers.

Financial risk transfer from commissioners to investors did occur and providers were satisfied that they were sufficiently shielded from financial incentives. Legacy also occurred and can be explored in the previous report (Hulse et al., 2024). These were all positive attributes of the MHEP SOP model. Regardless of the investor's frustrations and concern over an intermediated SOP, MHEP managed to provide sufficient relational, performance and contractual management for providers over ten years.

Discussion

The primary objectives of the MHEP longitudinal evaluation were to assess whether the use of a social outcomes partnership (or social impact bond) made a difference to the desired outcomes, compared to traditional commissioning approaches and, if so, how and why. Due to inaccessible data, we were unable to directly compare MHEP (and its impacts) with traditional commissioning approaches or other SOPs. However, this third report of the evaluation has 1) analysed the performance of outcomes achievement across the MHEP SOPs, 2) tested the effect of (performance management and) incentives on outcomes achievement, and 3) understood the time-cost required for a social outcomes partnership in addition to the cost of the service. Furthermore, Appendix B provides reflections on the contribution of the MHEP SOPs to the achievement of the overarching objectives of the LCF.

The quantitative analysis found a significant relationship between performance incentives and job outcome achievement. From qualitative interviews, we found the actual incentive that is effective sits at the intermediary level. Providers were less motivated by the outcome payments than by their intrinsic motivation, as already articulated in the previous report. However, as we found in the most recent interviews, the intermediary's incentives were driven by outcome payments, investor scrutiny, their role as a system leader, aligning stakeholders and gaining relational trust, and an impact-driven mission.

In order for the incentives for the intermediary to be effective in the MHEP SOP, several conditions needed to be met. There needed to be clear role separation between the social impact investor and the MHEP performance managers, transparency, a governance board that was capable of holding people accountable, a robust reporting and contractual framework, joint goal alignment with commissioners, and centralised oversight. On the other hand, time constraints and resource limitations may have weakened the strength of the intermediary's incentive.

This evaluation was unable to find a robust significant relationship between the performance management dose, measured through the duration of the provider's exposure to MHEP performance management, and job start achievement. This is a limitation to our analysis, and we recommend that future studies explore the potential effect of exposure to different performance management styles on outcomes achievement.

While the actual costs display some differences to the predicted costs of the SOP, this may be explained by COVID-19, over-ambitious targets at baseline and service underperformance to those targets, and delayed start of delivery.

Searching and negotiating with commissioners for SOP contractual setup was described as a 'circular' process rather than a straightforward back-and-forth. This

took about nine months, involving searching, agreeing on structures and co-developing specifications. The modelling of outcomes values to ensure financial viability took up about 80% of the search and negotiation stage. Once this was resolved, the contracting proceeded more quickly. MHEP directors interviewed stated that time and setup costs for SOPs that eventually signed contracts were equal to those of SOPs where contracts were not signed and fell through at the last minute. It is estimated that at least five additional SOPs fell through in the late stages.

Delays in MHEP SOP contractual setup often occurred due to approval processes, procurement timelines and staff turnover. For instance, external approvals lags (eg LCF approval processes) and Commissioner role changes meant that several commissioners rescinded their interest. MHEP developers expressed frustration that the centrally offered top-up funding model did not really sink in with local commissioner contract timing, and they suggested that addressing this issue could reduce costs.

Social Finance highlighted the fact that the end-of-grant reporting for an outcomes fund requires reconciling original projections of outcomes and finances of a project. For MHEP, this involved returning to five-year-old documentation often produced by a member of staff who was no longer at Social Finance. This problem may suggest the importance of appropriately managing baseline SOP expectations documentation when applying to an outcomes fund like LCF. These reconciliations required significant effort, exacerbated by the presence of caps and LCF engagement requirements. Furthermore, intermediaries (and other investment fund managers in the social investment market) often expressed criticism of rigidity over caps and argued for caps to be raised or to be considered over the duration of the projects rather than on an annual basis.

In terms of generalisability, MHEP projects structured the contractual relationships of the SOPs through the SPV run by Social Finance whilst in some SOPs there is a direct relationship between the outcome payer and service providers. The SPV structure of MHEP was not unique; a majority of Life Chances Fund projects (61%, or 19 out of the original 31) had this structure rather than a direct contract between providers and outcome payers. Nonetheless, Social Finance's role in managing the SOPs as a SPV via MHEP means that we need to be cautious in interpreting the wider generalisability of the result.

Overall, MHEP left a legacy: it was the world's first SOP aimed at helping people with mental health issues into paid employment and is the longest-running SOPs project globally, having run for a record-breaking nine years. It also created IPS Grow, a national programme to support the NHS scale-up of IPS across the country.

This report has been structured around key policy issues of SOPs, and we refer the reader to the summaries in each section:

- **How do MHEP SOPs deliver IPS compared with traditionally delivered IPS?** See page 26
- **Did SOPs overspend/underspend more than expected?** See page 40
- **What were the key factors that affected a difference in actual expenditure versus expected expenditure?** See page 51
- **Do performance incentives in the SOPs cause better job outcomes**? See page 59
- **Where is the effective incentive in the SOPs and what are barriers and facilitators?** See page 71

Box 9: Key findings of the longitudinal MHEP evaluation

Report 1: The first report found that:
MHEP SOPs provided additional value compared to traditional commissioning via:

- a dedicated performance management function that was seen to drive additional focus on achieving outcomes
- more effective working culture within each local partnership
- identifying and successfully unlocking the LCF funding. This was understood to bring additional financial and human resources to projects.

Report 2: The second report found that:

- The SOPs improved accountability and commissioning practice compared to traditional contracts of IPS.
- Social Finance's MHEP contributed to the national scaling of IPS in the NHS.
- MHEP SOPs' contractual and payment structures were seen as unnecessarily complex and could be simplified via earlier buy-in for design principles, annual caps, and more realistic expectations on forecasting outcomes performance.
- MHEP brought enhanced capacity to providers by building an IT data system, talent pipelines for staff, and efficient data routines. It brought enhanced capacity to commissioners via experience in partnership working and the creation of a new baseline for expected IPS outcomes
- MHEP SOPs' incentives on providers were more muted than expected.

Report 3: The third report found that:

- As the price per outcome increases, the likelihood of outcomes achievement increased in MHEP SOPs. However, the statistical analysis also indicated there were missing explanatory variables, pointing out that, from a quantitative analysis alone, we have a very partial understanding of the SOP mechanism.
- Incentives in MHEP were effective because the intermediary translated the financial incentive into softer incentive for providers, thereby protecting their intrinsic motivation.
- The incentive on the intermediary to achieve outcomes in the SOP came from:

    - Outcomes payments
    - Investor scrutiny
    - Reputational success through its role as a system leader, its impact-driven mission and relational trust.
- The SOP facilitated the implementation of IPS delivery through longer contract duration, reduced contract insecurity and more commissioner engagement than is faced in traditional commissioning.
- Management costs were higher than expected in the SOP, and the setup and wrap-up of the contract took significant effort.

References

Addicott, R. (2016). Challenges of commissioning and contracting for integrated care in the National Health Service (NHS) in England. Australian Journal of Primary Health, 22(1), 50-54.

Ariely, D., Bracha, A., and Meier, S. (2009). "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially." American Economic Review, 99 (1).

Becker, DR., Swanson, SJ., Reese, SL., Bond, GR., McLeman, BM. (2019). SUPPORTED EMPLOYMENT FIDELITY REVIEW MANUAL A companion guide to the evidence-based IPS Supported Employment Fidelity Scale. The IPS Employment Center. Available at: https://www.centreformentalhealth.org.uk/wp-content/uploads/2023/11/final-fidelity-manual-fourth-edition-112619_0.pdf (Accessed 27 January 2025).

Böhm, E., Backhaus, C., Eggert, A., & Cummins, T. (2016). Understanding outcomes-based contracts: benefits and risks from the buyers' and sellers' perspective. Journal of Strategic Contracting and Negotiation, 2(1-2), 128-149.

Bond G. R., Clark R. E., & Drake R.E. (1995). Cost effectiveness of psychiatric rehabilitation. Psychotherapy and Rehabilitation Research Bulletin, 4, 26-31.

Bond, G. R. (2023). Cost-Effectiveness of Individual Placement and Support. Christensen, T. N., Kruse, M., Hellström, L., & Eplov, L. F. (2021). Cost-utility and cost-effectiveness of Individual Placement Support and cognitive remediation in people with severe mental illness: Results from a randomised clinical trial. European Psychiatry, 64, e3. doi:10.1192/j.eurpsy.2020.111

Brinchmann, B., Widding-Havneraas, T., Modini M, Rinaldi M, Moe CF, McDaid D, Park A-L, Killackey E, Harvey SB, Mykletun A. (2019). A meta-regression of the impact of policy on the efficacy of individual placement and support. Acta Psychiatrica Scandinavica, 141 (3). 206-220. Doi.org/10.1111/acps.13129

Bond, G. R., Drake, R. E., & Becker, D. R. (2012). Generalizability of the Individual Placement and Support (IPS) model of supported employment outside the US. World psychiatry, 11(1), 32-39.

Browne, D. J., & Waghorn, G. (2010). Employment services as an early intervention for young people with mental illness. *Early intervention in psychiatry*, *4*(4), 327-335.

Clark R. E., & Bond G. R. (1996). Costs and benefits of vocational programs for people with serious mental illness. In: Moscarelli M, Rupp A, Sartorius N, eds. Schizophrenia. 219-237.

Clark, R. E., Bush, P. W., Becker, D. R., & Drake, R. E. (1996). A cost-effectiveness comparison of supported employment and rehabilitative day treatment. Administration and Policy in Mental Health, 24, 63-77. doi:10.1007/BF02106484

Clark, R. E. (1998). Supported employment and managed care: Can they coexist? Psychiatric Rehabilitation Journal, 22(1), 62-68. doi:10.1037/h0095263

Clark, R. E., Xie, H., Becker, D. R., & Drake, R. E. (1998). Benefits and costs of supported employment from three perspectives. Journal of Behavioral Health Services and Research, 25, 22-34. doi:10.1007/BF02287497

Dayson, C., Fraser, A., & Lowe, T. (2020). A comparative analysis of social impact bond and conventional financing approaches to health service commissioning in England: the case of social prescribing. Journal of Comparative Policy Analysis: Research and Practice, 22(2), 153-169.

Dixon, L. B., Hoch, J. S., Clark, R., et al. (2002). Cost-effectiveness of two vocational rehabilitation programs for persons with severe mental illness. Psychiatric Services, 53, 1118-1124. doi:10.1176/appi.ps.53.9.1118

Frederick, D. E., & VanderWeele, T. J. (2019). Supported employment: Meta-analysis and review of randomized controlled trials of individual placement and support.

Gongora-Salazar, P., Glogowska, M., Fitzpatrick, R., Perera, R., & Tsiachristas, A. (2022). Commissioning [Integrated] Care in England: An Analysis of the Current Decision Context. International Journal of Integrated Care, 22(4): 3, 1–16.

Heslin, M., Howard, L., Leese, M., et al. (2011). Randomized controlled trial of supported employment in England: 2 year follow-up of the Supported Work and Needs. World Psychiatry, 10, 132-137. doi:10.1002/j.2051-5545.2011.tb00035.x

Hoffmann, H., Jäckel, D., Glauser, S., Mueser, K. T., & Kupper, Z. (2014). Long-term effectiveness of supported employment: five-year follow-up of a randomized controlled trial. American Journal of Psychiatry, 171, 1183-1190. doi:10.1176/appi.ajp.2014.13070857

Hulse, ESG., Atun, R., McPake, B., Tayu Lee, J. (2021). Use of social impact bonds in financing health systems responses to non-communicable diseases: scoping review. BMJ Global Health, 6:e004127. doi: 10.1136/bmjgh-2020-004127

Hulse, E., Fraser, A. (2024). How can we scale up evidence-informed health care interventions through social outcomes partnerships? The evolution of social outcomes partnerships in the UK: Distilling fifteen years of experience from Peterborough to Kirklees. Government Outcomes Lab

Hulse, E., Nagarajan, S., Carter, E. (2024). Mental Health and Employment Partnership: second interim report as part of the LCF Evaluation. Government Outcomes Lab.

Hulse, E., Shiva, M., Hameed, T. and Carter, E. (2023). Mental Health and Employment Partnership LCF Evaluation. Government Outcomes Lab, Blavatnik School of Government.

Hutchinson, J. (2022). Individual Placement and Support A guide for Integrated Care Systems. Centre for Mental Health.

IPS Grow. (2025). IPS Fidelity Reviews. IPS Grow. Available at: https://ipsgrow.org.uk/providers/fidelity-reviews/ (Accessed 9 April 2025).

Jordan, N., Stroupe, K. T., Richman, J., et al. (2022). Comparing service use and cost of Individual Placement and Support to usual vocational services in veterans with PTSD. Psychiatric Services, 75, 1109-1116. doi:10.1176/appi.ps.202100197

Knapp, M., Patel, A., Curran, C., et al. (2013). Supported employment: cost-effectiveness across six European sites. World Psychiatry, 12, 60-68. doi:10.1002/wps.20017

Latimer, E., Bush, P., Becker, D. R., Drake, R. E., & Bond, G. R. (2004). The cost of high-fidelity supported employment programs for people with severe mental illness. Psychiatric Services, 55, 401-406. doi:10.1176/appi.ps.55.4.401

Levitt, A., Carter, E., Outes Verlarde, J. (2023). How much does ‘it’ cost? Developing an understanding of transaction costs for impact bonds and social outcomes contracts. Available at: https://golab.bsg.ox.ac.uk/community/blogs/how-much-does-it-cost-transaction-costs-impact-bonds/ (Accessed on 12 November 2024).

Modini, M., Joyce, S., Mykletun, A., et al. (2016). The mental health benefits of employment: Results of a systematic meta-review. Australasian Psychiatry, 24, 331-336. doi:10.1177/1039856215618523

NHS England(2023). Individual placement and support for severe mental illness. NHS England. Available at: https://www.england.nhs.uk/long-read/individual-placement-and-support-for-severe-mental-illness/. (Accessed on 13 January 2024 - 7 January 2025).

NOMIS : Official census and labour market statistics. (2024) *Annual Population Survey/Labour Force Survey*. Available at: https://www.nomisweb.co.uk/sources/aps. (Accessed on 8th January 2025).

OECD. (2012) *Sick on the Job?: Myths and Realities about Mental Health and Work, Mental Health and Work*. OECD Publishing. https://doi.org/10.1787/9789264124523-en.

Office for National Statistics. (2024) *Dataset : VACS01: Vacancies and unemployment*. Available at: https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/datasets/vacanciesandunemploymentvacs01. (Accessed 8th January 2025).

Peckham, S. (2020). The changing public health system: an examination of the new commissioning infrastructure. In Commissioning Healthcare in England (pp. 123-142). Policy Press.

Porteous, N., & Waghorn, G. (2007). Implementing evidence-based employment services in New Zealand for young adults with psychosis: progress during the first five years. *British Journal of Occupational Therapy, 70*(12), 521-526.

Priest, B., & Lockett, H. (2020). Working at the interface between science and culture: The enablers and barriers to individual placement and support implementation in Aotearoa/New Zealand. *Psychiatric Rehabilitation Journal, 43*(1), 40–52. https://doi.org/10.1037/prj0000388

Rinaldi, M., Miller, L., & Perkins, R. (2010). Implementing the individual placement and support (IPS) approach for people with mental health conditions in England. International Review of Psychiatry, 22(2), 163-172.

Richter, D., & Hoffmann, H. (2019). Effectiveness of supported employment in nontrial routine implementation: systematic review and meta-analysis. Social psychiatry and psychiatric epidemiology, 54(5), 525-531.

Sanderson, M., Allen, P., Gill, R., & Garnett, E. (2018). New models of contracting in the public sector: A review of alliance contracting, prime contracting and

outcomes-based contracting literature. Social Policy & Administration, 52(5), 1060-1083.

Shaw, S. E., Smith, J. A., Porter, A., Rosen, R., & Mays, N. (2013). The work of commissioning: a multisite case study of healthcare commissioning in England's NHS. BMJ open, 3(9), e003341.

Shi, Y. (2012). Cost-effectiveness of the individual placement and support model of supported employment for people with severe mental illness: Results from a Canadian randomized trial. McGill University. ProQuest Dissertations Publishing, MR84291.

Social Finance (2023). Building England's IPS sector to help more people into jobs. Available at: https://www.socialfinance.org.uk/work/building-englands-ips-sector-help-people-into-jobs (Accessed 8th January 2025)

Stroupe, K. T., Jordan, N., Richman, J., et al. (2022). Cost-effectiveness of Individual Placement and Support compared to Transitional Work Program for veterans with post-traumatic stress disorder. Administration and Policy in Mental Health and Mental Health Services Research, 49, 429-439. doi:10.1007/s10488-021-01173-z

Waghorn, G., Stephenson, A., & Browne, D. (2011). The importance of service integration in developing effective employment services for people with severe mental health conditions. *British Journal of Occupational Therapy, 74*(7), 339-347.

Waghorn G, van Veggel R, Chant D, Lockett H. The utility of item level fidelity scores for developing evidence based practices in supported employment. Journal of Vocational Rehabilitation. 2018;48(3):387-391. doi:10.3233/JVR-180946

Wharakura, M. K., Lockett, H., Carswell, P., Henderson, G., Kongs-Taylor, H., & Gasparini, J. (2022). Collaboration in the context of supporting people with mental health and addiction issues into employment: A scoping review. *Journal of Vocational Rehabilitation, 56*(3), 271-287.

Williamson, Oliver E. (1996) The Mechanisms of Governance. Oxford: Oxford University Press.

Yamaguchi, S., Sato, S., Horio, N., et al. (2017). Cost-effectiveness of cognitive remediation and supported employment for people with mental illness: a randomized controlled trial. Psychological Medicine, 47, 53-65. doi:10.1017/S0033291716002063

Appendices

12. Appendix A: Methods

12.1. Quantitative analysis: dose-response analysis & performance achievement

The ambition of this final MHEP evaluation report is to offer a quantitative assessment of the impact of the SOP model (RQ1), with reference to more conventional commissioning approaches. Ideally, this assessment should compare an IPS service provided through a SOP (via MHEP), with a non-SOP 'business as usual' IPS service taking place at the same time and subject to similar external events. This ideal setup was not viable due to data constraints. Based on insights from our qualitative research, we conceptualised the SOP as a 'dose' operating at varying intensity (Callaway et al., 2024). In simple terms, we attempted to operationalise some of the key differences between delivering IPS in a traditional contract and delivering it in a SOP. This difference was stronger in some places and at particular times in the contracts (ie at a 'higher dose') than other places or times. This is further explored in the limitations section.

Our operationalisation of the 'dose' for MHEP 's SOPs was informed both by conceptual literature (Carter, 2020) and by the qualitative process evaluation of MHEP's SOPs (Hulse et al., 2023).

Our longitudinal qualitative evaluation provides justification for taking a dose-approach to our treatment of the SOP, finding that the intensity of implementation of the SOP model (operationalised through level of performance management and level of performance incentives, see definition below) was highly variable between sites and over time.

The variation in implementation intensity at the **start of a contract** was emphasised by a service provider who said, '*[I]t took us a year to properly, to get used to it and to understand MHEP.'* Cumulative dose exposure to MHEP management was, of course, lower for new providers than for pre-existing providers (prior to LCF funding): '*[F]rom MHEP's point of view, they've got no previous relationship with us, so they weren't necessarily able to say we know how we operate and we know that the quality is there and we can have confidence that things will get back to normal*.'

Dose can be lower even at the **end of SOP contracts,** depending on the experience of the SOP's performance management and performance incentive. One quote summarises a provider's experience three years into a four year contract: *'I think*

we would have had more outcomes if it was pointed out some detail of what was required to capture them, because it feels like it's only recently we've known.'

Guided by the available project level data, we have identified two sub-components to the dose by which MHEP's SOPs are characterised and will be explored here:

- Intensive performance management (a quarterly cycle of active, data-led performance management routines)
- Performance incentives for the successful achievement of employment outcomes (ie sites are underpinned by a social outcomes contract).

Our hypothesis is that employment outcomes for service users will improve under a higher 'dose' of the MHEP's SOP approach.

We have created a proxy for *performance management*, where the dose is measured through the cumulative exposure of the service provider delivery teams to MHEP's performance management of the SOP (eg the number of months from start of a delivery site's first MHEP contract to client's referral date).

Performance incentives can be measured in several ways. We can derive the monetary value of successful employment outcomes at the time point when a service user was referred, ie the monetary value for a job start in the rate card, or we can calculate the proportion of provider payment that is conditional on the successful achievement of job outcomes.

The outcomes of interest are the job start outcomes, and the unit of analysis is all the LCF MHEP SOPs' service users, ie people with severe mental illness (SMI) who have been referred into the IPS service. Our sample size is 4,176 service users after data cleaning; it includes only service users who were being supported for mental health disorders. To reduce confounding, the MHEP Tower Hamlets Learning Disabilities SOP has been excluded from this part of the analysis, as this project supported clients with learning disabilities.

We conducted the following inferential analysis:

1. Logistic regression to assess the effect of performance incentive on job start achievement
2. Survival analysis to explore the time to job start achievement for different performance incentives
3. Logistic regression to assess the effect of performance management on job start achievement
4. Two-proportion z-test comparing site association with job start achievement.

We complemented the dose-response analysis with basic descriptive statistics of the final validated outcomes achievement of the service providers in the MHEP

SOPs across the contract duration from 2019–2024. We sourced this data from each MHEP SOP service provider, and figures were validated by Social Finance. Achievement is captured by three main metrics:

1. success rates[23]
2. conversion rates[24]
3. fidelity score comparisons[25].

Box A-1: Simple explanation of the key methods in the final reports of MHEP's longitudinal evaluation

A **logistic regression** is a statistical method used to predict the likelihood of an event occurring, such as whether a policy will succeed or if a voter will support a candidate. Unlike regular regression, which predicts continuous numbers (like income or temperature), logistic regression estimates probabilities and classifies outcomes into categories – typically 'yes' or 'no.' For example, it can help policymakers assess how factors like education, income or location influence the probability of someone supporting a policy change (Peng et al., 2002). In this report, the categorical outcomes variable is job start (either achieved or not achieved) and the categorical predictor variables are performance incentive/management, service delivery sites, local economic conditions and national unemployment rates. The regressions have a Pseudo R-squared measure, which is reflective of how the model fits the data. This measure ranges from 0 to 1 and the higher its relative value, the better the model fit. This is a measure which suggests that regressions models effectively capture the wide-ranging factors which influence outcomes achievement.

A **survival analysis** is a statistical method used to predict the timing of an event, such as how long people remain unemployed, how long a policy remains effective or when a business might close. Instead of just measuring whether something happens, it focuses on **when** it happens. This is especially useful in policy research for understanding durations. This 'event' in this report is the start date of employment, referred to in this report as 'job start.' This analysis

[23] Success rates are the proportion of actual number of outcomes achievement over targeted number of outcomes achievements.
[24] Outcomes conversion rate is estimated as the rate at which one type of outcome transitions into the next successive outcome in a causal chain, eg engagement to job start. For instance, at each quarter of MHEP service delivery, this measure was estimated by summing job starts for each project up to that point, then dividing that figure by engagement achievements for the same period, and lastly multiplying the result by 100 to express the rate as a percentage.
[25] Fidelity is a measure of the level to which an intervention is delivered as intended. The IPS Fidelity Scale is a prominent part of implementing IPS services. The fidelity scale is a translation of the 8 IPS principles into 25 items that can be scored. The IPS fidelity scale is sometimes used in performance management, especially amongst service providers. Formal fidelity reviews were not mandated as part of the MHEP contract, but some fidelity elements were included in the meetings between Social Finance, the commissioner and the provider.

is used to describe or predict the probability that the event of interest has or has not occurred by each time point.

A **dose-response analysis** examines how different levels of exposure to a policy, programme or intervention affect outcomes. It helps policymakers understand whether increasing the 'dose' (such as more funding, longer programme participation or stricter regulations) leads to better results. Though it originates in pharmacology, dose-response analysis is gaining popularity in causal analysis of social sciences.

12.2. Transaction costs analysis

There is concern that the time and resource taken to develop, contract and manage social outcomes partnerships is high, relative to more conventional forms of commissioning (Hulse et al, 2021; Levitt et al, 2023). These costs are labelled 'transaction costs', ie the expenses incurred when buying or selling a service, over and above the operational costs of the service itself (Williamson, 1996). They can be described as the resources expended on finding and selecting vendors, negotiating contract terms, monitoring performance, and ensuring that the delivered intervention meets specifications. There is currently very little evidence that systematically outlines the 'transaction costs' in social outcomes partnerships.

Inspired by Petersen et al, 2019, we created an analytical framework consisting of a full comprehensive list of transaction costs borne by the co-commissioner (buyer: Social Finance), local commissioner (buyer: local authority or NHS), and the provider (seller). We separated costs under:

- Search:
    - Scanning the market for potential vendors
    - Developing requirement and specifications (performance metrics and scenario targets)
    - Incentivising potential bidders
- Negotiation:
    - Evaluating formal bids
    - Conducting reference checks of proposers
    - Negotiating contract terms and compensation
- Monitoring:
    - Assessing deliverables (payable outcomes)
    - Gathering information/data from users
- Enforcement:
    - Executing contract options (or termination)
    - Implementing performance incentives
    - Resolving disputes ((re)negotiation, arbitration, litigation).

To obtain transaction costs, we used two sources. Firstly, we used a structured survey completed through researcher-respondent interviews. This research is the first systematic attempt to develop and test a dataset on SOP specific transaction costs. We worked with commissioners, providers and MHEP managers/Social Finance as the intermediary who were purposively selected for their experience in the MHEP SOPs' design and/or implementation to have a robust understanding of the transaction costs. Secondly, for triangulation, we also extracted costs from the following documents: the outcomes-based contracts, the LCF grant agreements, LCF's end-of-grant documents, and Social Finance's internal excel financial models.

Our sample size was 10 completed structured surveys, each lasting approximately 1 hour. We contacted all MHEP staff and commissioners as the buyer-counterparts in the SOP contract for comprehensiveness. Unfortunately, only one local commissioner responded, meaning that the commissioner's perspective was taken from other stakeholders and previous interviews in 2021 and 2023. All others declined or had moved on from their role within the SOP. In this report we present the transaction costs as ranges according to each cost category across each MHEP SOP alongside quotes from interviewees to help interpret the findings.

12.3. Interviews

The aim of the interviews was to seek to understand how traditional commissioning and funding arrangements such as grant funding and fee-for-service contracts compare with SOPs. Our areas of interest include the IPS service's governance and coordination, performance management, and leadership according to providers and local commissioners (in local authorities and the NHS). The interview sampling was purposive, covering the boroughs in London and Shropshire, consistent with the areas covered by the MHEP's SOPs.

From July 2024 to October 2024, we included 5 interviewees (3 provider managers and 2 commissioners in traditional IPS contracts), adding to longitudinal data from a total of 49 interviewees across Phase 1 (2022-2023) and Phase 2 (2023-2024). The interviews were semi-structured, according to 8 categories of questions: commissioning mechanism, partnership and governance, team dynamics and integration, contract management and accountability, performance measurement, learning and evaluation, COVID-19, and future directions. To build a new theory of change diagram, the questions were framed to seek answers for each comparator site's long-term impact, outcomes activities, behavioural change of staff, mechanism of action, assumptions and conditions for impact.

13. Appendix B. Reflections on the MHEP SOPs alignment within the broader LCF objectives

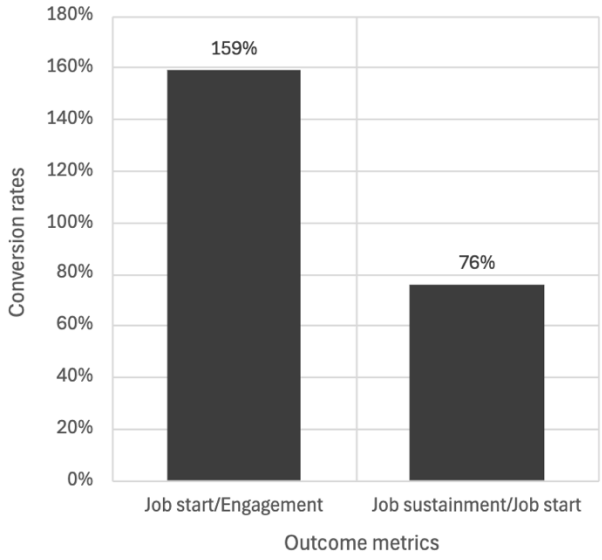
LCF Objectives	MHEP alignment with LCF objectives
Increasing the number and scale of SOPs in England	Three SOPs were created in 2016 through MHEP, which expanded to nine additional SOPs from 2017 onwards. MHEP, supported by LCF, refocused its scaling objective to create IPS Grow and contribute to clinical trials to support IPS in different cohorts.
Making it easier and quicker to set up a SOP	The use of standard contracts, service specifications, and the retention of the original social investment partner meant that Social Finance could decrease the time needed to set up additional SOPs under the same SPV.
Generating public sector efficiencies by delivering better outcomes and using this to understand how and whether cashable savings can be achieved	MHEP generated some efficiencies in the SOPs by building capacity in commissioning units through experience in partnership working, creating a new baseline for expected outcomes and support in IPS services (Hulse et al., 2024). The five MHEP SOPs under the LCF delivered a total of 4,185 outcomes and 954 job starts.
Increasing social innovation and building a clear evidence base of what works	Through its MHEP SOPs, Social Finance founded IPS Grow, which is building a clearer evidence base of what works in IPS services. (Social Finance, 2023)
Increasing the amount of capital available to a wider range of voluntary, community and social enterprise (VCSE) sector providers to enable them to compete for public sector contracts	Prior to MHEP's involvement, the majority of the VCSE providers were facing contract insecurity and short-term funding cycles. The SOP allowed for catalytic capacity development in their organisations in terms of data systems for evidence collection and validation of KPI/outcomes achievement compared to other IPS providers. (Hulse et al., 2024)
Providing better evidence of the effectiveness of the SOP mechanism and the savings that are being accrued	Social Finance contributed to several evaluations run on MHEP.

Growing the scale of the social investment market	MHEP did not scale the social investment market as they partnered with the same social impact investor in all SOPs. However, the loans from the investor grew from £360,000 at the first sites to £1,185,000 at the last sites.

14. Appendix C: Tower Hamlets Learning Disabilities

In the following Figures A-1 and A-2, Tower Hamlets Learning Disabilities achieved more than double the expected outcomes in achieving job sustainment for their clients. However, all other metrics were below best-case scenario expectations. The main reason for this was the lower caseload than expected due to the limited size of a fixed cohort of people with learning disabilities. The ongoing in-work support requirement for a high-need cohort means employment specialists spent longer than in a traditional IPS service, which was not accounted for in the early design of the SOP.

Figure A-1: Tower Hamlets Learning Disabilities outcomes metrics conversion rates



Box A-2: Tower Hamlets Learning Disabilities actual and expected best-case scenario outcomes metric achievement and percentage success rate

	Tower Hamlets Learning Disabilities		
	Target	**Actual**	**Success rate**
Engagement	370	85	23%
Job start	182	135	74%
Job sustainment (<16 hours per week)	40	93	233%
Job sustainment (>16 hours per week)	17	9	53%

Outcomes claimed by Tower Hamlets Learning Disabilities

Figure A-2: Number of outcomes claimed by year

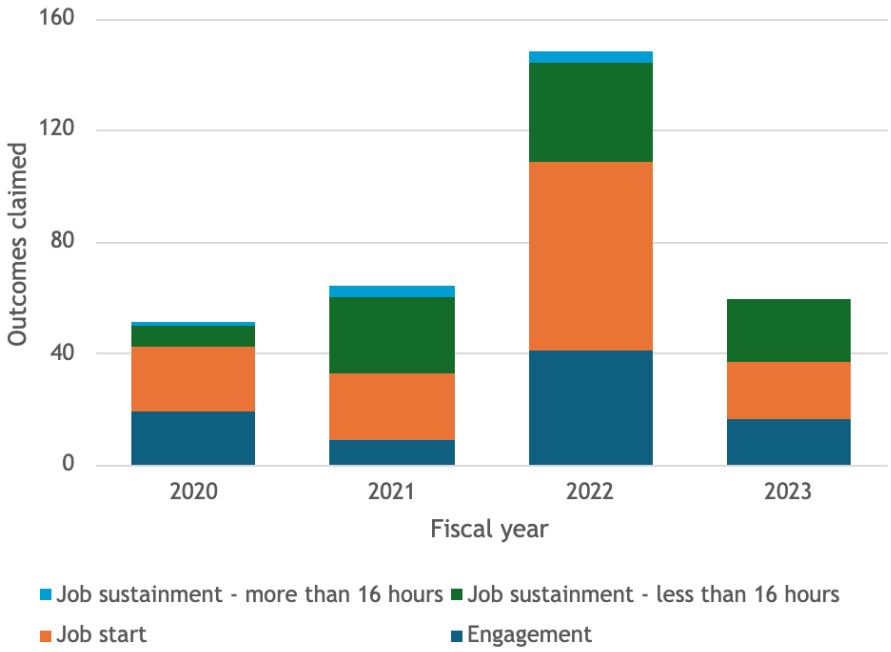


Figure A-3: Outcomes payments

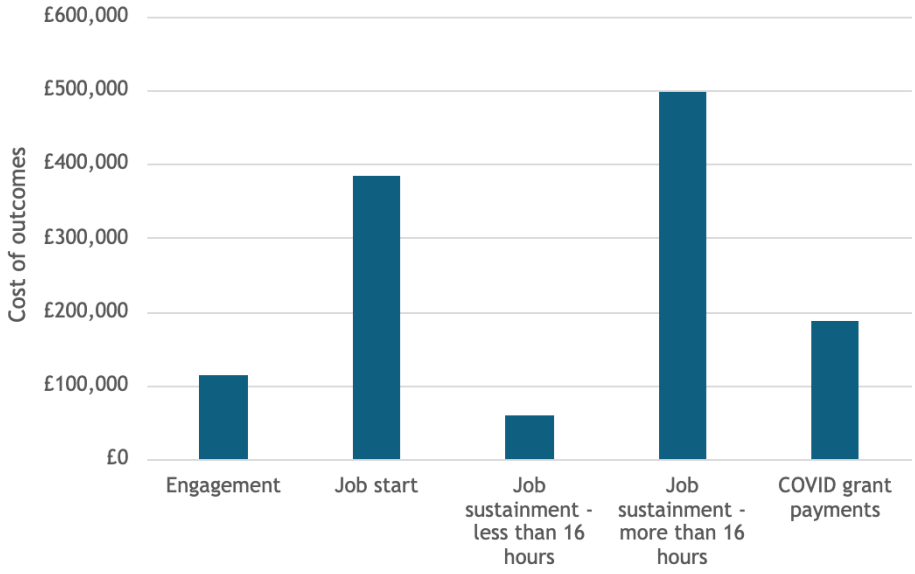
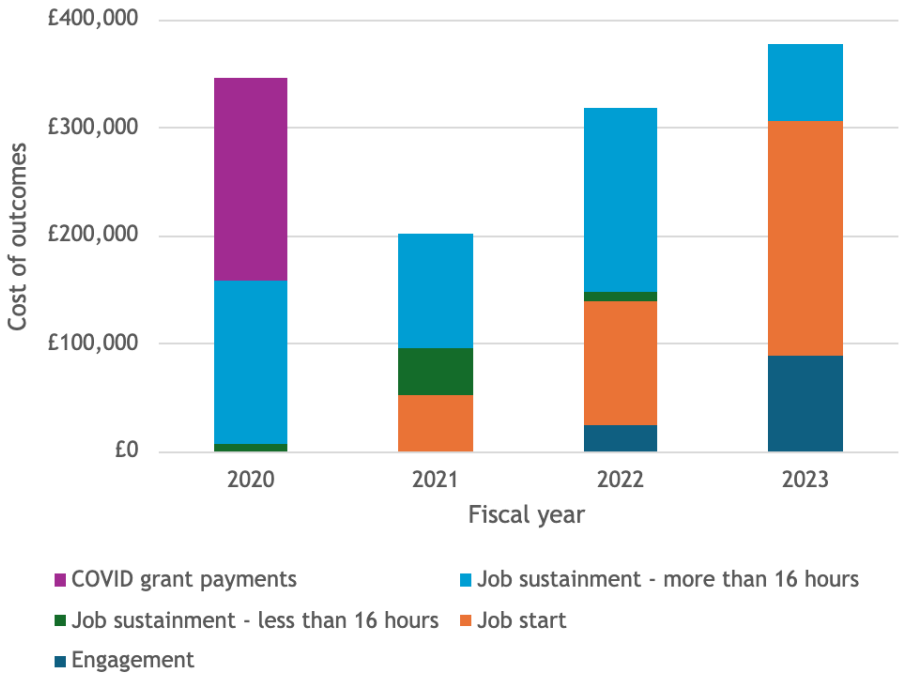


Figure A-4: Cost of outcomes by year



15. Appendix D: Details of the dose-response analysis

Performance incentive dose

Logistic regression

1) Two proportion z-test – Figure A-5's left bar chart (below) illustrates that there is a statistically significant difference between the proportion of job starts achieved for clients with performance incentive doses of less than and more than £2,000. A two proportion z-test was used to assess the difference in proportions of the two groups.

2) Logistic regression – This regression assesses the association between the job start outcomes and performance incentive variables whilst also controlling for potentially confounding variables. Regression reference codes were selected as the categories with the highest sample sizes. The following table explores the effect of the inclusion/exclusion of available control variables into the model.

Table A-1: Testing model selection for performance incentive and job start outcomes regression.

Model No.	Model variables	Logistic regression outputs:	Coefficients for performance incentive:
1	Job start outcomes, Performance incentive	n = 3,535 Log likelihood = -1685.0 p-value = 0.00** Pseudo R-squared= 0.01	0.20 (p-value = 0.00**)
2	[Model 1] + Sites	n = 3,535 Log likelihood = -1666.4 p-value = 0.00** Pseudo R-squared= 0.02	0.15 (p-value = 0.00**)
3	[Model 2] + Economic inactivity (%) by local authority district, Number of unemployed people per vacancy nationwide.[26]	n = 3,533 Log likelihood = -1660.3 p-value = 0.00** Pseudo R-squared= 0.03	0.19 (p-value = 0.00**)
4	[Model 3] + Gender, Ethnicity, Age[27]	n = 2,035 Log likelihood = -1146.6 p-value = 0.00** Pseudo R-squared= 0.03	0.13 (p-value = 0.01**)
5	[Model 3] + Gender, Ethnicity, Age Trialled including missing data within 'Unspecified' category	n = 3,522 Log likelihood = -1514.3 p-value = 0.00** Pseudo R-squared= 0.11	0.16 (p-value = 0.00**)
6	[Model 5] + Religion, Sexuality, Relationship Trialled including missing data within 'Unspecified' category	n = 3,522 Log likelihood = -1496.5 p-value = 0.00** Pseudo R-squared= 0.12	0.16 (p-value = 0.00**)

We have selected Model 3 to be detailed below and used in subsequent survival analysis testing. Demographic variables have not been included in the main analysis as the demographic data had a high proportion of missing values and their addition substantially reduced the sample size.

Regression assumption tests:

Investigations have been completed to check whether the model complies to logistic regression assumptions. For example, no problems have been found regarding multicollinearity or model specification (via STATA's correlation, collin

[26] Economic inactivity (%) by local authority district (NOMIS: Official census and labour market statistics, 2024) and number of unemployed people per vacancy nationwide (Office for National Statistics, 2024) have been added to the model to control for changing competitiveness of job attainment through time.

[27] Demographic information collected by site was appreciably compromised by missing values and inconsistent coverage across sites. Gender (n=2,714), Ethnicity (n=2,535) and Age (n=2,409) had the best coverage, compared to Religion (n=1,537), Sexuality (n=1,405), and Relationship (n=1,435). Including all 5 demographic variables results in the model having a substantially reduced sample size of 794. Below, we have investigated the effect of adding further variables to the model by including missing data and 'Prefer not to say' responses within an 'Unspecified' category.

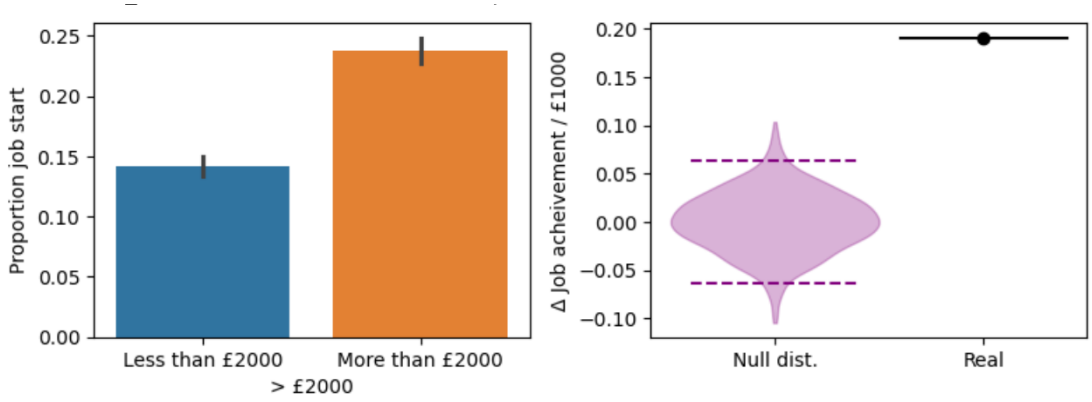
and link tests); there is a linear relationship between the independent variable and log odds of the dependent variable; and potential outliers have been removed from the model.

Nevertheless, the model has a low Pseudo R-squared, which suggests that the model does not effectively capture the plethora of factors which influence employment obtainment. This model was not created to account for all the influences which contribute to job start but to investigate whether there is evidence of a relationship between performance incentive and job start and the directionality of a potential association using the available data. However, this model and its goodness of fit have been hindered by the high degree of missing client-level data, which resulted in the inclusion of only a select number of control variables. Hence, the proceeding results need to be interpreted with caution and model specification robustness checks have been completed – ie it has been checked that results hold despite the inclusion of demographic variables with the model. For model 3, the same trends for performance incentive and job start have been observed throughout our preliminary model specification investigations – as detailed in Table A-1.

Model 3 results:

The logistic regression model coefficients will refer to £1,000 increases/decreases of the performance incentive variable. In other words, the odds ratio of 1.21 (p-value<0.01) associated with the performance incentive variable (see this model's output below) indicates that for each £1,000 increase in the performance incentive variable, the **odds** of starting a job increase by approximately 21% – holding all other control variables constant. Ensuring that the regression outputs are in relation to £1,000 incremental changes in the performance incentive variable allows the regression coefficients to accurately reflect how performance incentive fluctuates over time and across sites.

Figure A-5: Proportion of job start for clients with less than and more than £2,000 performance incentives, and performance incentive logistic regression model outputs compared to permutation testing



Performance incentive dose logistic regression model output

Number of observations = 3,533
Log likelihood = -1660.3
p-value = 0.00**
Pseudo R-squared = 0.03

Job start outcomes	Coefficient	Standard error	z	P>|z|
Performance incentive	0.191	0.047	4.062	0.000**
Site (Reference: Tower Hamlets)				
Shropshire	0.032	0.154	0.209	0.834
Haringey and Barnet	0.110	0.195	0.567	0.571
Enfield	1.085	0.171	6.349	0.000**
Economic inactivity (%) by local authority district	-0.0278	0.020	-1.358	0.174
Number of unemployed people per vacancy nationwide	0.134	0.055	2.424	0.015*

Significance ratings: ** = <0.01 and * = <0.05.

Figure A-5's right graph depicts the results of the performance incentive logistic model. Specifically, the black dotted line marks the coefficient for the performance incentive relationship with job start achievement (when controlled for the covariates detailed in the above full logistic model output).

Permutation testing – The parallel null distribution illustrates the output of a permutation test, and the dashed lines indicate error bands of 2.5 to 97.5% (in line with an alpha value of 0.05). This permutation test (which has also been used in the subsequent analysis) involves running a statistical model 1,000 times, each time shuffling the independent variable contents. In other words, each client had 1,000 random independent variable values for the repeat statistical analysis, whilst all other variables held their original value. This simulation is to investigate whether the effects found in the analysis are specifically due to the independent variable and not caused by random effects or the model assumptions.

For this model, the null distribution and error bands hovers around 0, whereas the real performance incentive variable has a coefficient of 0.191 (equal to an odds ratio of 1.21, p-value<0.001). This indicates that performance incentive is responsible for the relationship found. The robustness of this finding is also supported by the fact that performance incentive consistently had a positive, statistically significant relationship with job start achievement across all model specifications included in Table A-1, with the coefficient ranging from 0.13-0.20.

Nevertheless, the models' low Pseudo R-squared limits the results found and the lack of goodness of fit could reflect that the independent variable has a more nuanced or indirect effect on job start rate.

Survival analysis

Survival analysis analyses and predicts the time until an event occurs, whilst testing the effect of variables on the likelihood of that event happening.

1) Kaplan Meier estimator – Provides nonparametric estimation for two group comparisons. This has been used to explore the rate of job achievement over time for high (>£2,000) and low (<£2,000) performance incentives.

2) Cox proportional hazards regression – This parametric test investigates the rate of job achievement over time for different performance incentives (with £1,000 increments). This analysis facilitates the inclusion of covariates, and the regression will use variables from model 3.

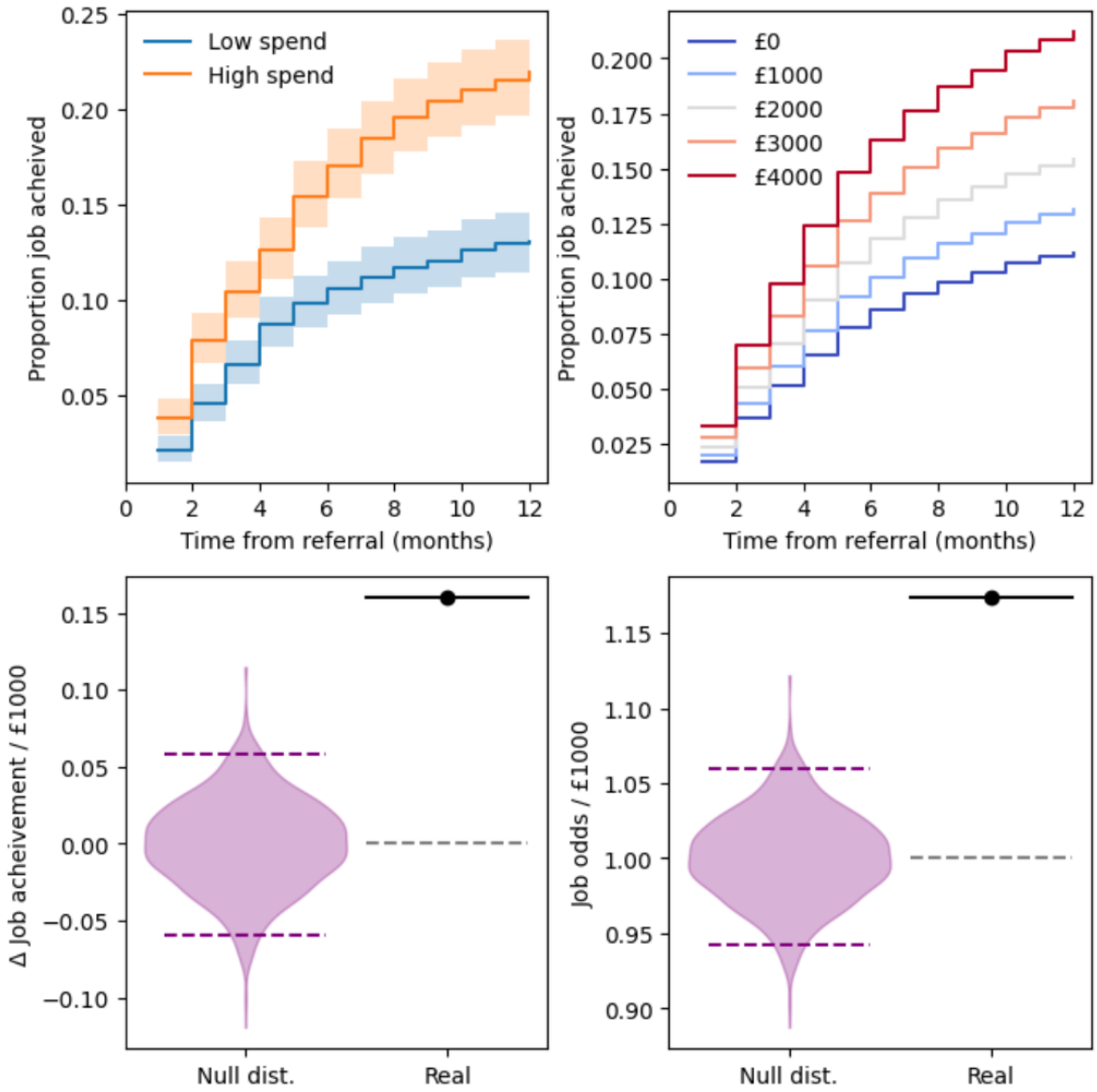
Analysis assumption: 12-month service length for clients. This is because survival analysis requires information on the duration of time a client is observed until the event of interest occurs. NHS England have stated that there are no fixed time-limits to the length of IPS support received (NHS England, 2023), and Figure A-6 shows that the majority of service users who achieved the job start outcomes obtained their job in the space of a year from their referral date. Hence, we have used a 12-month cut-off (clients who started a job after 12-months from their referral were not included in this survival analysis). Nevertheless, other service lengths have been trialled in the analysis, and the same trends were observed.

Figure A-6: Months from referral to job start (n=806)

Consistent trends have been found for both survival analysis tests, suggesting that trends hold with or without the inclusion of covariates and parametric assumptions.

Figure A-7 Survival analysis (Kaplan Meier estimator and Cox proportional hazards regression) outputs

The top two charts represent the speed of job outcomes achievement in for sites with low (<£2,000) or high (>£2,000) performance incentives (top left) or in £1,000 increments (top right). The bottom two charts report results from permutation testing.



The top left figure illustrates outputs from the Kaplan Meier estimator. This test suggests that proportion of job achievement is greater for clients with performance incentives of over £2,000 than for clients with performance incentives of less than £2,000. The top right figure depicts the Cox proportional hazards regression output, which reaffirms the trend found by the Kaplan Meier estimator – whilst also holding the control variables constant and depicting trends across £1,000 increments.

The Cox proportional hazards regression and permutation testing outputs are depicted in the bottom left and right figure, which respectively, from left to right, depict coefficient and odds ratio outputs. The permutation testing suggests that the performance incentive variable is responsible for the effect found in the regression output. The real performance incentive odds ratio of 1.17 (p-value<0.01), suggests that for a £1,000 increase of performance incentive, the likelihood of job start occurring increases by a factor of approximately 1.17 – whilst holding all other control variables constant.

Cox proportional hazards regression output:

Job start outcomes	Coefficient	Standard error	z	P>|z|
Performance incentive	0.159	0.044	3.610	0.000**
Site (Reference: Tower Hamlets)				
Shropshire	0.003	0.145	0.022	0.983
Haringey and Barnet	0.075	0.182	0.412	0.680
Enfield	0.936	0.150	6.253	0.000**
Economic inactivity (%) by local authority district	-0.032	0.018	-1.723	0.085
Number of unemployed people per vacancy nationwide	0.067	0.051	1.308	0.191

Significance ratings: ** = <0.01 and * = <0.05.

The Cox proportional hazards regression has been tested using variables from models 1, 2, 4 and 5 (detailed in the table above) and the same trends have consistently been observed despite the variables included. Given that gender, ethnicity and age are the demographic variables with the highest coverage for clients, they were the only demographics included in this robustness test.

Site level comparisons

1) Logistic regression model best fit – The following logistic regression model has been run with and without the site variable. The model fit of these 2 regressions has then been tested through the Log-Likelihood Ratio test.

Model variables Job start outcomes, Performance incentive, Sites, Economic inactivity (%) by local authority district, and Number of unemployed people per vacancy nationwide.

	Model without Site	Model with Site
Log likelihood	-1679.5	-1660.3
p-value	0.00**	0.00**
Pseudo R-square	0.0146	0.0261

The log likelihood results suggest better model fit with inclusion of site in the model. The subsequent site level comparison analysis uses the performance incentive regression model's sample size (n=3,533).

2) Post hoc test, Chi-Square – Two by four chi-square test assesses the overall association between two categorical variables.

Job start achievement by site	No achievement count	Achievement count
Enfield	140	74
Haringey and Barnet	263	96
Shropshire	633	165
Tower Hamlets	1835	327

Chi-square: 70.88
P-value: 0.00**

The chi-square test suggests that there is an association between sites and job start outcomes achievement (p-value<0.01) – ie that the variables are dependent on each other.

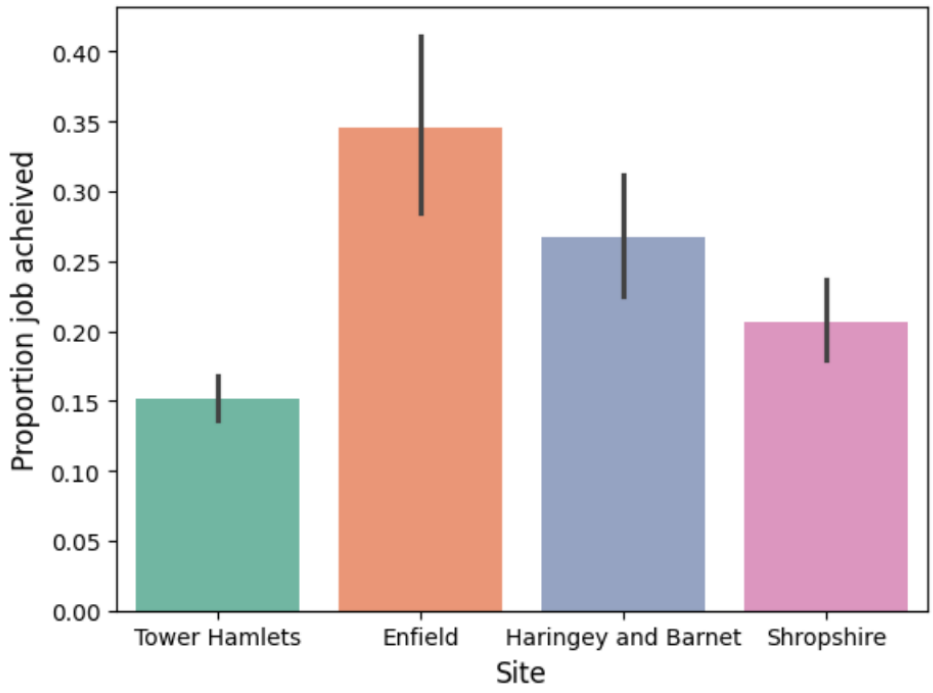
3) Post hoc test, two proportion z-test – The two proportion z-test compares the proportions of a binary outcome between two independent groups. Benjamini-Hochberg adjustments have been applied to account for multiple comparisons. This reduces the probability of false positive results by adjusting and lowering the significance levels.

Comparison	z	p-value	Adjusted p-value
Enfield vs Haringey and Barnet	1.99	0.05	0.05
Enfield vs Shropshire	4.25	0.00**	0.00**
Enfield vs Tower Hamlets	7.25	0.00**	0.00**
Haringey and Barney vs Shropshire	2.28	0.02*	0.03*
Haringey and Barnet vs Tower Hamlets	5.45	0.00**	0.00**
Shropshire vs Tower Hamlets	3.60	0.00**	0.00**

By comparing the proportions of the job start outcomes achievement between two sites, the two proportion z-test suggests that five out of the six site comparisons have statistically significant different outcomes. The comparisons are always statistically significant when Tower Hamlets is included in the comparison.

Subsequently, Figure A-8 shows that Tower Hamlets has the lowest proportion of job achievement and Enfield has the highest proportion of job achievement.

Figure A-8 Job start achievement by site (for sample size of performance incentive regression model)



Performance management dose

Logistic regression

1) Logistic regression – This regression will explore whether there is an association between performance management and job start outcomes.

Model variables: Job start outcomes, Performance management, Wind down dummy, Sites, Economic inactivity (%) by local authority district, and Number of unemployed people per vacancy nationwide.

Alongside the control variables from the main performance incentive analyses, a wind down dummy variable has been included. This variable logged clients who

were referred in the last 6 months of project service delivery, to take into account the potential wind down effect which we observed in the data.

Regression assumption tests:

Similarly to the performance incentive logistic regression model (3), no problems have been found regarding multicollinearity or model specification (via STATA's correlation, collin and link tests); there is a linear relationship between the independent variable and log odds of the dependent variable; potential outliers have been omitted; and the model has a low Pseudo R-square.

In contrast, logistic regression assumes a monotonic linear relationship between the dependent and independent variable and a monotonic linear relationship has not been found between the performance management and job start variables (see Figure 10, pg 62). This means that the regression results are compromised and are not robust as outputs could be influenced by logistic regression assumptions not being effectively met. As a result, performance management testing is self-contained, and the variable has not been applied to any of the other regression analysis.

Performance management dose logistic regression model output

Number of observations = 4,167
Log likelihood = -1957.6
p-value = 0.00**
Pseudo R-squared= 0.04

Job start outcomes	Coefficient	Standard error	z	P>|z|
Performance management	-0.015	0.003	-5.12	0.000**
Wind down dummy	-1.270	0.262	-4.84	0.000**
Site (Reference: Tower Hamlets)				
Shropshire	-0.536	0.186	-2.88	0.004**
Haringey and Barnet	0.353	0.143	2.47	0.014*
Enfield	0.639	0.200	3.19	0.001**
Economic inactivity (%) by local authority district	-0.024	0.009	-2.73	0.006**
Number of unemployed people per vacancy nationwide	-0.150	0.061	-2.48	0.013*

Significance ratings: ** = <0.01 and * = <0.05.

Although the aforementioned model has suggested an extremely weak negative statistically significant relationship between performance management and job start, this does not hold when the model also controls for performance incentive. Without controlling for performance incentive, any effect found in the above regression could be due to fluctuating outcomes metric prices. The inclusion of this control variable results in a better model fit – as suggested by the log likelihood results – and no statistically significant association between the performance management and job start outcomes variables.

Model variables: Job start outcomes, Performance management, Wind down dummy, Performance incentive, Sites, Economic inactivity (%) by local authority district, and Number of unemployed people per vacancy nationwide.

Performance management dose logistic regression model output, with performance incentive included

Number of observations = 3,533
Log likelihood = -1633.6
p-value = 0.00**
Pseudo R-squared = 0.04

Job start outcomes	Coefficient	Standard error	z	P>|z|
Performance management	-0.008	0.006	-1.28	0.200
Wind down dummy	-1.393	0.279	-5.00	0.000**
Performance incentive	0.072	0.054	1.35	0.178
Site (Reference: Tower Hamlets)				
Shropshire	-0.113	0.297	-0.38	0.703
Haringey and Barnet	0.306	0.196	1.56	0.118
Enfield	0.886	0.353	2.51	0.012
Economic inactivity (%) by local authority district	-0.007	0.010	-0.74	0.457
Number of unemployed people per vacancy nationwide	-0.007	0.085	-0.08	0.937

Significance ratings: ** = <0.01 and * = <0.05.

Further research is required to explore the relationship between performance management and job start due to the model's limitations and the complexity of investigating performance management.

16. Appendix E: Table of acronyms and glossary

Acronym	Definition
A&E	Accidents & Emergencies
BII	Big Issue Invest
CBO	Commissioning Better Outcomes
DCMS	Department for Culture, Media and Sport
DWP	Department for Work and Pensions
GO Lab	Government Outcomes Lab
HLT	The Health Led employment trials
ICB	Integrated Care Board
INDIGO	International Network for Data on Impact and Government Outcomes
IPS	Individual Placement and Support
IPS-AD	Individual Placement and Support for Alcohol and Drug Dependence
ICS	Integrated Care Systems
IT	Internet Technology
KPIs	Key Performance Indicators
LA	Local Authorities
LCF	Life Chances Fund
MHEP	Mental Health and Employment Partnership
NHS	National Healthcare Service
OBC	Outcomes-Based Contracts
OHID	The Office for Health Improvement and Disparities
RBF	Results-Based Financing
SIB	Social Impact Bond
SOP	Social Outcomes Partnership
SPV	Special Purpose Vehicle
THLD	Tower Hamlets Learning Disabilities
THSMI	Tower Hamlets Severe Mental Illness
TNLCF	The National Lottery Community Fund
VCSE	Voluntary, Community, & Social Enterprise

Block payments Payments are made for the service, regardless of outcomes. Block payments have been used in traditionally commissioned contracts in health and social care. It is payment made to a provider to deliver a specific yet broadly defined service, made on a regular basis. Typically, they do not provide incentives for improved care. Historically it is the most common payment system in the NHS.

Cap Social outcomes contracts often use caps to establish a maximum monetary limit on outcome payments. Caps can be designed in several forms. For instance, some social outcomes contracts include caps at the outcome level (ie: In Hounslow, Enhanced Dementia Care Service capped the payment of 'completion of integrated care plans' at 300 outcomes. Although they achieved more integrated care plans, they only got paid for 300.) Other contracts include caps at the participant level. In Midlands Regional Pause Hub, the outcome 'engagement with the Pause programme' can be achieved up to 10 times by one individual.

Cohort The targeted population of beneficiaries or service users.

Commissioning The cyclical process by which entities assess the needs of people in an area, determine priorities, design and contract appropriate services, and monitor and evaluate their performance. This term is used widely in the UK public sector context, but less so elsewhere. It is sometimes used interchangeably with 'contracting.'

DCMS The Department for Culture, Media and Sport (DCMS) is a department of the United Kingdom government. It hosts the Civil Society and Youth Directorate and VCSE Public Sector Commissioning Team (formerly the Centre for Social Impact Bonds), which holds policy responsibility for this policy area within UK central government. In 2016, DCMS launched the Life Chances Fund (LCF), within which it acted as the central government outcome payer.

DCMS Data Portal A dedicated data portal (created and owned by DCMS) set up for social outcomes partnerships within the LCF to capture detailed baseline and performance data for individual SOP projects. It aimed to facilitate a more streamlined application process and grant management. The portal supported outcome and payment reporting and grant management by The National Lottery Community Fund, as well as the GO Lab evaluation activity.

Delivery cost In the End of Grant Form, projects were required to report the total cost of their projects and the distribution of this cost across investment, delivery, management and evaluation and learning. In delivery cost, projects included the cost of all items related with the implementation of the intervention, such as cost

of front-line personnel, special material and licenses to deliver the programme, training costs, etc.

End of Grant Form When finishing the award, LCF projects were required to complete an End of Grant Form. In this form, projects compared the initial figures on investment, costs, and outcomes that they shared in the 'Grant Baseline Block' with the final figures on outcome achievements and actual costs. This form was stored in the DCMS Data Portal.

Intermediary Social outcomes partnerships are often supported by experts that provide specific advice. These are typically all referred to as "intermediaries" but can encompass at least four quite different roles: consultancy to develop business cases, social investment fund managers, performance management experts, and special purpose vehicles.

Investment cost Investment costs refer to the cost of items such as setting up and maintaining a Special Purpose Vehicle (SPV), expenses related with setting up the investment, the return to social investors, etc.

Investment Fund Manager Responsible for providing the project finance and managing the investment strategy on behalf of the social investors.

IPS (Individual Placement and Support) is a service that uses employment specialists in mental health teams to promote the return to work for people experiencing mental health and addiction issues. It is a strength-based approach and individually tailored support to help people find the right job with ongoing support. It is based on eight principles which includes a focus on competitive employment, zero exclusion, and providing unlimited support and integrated services regardless of diagnosis, symptoms, or substance misuse

IPS Grow is a national programme designed to support the expansion of Individual Placement and Support (IPS) services in mental health, primary care, and drug and alcohol teams across England. Includes operational support, workforce development, and tools to improve data and outcomes reporting. The programme is led by Social Finance in partnership with the Centre for Mental Health. It is funded by NHS England and Improvement (NHSE/I), the Department for Work and Pensions (DWP), and the Office for Health Improvement and Disparities (OHID).

Investment return A ratio that measures the profitability of an investment. This is typically the ratio of the income from the investment over the cost of the investment. A positive return indicates that a social investor has made a profit, while a negative return indicates that the social investor lost their capital. For

example, a project that reported an initial investment of £1,000,000 and return of £1,200,000, has achieved an investment return of 20%.

Life Chances Fund The LCF was launched as an £80m outcomes fund committed in 2016 by UK central government (DCMS) to tackle complex social problems. It provided top-up contributions to locally commissioned outcomes-based contracts involving social investment, referred to as social outcomes partnerships (SOPs). The overall fund spend of the LCF was reduced to £70m from £80m as part of the DCMS budget negotiations in September 2020. This did not affect the ability to deliver existing commitments to projects in the Fund.

Management cost In management cost, projects included the cost of items such as cost of coordination and oversight personnel, cost of performance management systems, financial management systems, cost of resources spent on governance discussions and partnership building, etc.

Outcome payment Total amount of outcome payments that could be paid to a project if all potential outcomes were achieved. Practitioners often refer to the maximum potential outcome payment as the 'contract cap' or the 'size of contract'. Also referred to as outcomes-based payments.

Outcome (outcome metrics/outcome payment triggers) The outcome (or outcome metric) is a result of interest that is typically measured at the level of service users or programme beneficiaries. In evaluation literature, outcomes are understood as not directly under the control of a delivery organisation: they are affected both by the implementation of a service (the activities and outputs it delivers) and by behavioural responses from people participating in that programme. Achieving these outcomes 'triggers' outcome payments within an outcomes contract or SOP arrangement.

Outcomes-based contract (OBC) 'Outcomes' can feature in a contractual arrangement in a range of ways. Typically, an outcomes-based contract is understood as a contract where payments are made wholly or partly contingent on the achievement of pre-defined and measured outcomes. Also known as an outcomes contract.

Outcomes fund Outcomes funds pool capital from one or more funders to pay for a set of pre-defined outcomes. Outcome funds allow the commissioning of multiple social outcomes partnerships under one structure. Payments from the outcomes fund only occur if specific criteria agreed ex-ante by the funders are met. Recent examples of outcome funds in the UK include the Refugee Transitions Outcome Fund (hosted by the Home Office), Commissioning Better Outcomes Fund and the Life Chances Fund, both administered by the National Lottery Community Fund.

Outcome metric Outcome metrics are the specific ways the commissioners choose to determine whether an outcome has been achieved. Outcome metrics often encompass a single dimension of an outcome. For example, the outcome metric for an employment outcome can be a job contract. In the Life Chances Fund, outcome metrics are referred to as 'payment triggers', as they trigger a payment for a project.

Outcome payer The organisation that pays for the outcomes in an outcomes contract or social outcomes partnership. Outcome payers are often referred to as commissioners or outcome funders.

Outcome payment Payment by outcome payers for achieving pre-agreed outcomes. Payments may be made to a special purpose vehicle or management entity in an impact bond or to service providers in other forms of outcomes-based contracts.

Provider Also known as service provider, service delivery organisation or delivery partner. A provider can be a private sector organisation, social enterprise, charity, NGO or any other legal form.

Rate Card A schedule of payments for specific, pre-agreed outcome measures that a commissioner (outcome payer) is willing to make for each user, cohort or specified improvement that verifiably achieves each outcome.

Service provider Service providers are responsible for delivering the intervention to participants. A provider can be a private sector organisation, social enterprise, charity, NGO, or any other legal form.

Service users See Cohort.

Social impact bond (SIB) See SOP

Social investor cost In the End of Grant Form, projects were required to report the total cost of their projects and the distribution of this cost across investment, delivery, management and evaluation and learning. In investment cost, projects included the cost of items such as setting up and maintaining a Special Purpose Vehicle (SPV), expenses related with setting up the investment, the return to social investors, etc.

Social Investor (or investor) An investor seeking social impact in addition to financial return. Social investors can be individuals, institutional investors, and philanthropic foundations, who invest through their endowment. In UK SOPs, these

assets are often managed by 'investment fund managers' rather than the original investing institutions or individuals who provide the capital

Social outcomes partnership (SOP) While there is no single, universally agreed definition of social outcomes partnerships (often referred to as social impact bonds, SIBs, or social outcomes contracts, SOCs), they are best understood as cross-sector partnerships that bring organisations together in the pursuit of measurable social outcomes. Typically, SOPs are defined as contractual arrangements that have two key characteristics: (1) Payment for social or environmental outcomes achieved (an outcomes contract), (2) Up-front repayable finance provided by a third party, the repayment of which is (at least partially) conditional on achieving specified outcomes.

Special purpose vehicle (SPV) A legal entity (usually a limited company) that is created solely for a financial transaction or to fulfil a specific contractual objective. Special purpose vehicles have sometimes been used in the structuring of **social outcomes partnerships**

Target When awarded funding by the Life Chances Fund, projects had to complete a 'Grant Baseline Form'. In this form, they were asked to report baseline targets for every outcome that they were expected to achieve. These targets indicate the amount of outcomes that a project could potentially achieve in a best-case scenario.

The National Lottery Community Fund (TNLCF) TNLCF, previously legally named the Big Lottery Fund, is a non-departmental public body responsible for distributing funds raised by the National Lottery. The Community Fund aims to support projects which help communities and people it considers most in need. TNLCF managed the Life Chances Fund on behalf of DCMS.

Top-up funding An outcomes fund may provide a partial contribution to the payment of outcomes where the remainder of outcomes payments are made by another government department, local government, or public sector commissioner. In the LCF the partial contribution from DCMS 'tops up' the locally funded payment for outcomes and is intended to support the wider adoption of social outcomes partnerships (SOPs) commissioned locally.

Voluntary, community and social enterprise (VSCE) sector A 'catch all' term that includes any organisation working with social objectives ranging from small community organizations to large, registered charities operating locally, regionally and nationally

Government Outcomes Lab
Blavatnik School of
Government

University of Oxford

golab.bsg.ox.ac.uk