



Study on the benefits of using social outcome contracting in the provision of social services and interventions – a cross-country comparative assessment of evolving good practice in cross- sectoral partnerships for public value creation

Annexes

Written by Klimavičiūtė L. (PPMI), Chiodo V. (Politecnico di Milano), De Pieri B. (Politecnico di Milano),
Gineikytė V. (PPMI)
January – 2021

PPMi



**POLITECNICO
DI MILANO**

EUROPEAN COMMISSION

Directorate-General for Employment, Social Affairs & Inclusion

Directorate Social Affairs Directorate

Unit C.1 — Social Investment Unit

Contact: Monika Chaba

E-mail: Monika.Chaba@ec.europa.eu

*European Commission
B-1049 Brussels*

**Study on the benefits of using
social outcome contracting in the
provision of social services and
interventions – a cross-country
comparative assessment of
evolving good practice in cross-
sectoral partnerships for public
value creation**

Annexes

Authors: Luka Klimavičiūtė (PPMI), Veronica Chiodo (Polimi), Benedetta De Pieri (Polimi), Vaida Gineikytė (PPMI)

Scientific supervisors: Egidijus Barcevičius (PPMI), Mario Calderini (Polimi)

Acknowledgements: The authors would like to thank PPMI colleagues Laurynas Stankevičius, Gabija Skardžiūtė, Ugnė Pilkionytė, Loreta Matulevič, colleagues at Politecnico di Milano Alice Borrello, Federica Falomi, Francesco Gerli, and Dr Lisa Knoll, Dr Reinhard Millner, Dr Frédéric Dufays, Dr Niels Bosma and Dara van Buiten for their research, contributions in drafting the Annex of this report, and peer review.

Manuscript completed in February 2021

This document has been prepared for the European Commission however it reflects the views only of the authors, and the European Commission is not liable for any consequence stemming from the reuse of this publication. More information on the European Union is available on the Internet (<http://www.europa.eu>).

Luxembourg: Publications Office of the European Union, 2021

© European Union, 2021



The reuse policy of European Commission documents is implemented based on Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Except otherwise noted, the reuse of this document is authorised under a Creative Commons Attribution 4.0 International (CC-BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated.

For any use or reproduction of elements that are not owned by the European Union, permission may need to be sought directly from the respective rightholders.

PDF ISBN 978-92-76-30287-2

doi: 10.2767/412285

KE-01-21-066-EN-N

Table of Contents

Annex 1: Individual analysis of social outcomes contracting schemes ... 13

1.1. Eleven Augsburg (Germany).....	14
1.1.1. Background	14
1.1.2. Equivalent TF scheme	16
1.1.3. Outcomes measurement.....	17
1.1.4. Effectiveness	19
1.1.5. Efficiency	20
1.1.6. Design of the scheme	22
1.2. Perspektive:Arbeit (Austria)	24
1.2.1. Background	24
1.2.2. Equivalent TF scheme	26
1.2.3. Outcomes measurement.....	27
1.2.4. Effectiveness	30
1.2.5. Efficiency	31
1.2.6. Design of the scheme	32
1.3. DUO for a JOB (Belgium)	33
1.3.1. Background	33
1.3.2. Equivalent TF scheme	35
1.3.3. Outcomes measurement.....	36
1.3.4. Effectiveness	38
1.3.5. Efficiency	40
1.3.6. Design of the scheme	42
1.4. Buzinezzclub SIB (Netherlands)	43
1.4.1. Background	43
1.4.2. Equivalent TF scheme	45
1.4.3. Outcomes measurement.....	46
1.4.4. Effectiveness	47
1.4.5. Efficiency	50
1.4.6. Design of the scheme	50
1.5. BOAS Werkt (Netherlands and Germany)	52
1.5.1. Background	52
1.5.2. Equivalent TF scheme	53
1.5.3. Outcomes measurement.....	54
1.5.4. Effectiveness	56
1.5.5. Efficiency	57
1.5.6. Design of the scheme	59
1.6. Adolescent Behavioral Learning Experience (ABLE) Program at Rikers Island (USA)	60
1.6.1. Background	60
1.6.2. Equivalent TF scheme	63
1.6.3. Outcomes measurement.....	64
1.6.4. Effectiveness	66
1.6.5. Efficiency	68

1.6.6. Design of the scheme	70
1.7. The Benevolent Society Social Benefit Bond (Australia)	71
1.7.1. Background	71
1.7.2. Equivalent TF scheme	75
1.7.3. Outcomes measurement.....	76
1.7.4. Effectiveness	81
1.7.5. Efficiency	82
1.7.6. Design of the scheme	83
1.8. Mental Health and Employment Partnership (the UK)	84
1.8.1. Background	84
1.8.2. Equivalent TF scheme	88
1.8.3. Outcomes measurement.....	89
1.8.4. Effectiveness	93
1.8.5. Efficiency	95
1.8.6. Scheme design	98
1.9. JobPath (Ireland)	99
1.9.1. Background	99
1.9.2. Equivalent TF scheme	102
1.9.3. Outcomes measurement.....	105
1.9.4. Effectiveness	109
1.9.5. Efficiency	112
1.9.6. Design of the scheme	114
1.10. Provider-led Pathways to Work (the UK)	116
1.10.1. Background	116
1.10.2. Equivalent TF scheme	119
1.10.3. Outcomes measurement.....	120
1.10.4. Effectiveness	124
1.10.5. Efficiency	127
1.10.6. Design of the scheme	129
1.11. The Work Programme (UK)	131
1.11.1. Background	131
1.11.2. Equivalent TF scheme	137
1.11.3. Outcomes measurement.....	139
1.11.4. Effectiveness	144
1.11.5. Efficiency	148
1.11.6. Design of the scheme	150
1.12. The Youth Contract (UK)	153
1.12.1. Background	153
1.12.2. Equivalent TF scheme	156
1.12.3. Outcomes measurement.....	157
1.12.4. Effectiveness	160
1.12.5. Efficiency	161
1.12.6. Design of the scheme	164
1.13. Drug and Alcohol Recovery Pilots (UK)	166
1.13.1. Background	166
1.13.2. Equivalent TF scheme	168

1.13.3. Outcomes measurement.....	169
1.13.4. Effectiveness	174
1.13.5. Efficiency	176
1.13.6. Design of the scheme	178
1.14. Transforming Rehabilitation (UK).....	179
1.14.1. Background	179
1.14.2. Equivalent TF scheme	182
1.14.3. Outcomes measurement.....	184
1.14.4. Effectiveness	189
1.14.5. Efficiency	193
1.14.6. Design of the scheme	195
1.15. Troubled Families – Phase II (UK).....	196
1.15.1. Background	196
1.15.2. Equivalent TF scheme	199
1.15.3. Outcomes measurement.....	200
1.15.4. Effectiveness	208
1.15.5. Efficiency	210
1.15.6. Design of the scheme	210
Annex 2: Methodology	212
1.1. Task 1: Long-list of completed and diverse SOC.....	214
1.2. Task 2: Shortlist of SOC operations for which causalities of impact can be attributed.....	218
1.3. Data collection process.....	225
1.4. Task 3: Mapping and review of outcomes measurement methods	227
1.5. Task 4: Effectiveness and added value	228
1.6. Task 5: Analysis of efficiency.....	229
1.7. Task 6: Policy recommendations	231
1.8. Limitations	232
Annex 3: Long list of SOC schemes.....	233
References	241

LIST OF ABBREVIATIONS

AA Pilots	Activity Agreement Pilots
ABLE	Adolescent Behavioral Learning Experience
AGE	Apprenticeship Grant for Employers
AMS	Austrian Public Employment Service (Arbeitsmarktservice)
BII	Big Issue Invest
CBO	Commissioning Better Outcomes Fund
CCG	Clinical Commissioning Group
CMP	Condition Management Programme
CPA	Contract package area
CRC	Community Rehabilitation Companies
DCLG	Department for Communities and Local Government (UK)
DEASP	Department of Employment Affairs and Social Protection (Ireland)
DfE	Department for Education (UK)
DOC	The New York City Department of Correction (USA)
DWP	Department for Work and Pensions in the UK
EFA	Education Funding Agency in the UK
ER	Employment rate
ESA	Employment and Support Allowance
ESF	European Social Fund
EU	European Union
FACS	Family and Community Services
FND	Flexible New Deal
GCSE	General Certificate of Secondary Education
HM	Her Majesty (UK)
HMI	Her Majesty's Inspectorate (UK)
HMPPS	Her Majesty's Prison and Probation Service (UK)
HMRC	Her Majesty's Revenue and Customs (UK tax authority)
IB	Incapacity Benefit
IES	Institute of Employment Studies
IFP	Intensive Family Preservation
IPS	Individual placement and support
IS	Income Support
IT	Information technology
ITT	Invitation to Tender
JLD	Jobseekers Longitudinal Dataset
JSA	Jobseeker's Allowance
LA	Local authority
LASARS	Local Area Single Assessment and Referral Service
LBC	London Borough of Camden

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

LTU	Long-term unemployed
MHCLG	Ministry of Housing, Communities and Local Government (UK)
MHEP	Mental Health and Employment Partnership
MoJ	Ministry of Justice (UK)
MPL	Minimum performance level
MRT	Moral reconnection therapy
MSG	Minimum service guarantee
NAO	National Audit Office (UK)
NatCen	National Centre for Social Research (UK)
NCCIS	National Client Caseload Information System (UK)
ND	New Deal
ND25PL	New Deal for those aged 25+
NDDP	New Deal for Disabled People
NDTMS	National Drug Treatment Monitoring System
NDYP	New Deal for Young People
NEET	Not in education, employment or training
NFA	No fixed abode
NGO	Non-governmental organisation
NHS	National Health Service (UK)
NIR	Non-intervention rate
NPO	Non-profit organisation
NPS	National Probation Service (UK)
NSW	New South Wales (Australia)
OECD	Organisation for Economic Co-operation and Development
OGRS	Offender Group Reconviction Scores
OOHC	Out-of-home care
PbR	Payment-by-Results
PL	Provider-led
PSI	Policy Studies Institute
PSM	Propensity score matching
RBD	Readmission bed days
RF	Resilient Families
SARA	Safety and Risk Assessment
SBB	Social Benefit Bond
SE	Social enterprise
SIB	Social impact bond
SME	Small and Medium Enterprise
SOC	Social outcomes contracting
SOF	Social Outcomes Fund
SPV	Special purpose vehicle
SSP	Statistically significant points

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

TBS	The Benevolent Society
TF	Traditionally financed
TOP	Treatment Outcome Profile
TR	Transforming Rehabilitation
TTG	,Through the Gate' support
TUPE	Transfer of Undertakings (Protection of Employment)
UWV	Employee Insurance Agency in the Netherlands (Uitvoeringsinstituut Werknemersverzekeringen)
VAT	Value Added Tax
WFI	Work-focused interview
WHP	Work and Health Programme
WP	Work Programme
YC	Youth Contract

Annex 1: Individual analysis of social outcomes contracting schemes

This annex provides detail of 15 selected social outcomes contracting (SOC) schemes, and discusses the following features of each scheme:

- background;
- comparable traditionally financed (TF) intervention;
- outcomes measurement methods;
- effectiveness and added value, compared to those of the equivalent TF scheme;
- efficiency; and
- scheme design.

In the background section for each scheme, we discuss the reasons for selecting outcomes-based contracts to deliver the intervention; the main goals and specific objectives of the intervention; the services that were offered to participants in the intervention; the parties involved in the delivery of the scheme; the payment arrangements; and, briefly, the outcomes achieved.

Following each scheme's background information, we describe the TF scheme chosen for comparison, together with elements that are comparable between the two programmes, as well as issues relating to the comparability of the paired schemes.

We then detail the methods used to measure the SOC scheme's outcomes, together with the indicators used to determine whether or not payments should be issued to providers and investors. This section also covers non-payment-related evaluations and indicators that were used to assess the broader impacts of these programmes, as well as the strengths and weaknesses of the outcomes measurement process.

In the sections relating to the effectiveness of each SOC scheme, we elaborate further on the outcomes achieved by the scheme. Where possible, we provide a comparison of similar outcomes achieved by the scheme's TF equivalent. We also detail the reasons for differences in performance, and determine whether each programme managed to achieve the targets set by its respective commissioner. We further discuss the added value of each SOC scheme, particularly in relation to its ability to deliver an innovative service and meet the needs of users. In those cases for which the data show that SOC models did not bring added value in these areas, we outline the reasons why (e.g. the presence of perverse incentives and cherry picking or creaming). Lastly, we identify whether the programme was replicated or scaled after its termination and, where possible, indicate what changes were implemented during the scheme's replication or scaling.

After discussing each programme's effectiveness, we turn our attention to their efficiency. Where possible, we first compare the SOC and TF schemes' intervention costs, including total costs, cost per participant and cost per outcome achieved. It is important to note that we do not control for participant characteristics, the different time periods in which the interventions were delivered and other factors that might have affected the comparability of each pair of SOC and TF programmes. We then assess whether the operating costs, which we define as the cost of setting up, managing and evaluating the programme, are higher for the SOC scheme than for its TF equivalent. Finally, we consider the design of each scheme, outlining its specific design features and what stakeholders say they would change if they were to re-design a similar programme again. We end each scheme's section with a discussion of the impact the design of the scheme had on third-sector organisations and social enterprises.

1.1. Eleven Augsburg (Germany)

1.1.1. Background

In September 2013, the Bavarian State Ministry of Labour, Social and Family Affairs and Integration launched the first social impact bond (SIB) in continental Europe. Named 'Eleven Augsburg/Jugendliche mit Perspektive (JuMP)', the project focused on **difficult-to-reach youth** in the districts of Augsburg and Aichach-Friedberg. The target group was at risk of developing a drug or gambling addiction, becoming violent and being imprisoned. The specific **objective** of the SIB was to bring at least 20 difficult-to-reach juveniles into work or apprenticeships (in both cases subject to social insurance), lasting at least nine months. There appears to have been some ambiguity with respect to the criteria for eligibility to participate in the programme. Six months after the start of the project, the commissioner specified that service providers were not allowed to work with youth who had reached out to local employment services.¹

The project's **activities** were not defined *ex-ante*, in order to allow for creative and flexible solutions². The flexibility inherent in the design of SIBs was one of the main attractions of the funding model. According to the evaluation that accompanied the project, the main **purpose** of the SIB was not to generate public savings, but to develop effective interventions in social areas which previously lacked solutions that were proven to work³. As the commissioner put it, "SIB is attractive in experimental situations... if I face an unsolved problem and I am looking for new ways to address it"⁴.

The SIB was implemented by various **stakeholders** (see the table below). The process was largely initiated by Juvat gGmbH, a non-profit subsidiary of the Benckiser Foundation Future⁵ that specialised in setting up outcomes-based project designs in the social sector, and introduced the SIB instrument to the Bavarian State Ministry of Labour, Social and Family Affairs and Integration. Juvat gGmbH served as the intermediary. The SIB also involved multiple investors, mostly from philanthropic organisations, as well as four service providers. Although the service providers were expected to work as a team with regard to individual cases, this turned out to be problematic throughout the process. It became clear that one service provider should lead the process and ask others to join in where necessary⁶. Nevertheless, the involvement of all service providers was necessary because some service providers had better contacts with the target group than others.

By December 2015, when the project ended, the goal of finding employment or apprenticeships lasting nine months for at least 20 hard-to-reach juveniles was **successfully achieved**, and the commissioner issued **payments** to the investors, including interest of 3%. In total, around 100 adolescents had been approached by the social service organisations, and 69 had participated the programme⁷.

¹ Scheck, Barbara (2017). Begleitevaluation des ersten deutschen Social Impact Bonds. JuMP Jugendliche mit Perspektive, 31.

² Scheck (2017), 27.

³ Scheck (2017), 2.

⁴ Jebabli, Jonas (2016). Social Impact Bonds: Collaboration Across Institutional Domains. Master Thesis. Vienna University of Economics and Business, 37, translated by the authors.

⁵ Scheuerle, Thomas & Anja Nieveler (2017). Implementing Social Impact Bonds in Germany – Challenges for Pay-for Success Models in the German Welfare System. German discussion paper. Interreg Alpine Space AlpSib. European Regional Development Fund, 3.

⁶ Scheck (2017), 34.

⁷ Scheuerle & Nieveler (2017), 7.

Table 1. Summary of Eleven Augsburg SIB

Personally targeted social service	Labour market-related
Target population	Unemployed youth that are difficult to reach, in the region of Augsburg, Bavaria
SOC scheme type⁸	Intermediated SIB
Payment model⁹	Binary
Cost	EUR 250,000 ¹⁰
Commissioner(s)	Bavarian State Ministry of Labour, Social and Family Affairs and Integration
Social service provider(s)	Main local social service providers: <ul style="list-style-type: none"> • Regionales Bildungszentrum Eckert gGmbH (Eckert Schulen) (NGO) • Kinder-, Jugend und Familienhilfe Hochzoll (NGO) “Two external organisations specialising in youth and educational support” ¹¹ : <ul style="list-style-type: none"> • apeiros e. V. (NGO) • Joblinge gAG München (NGO)
Investor(s)	<ul style="list-style-type: none"> • BHF-BANK-Stiftung (philanthropic organisation) • Eberhard von Kuenheim Stiftung (philanthropic organisation) • BMW Stiftung Herbert Quandt (philanthropic organisation) • BonVenture gGmbH (private investor)
Intermediary	Juvat gGmbH, “a non-profit subsidiary of the Benckiser Foundation Future” ¹² (non-profit organisation)
Evaluator(s)	<ul style="list-style-type: none"> • Evaluation of success: Law firm Spiegel RA WP StB Partnerschaft mbB (contract take-over from law firm Dr. Mohren & Partner, which was dissolved in 2015) • Accompanying evaluation: University of Hamburg, Prof. Dr. Barbara Scheck

⁸ Three types of SIB can be distinguished. In a **direct SIB**, a delivery contract is signed between the outcomes-payer and service provider. The service provider carries the risk if results are not achieved, as the investor lends funds and is repaid first with a loan secured against the enterprise’s assets. **Intermediated SIBs** also involve an intermediary to liaise between investors, service providers and the commissioner. The service delivery contract is signed between the outcomes payer and an investor-owned special purpose vehicle (SPV), which contracts the service provider, supports the performance management process and refines the financial model. The risk is shared within the SPV among the investors, and repayment depends on the SIB’s structure. A **managed SIB** is signed between the outcomes-payer and the prime contractor (usually an intermediary), who usually manages the entire process. The investors carry the risk but at the same time, they obtain profits if the results are achieved or exceeded. Please find more information on the conceptual framework developed in the Inception report.

⁹ In **binary** models, the provider has to achieve an absolute target, and no payment is granted for achieving lesser results. This can be contrasted with **frequency schemes**, in which rewards are staggered according to the agreed frequency of results, with payments increasing as results increase. Another model is a hybrid grant, in which the cost of delivering a service is funded, but additional payments are awarded as bonuses if ‘additional’ impacts are demonstrated at the end of a programme. Please find more information on the conceptual framework developed in the Inception report.

¹⁰ Juvat (2016), p. 2.

¹¹ Jebabli (2016), 27.

¹² Scheuerle & Nieweler (2017), 3.

1.1.2. Equivalent TF scheme

Eleven Augsburg is compared with ‘AMA-Zukunft’,¹³ a project funded by Germany’s Federal Ministry of Labour and Social Affairs and the European Social Fund (ESF). The two schemes share the following similarities:

- Both schemes were delivered by the same service provider (Regionales Bildungszentrum Eckert gGmbH [Eckert Schulen]);
- the provider delivered the same activities to participants in each scheme (although other service providers were also present during the SIB, delivering other tailor-made services);
- the schemes targeted the same group of participants (unemployed youth, though importantly, the SIB targeted youth that were considered the most “difficult to reach”); and
- both schemes had the same goal (to get young unemployed people into work or training).

The social service provider of ‘AMA-Zukunft’ employs a project coordinator, several social pedagogues and a psychotherapist. It runs a boxing club, and offers counselling on nutrition and debt. In addition, it conducts series of workshops for potential and existing employers and employees, in order to reduce the potential for conflict. The interventions are based on individually agreed target-setting and continuous low-threshold contact options in order to build trust.

The ESF-project AMA-Zukunft began in 2016 under the ESF-programme Integrationsrichtlinie Bund (en. *Federal Integration Guideline*), and was extended in 2019 until 2020. For a summary of stakeholders, please see the table below.

Table 2. Comparability of Ama-Zukunft (TF) with Eleven Augsburg SIB

	Same as SOC	Comments
Personally targeted social service	Yes	Labour market-related
Target population	Yes/No	Both programmes targeted unemployed youth, but the SIB targeted youth that were most “difficult to reach”
Number of participants	No	SOC: 69 (roughly 23 per year) ¹⁴ TF: 296 (roughly 59 per year) ¹⁵
Location	Yes	Augsburg
Cost	No	SOC: EUR 250,000 ¹⁶ for 2013-2015, with an annualised cost of roughly EUR 85,000

¹³ Retrieved from: <https://www.esf.de/portal/DE/Foerderperiode-2014-2020/ESF-Programme/bmas/IRL-Bund-Projekte/isa/ama-zukunft.html>

¹⁴ Scheuerle & Nieveler (2017), 7.

¹⁵ Personal communication with TF commissioner, 2020-07-27.

¹⁶ Juvat (2016), 2.

	Same as SOC	Comments
		TF: EUR 1.63 million ¹⁷ for 2015-2019, with an annualised cost of roughly EUR 326,000;
Commissioner(s)	No	SOC: Bavarian State Ministry of Labour, Social and Family Affairs and Integration TF: Integrationsrichtlinie Bund (co-financed scheme by the European Social Fund and Federal Ministry of Labour and Social Affairs)
Social service provider(s)	Yes/No	One of the service providers – Regionales Bildungszentrum Eckert gGmbH (Eckert Schulen) (NGO) – was the same in both the SOC and TF, but the SOC also included additional service providers (see 1.1.1).

1.1.3. Outcomes measurement

Two evaluations were conducted for Eleven Augsburg. One concerned the outcomes upon which payment was conditioned. This was carried out by the law firm Spiegel RA WP StB Partnerschaft mbB. The other evaluation accompanied the SIB with the aim of identifying success factors that could inform future projects. This was carried out by Dr Barbara Scheck of the University of Hamburg¹⁸. Both evaluations are briefly summarised below.

Measurement of outcomes relating to payment mechanisms

Method

The law firm Law firm Spiegel RA WP StB Partnerschaft mbB counted the number of juveniles who had been in work or apprenticeship for at least nine months, on the basis of employment contracts. In addition, the evaluator relied on information provided by the job centre to assess whether the “difficult to reach” criterion applied. The age of participants was determined using information from their passports. The success criterion in Eleven Augsburg as binary, based upon a “contractual ‘all-or-nothing’ approach regarding the disbursement”¹⁹.

Outcomes and indicators

The outcomes concerned placing hard-to-reach, disadvantaged youth (younger than 25 years-old, not in school, education or training, without any school-leaving qualification, and without employment) into apprenticeships or employment (both subject to social security contributions). The target was to place 20 hard-to-reach young people into work or apprenticeships lasting at least nine months.

As a success indicator, the target of placing 20 young persons into jobs or training was not very well justified. The evaluation report states that “clear criteria for the selection of the

¹⁷ Personal communication with TF commissioner, 2020-07-27.

¹⁸ Scheck (2017).

¹⁹ Jebabli (2016), 28

indicators were considered missing²⁰. The criteria were negotiated between the commissioner and the intermediary. The investors and service providers did not have much say in the setting of the target²¹. The commissioner stated in an interview that “it was decisive for [them] to have a measurable target with a clearly defined and explicit success criterion, which is purely objective, and does not entail estimations”²². This was also important because the commissioner did not prescribe any specific ways in which the service providers were asked to deliver services, so an objective outcome metric was required to track the performance of different providers. The service providers would, however, have preferred the broader involvement of qualitative factors to achieve a more comprehensive assessment of the participants’ wellbeing, including their quality of life.²³

Other evaluations

The accompanying evaluation drew on various reports published by the project’s stakeholders, as well as 17 semi-structured interviews, which were carried out shortly after the project began and once the project had ended²⁴. Often, multiple representatives from the same organisation took part in the interviews. The topics focused mostly on suggestions as to how to improve the project.

Barriers to and enablers of the measurement process

The stakeholders noted a few disadvantages in the way outcomes were defined and measured. First, participants who dropped out of the study could not be counted as successes (and hence no payment could not be claimed for them), even if those participants later returned to the programme.²⁵ This may have led to providers limiting the resources dedicated to returning participants.

Second, the providers were not used to reporting outcomes via quarterly reports. Investors, in turn, felt that asymmetries existed in information about the ways in which outcomes were tracked because the process was managed by an intermediary, and thus the investors lacked knowledge about the activities of the service providers.²⁶

Table 3. Summary of the measurement process in Eleven Augsburg

Measurement methodology	
Methods relating to payment mechanisms	Self-reported data
Other evaluations	Accompanying evaluation - qualitative
Control group	No
Causality of impact	No
Enabling factors and strengths of the evaluation process	n/a

²⁰ Scheck (2017), 31, translation by the authors.

²¹ Jebabli (2016).

²² Jebabli (2016), 40, translation by the authors.

²³ Jebabli (2016), 62.

²⁴ Scheck (2017), 7.

²⁵ Scheck (2017).

²⁶ Jebabli (2016).

Barriers to and weaknesses of the evaluation process	No experimental method, no control group Inability to count returning dropouts as successes may have led to limited services being given to these participants Burdensome and unfamiliar process for service providers; information asymmetries for investors
Governance of measurement	
Independence of evaluator	Yes
Evaluator	University of Hamburg (“accompanying evaluation”) Law firm Spiegel RA WP StB Partnerschaft mbB (contract take-over from law firm Dr. Mohren & Partner, which was dissolved in 2015) (“evaluation of success”)

1.1.4. Effectiveness

Achievement of outcomes

Both AMA-Zukunft and Eleven Augsburg succeeded in achieving their targets. In fact, Eleven Augsburg exceeded the agreed target at an operative level, as 22 hard-to-reach juveniles were hired. However, due to the strict application of the criteria regarding proof of employment, only 20 jobs were counted for the purposes of the SIB²⁷. AMA-Zukunft has not yet ended, but the commissioner of the scheme provided data regarding the current results of the scheme. According to these data, at the end of 2019, the scheme had reached out to 296 juveniles, and 136 had been successfully hired or placed in training – a success rate of around 46%. This means that the project is hitting its targets and compares favourably with Eleven Augsburg, which achieved a success rate of 29% (based on the 20 juveniles included in the official count)²⁸. Nevertheless, it is important to bear in mind that AMA-Zukunft receives greater funding, which could help to explain the difference in success rates. Furthermore, Eleven Augsburg targeted youth who were more disadvantaged than those addressed by AMA-Zukunft.

The services provided to juveniles under the SIB were the same as those provided afterwards by the service provider engaged in AMA-Zukunft. Thus, according to service provider interviewed, the needs of beneficiaries are equally well met under both the SIB and TF projects. Provider I said: “Our pedagogical manager would rather chop off his hand than treat the beneficiaries differently [depending on the funding stream].” However, another provider noted that the conditions of the SIB had some effect on the work of the providers – particularly the **eligibility criteria, which put pressure upon service providers** to find very difficult-to-reach youngsters.

There are no data concerning user satisfaction under either the SIB or the equivalent TF scheme.

Benefits and drawbacks of SOC compared with TF

In terms of effectiveness, the services provided to juvenile beneficiaries under SIB were the same as those provided afterwards, by the service provider engaged in AMA-Zukunft. Thus, as noted above, the interviewed service provider believed that beneficiaries’ needs were

²⁷ Juvat (2016), 3

²⁸ Interview with TF commissioner.

met equally well under both the SIB and TF projects. However, as noted above, the conditions of the SIB had some effect on the work of the providers, particularly the **eligibility criteria, which put pressure upon service providers** to find very difficult-to-reach youngsters. There are no data about user satisfaction under either the SIB or the equivalent TF scheme.

One service provider noted that, although the SIB design forced them to work with extremely difficult cases, the **flexibility** they had in designing and implementing the intervention, as well as in spending the funds, was very useful. This provider had experience in classic welfare state budget funding, where providers have to apply for funding for all extra expenses (such as solicitation materials). These processes are time-consuming and, in his opinion, reduce the effectiveness of the intervention, as providers have less time to spend working directly with beneficiaries. He stated: "I had my own pot of money that I could administer myself under the SIB. It was very relaxed in terms of daily working procedures. I could choose where to spend the money. This was gigantic!"

Another provider noted that the SIB encouraged providers to think outside the box, since it allowed them to design interventions from a **long-term perspective**. Thus, the provider thought that SIBs are particularly well suited to prevention projects, in which one has to consider the long-term effects of the intervention.

The main drawback of the SIB mentioned by stakeholders, particularly providers, concerned the definition of the **selection criteria** and the risks related to **creaming and cherry picking**. To prevent creaming, the commissioner intervened during the implementation of the SIB in order to better define the criterion of "difficult to reach". In the course of the interim evaluation, the criterion used to define "difficult to reach" was further specified that participants could not previously have contacted job centres and welfare youth offices, in order to guarantee that the intervention filled a gap in the existing institutional welfare infrastructure²⁹. Thus, some of the participants previously engaged had to be excluded from the target group³⁰. Notwithstanding this, service providers continued to work with these juveniles at their own expense.

Scalability and replicability

The same intermediary conducted another SIB in Austria, Perspektive:Arbeit (presented in detail in the following chapter), but the intervention carried out in Germany was not replicated.

1.1.5. Efficiency

Intervention costs

The table below presents a cost comparison between Eleven Augsburg SIB and its equivalent TF intervention, AMA-Zukunft. It is important to note that the estimates below do not account for various factors that could have influenced the outcomes. For example, although both projects target young unemployed people who are difficult to reach, we lack information to identify whether the project participants share key socio-demographic characteristics such as education, work experience and others, which might affect their prospects of finding a job. Furthermore, Eleven Augsburg SIB was implemented at a time when unemployment in Germany was slightly higher than during the subsequent years during which the AMA-Zukunft project was initiated. This more favourable economic climate may or may not have made it easier for AMA-Zukunft participants to find jobs than for the

²⁹ Interview with intermediary and interviews with service providers.

³⁰ Jebabli (2016), 48, and interviews service providers.

SIB participants. On the other hand, a favourable economic situation might result in the formation of an even harder to reach target group, as other potential clients might already have found jobs due to the prosperous economic environment. Moreover, although the activities delivered by one of the service providers were the same under both the SOC and the TF, the SIB also included additional activities involving other social providers. Lastly, the SIB and TF projects were commissioned by different government bodies. Different commissioners could result in different performance management, which could also affect the outcomes of either project. All of these, as well as other potential factors – in addition to the financing mechanism – could have impacted the number of jobs achieved.

Overall, the cost per participant who acquired a job is slightly lower for the AMA-Zukunft project, compared with the SIB. Although a total of 22 participants in the SIB intervention were initially reported as having found jobs (compared with the target of 20), two of these cases ultimately did not satisfy the criteria for success, due to insufficient documentation to prove that these participants were “difficult to reach”. If these extra two participants had been counted as successes, this would have brought the cost per job down to EUR 11,705 – a level roughly equal to that of AMA-Zukunft. However, according to the SIB service provider interviewed, the SIB funding was also insufficient to cover project costs and the provider cross-subsidised from other funding streams, amounting to roughly one annual salary of a social pedagogue (about EUR 48,000). If taken into account, these costs would further increase the cost per participant who found a job for the Eleven Augsburg SIB relative to AMA-Zukunft. The costs indicated below do not include the administrative expenses of the Federal Ministry of Labour and Social Affairs of Germany, such as the cost of public tendering or project oversight.

Table 4. Cost comparison between Eleven Augsburg and Ama-Zukunft (EUR)

	Eleven Augsburg (SIB) 2013-2015	AMA-Zukunft (TF) ³¹ 2015-2019
Participants	69 ³²	296
Participants who found jobs	20 ³³	136
Total cost	257,500 ³⁴	1,629,480
Cost per participant	3,732	5,505
Cost per participant who found a job	12,875	11,981

Source: elaborated by PPMI, based on available information.

Note: the costs have not been adjusted for inflation because yearly breakdowns are not available. Such adjustment would further increase the costs of the Eleven Augsburg SIB relative to those of AMA-Zukunft.

Operational costs

In addition to the project costs specified above, both projects also entail the operational costs required to set up and oversee the interventions. For the SIB, the total operational costs are unknown, although the service provider involved in both interventions believes

³¹ The information on the number of participants, job outcomes and the cost per participant were provided by the commissioner of the intervention. Total cost and the cost per job were derived by PPMI, based on the information provided.

³² Scheuerle & Nieveler (2017), 7.

³³ Juvat (2016), 3.

³⁴ Calculated by PPMI, given that the initial cost of SIB was EUR 250,000 and the 3% interest rate for the entire duration of the project was paid out to investors (Juvat (2016), 2).

them to be higher than for the TF project. According to the commissioner of the TF project, the operational costs for the AMA- Zukunft intervention amount to an additional cost of EUR 249 per beneficiary, amounting to a total of roughly EUR 74,000 over the duration of the project. Below, we break these costs down into set-up and oversight, providing comparisons between the two schemes where possible.

Detailed information is not available with regard to set-up costs, but **the process used to set up the SIB appears to have been lengthier than that for the TF intervention.** Contract negotiations between the commissioner and the intermediary for the Eleven Augsburg SIB took between a year and a year-and-a-half. The evaluation report reveals a “significant need for coordination in the beginning of the project.”³⁵ According to the intermediary interviewed, the SIB’s set-up costs were fully funded by the Benckiser Foundation, in order to pilot the innovative SIB funding model in Germany.³⁶ Nevertheless, these costs did not include remuneration for the working hours spent by the commissioner on contract negotiations. After the project was set up, the Benckiser Foundation hired a project coordinator. In 2014, project-related activities were equivalent to three-quarters of the coordinator’s full-time work, and decreased to roughly one-third of a full-time-equivalent in both 2015 and 2016.³⁷ Some of these coordination tasks included, for example, identifying suitable service providers. Given that the SIB was treated as an experiment, it could bypass standard tendering procedures.³⁸ By contrast, the AMA-Zukunft project involved a public tendering procedure that took roughly a year.

Not only the set up costs, but also **oversight appears to have been more strenuous under the SIB.** This is mostly due to negotiations between the service provider and the commissioner to determine which juveniles counted as ‘difficult to reach’. Otherwise, oversight was similar under each scheme. Both schemes required the documentation of outcomes (juveniles placed into jobs) and a financial statement. In each scheme, the commissioner checked to make sure that the juveniles placed into jobs no longer claimed unemployment benefits. The financials of the ESF-scheme also involved an administrative lump-sum in addition to monthly salaries, which could be used on a flexible basis. No additional payments were made for administrative expenses under the SIB, because the funds provided under the SIB could be used flexibly at the providers’ discretion.

1.1.6. Design of the scheme

Design features

From a design viewpoint, perhaps the main strength of this SIB was the intermediary’s ability to **bring together multiple investors**, thus in turn lowering the risk for each investor. Despite the involvement of multiple investors, the SIB nevertheless created **substantial risks for investors due to its binary criterion for success**: “the investors complained about an overly low risk-return ratio, given the high risk of default related to the successful mediation of the target group. They proposed instead, staggered payments or a more reasonable split of the economised savings on the part of the government.”³⁹ In fact, despite the SIB’s success, service the provider interviewed doubted that the investors would participate in another SIB with the same binary success criterion, because it posed too much risk. This was not only due to the binary success criterion, but also the rate of return (3%), which was considered by the investors to be too low, given the risks associated with the SIB.

³⁵ Scheck (2017), 34, translation L.K.

³⁶ Interview with the intermediary.

³⁷ Email from Niklas Ruf from the 10th of July 2020, translation LK.

³⁸ Scheck (2017), 29.

³⁹ Jebabli (2016), 41.

The commissioner refused the investors' demands for staggered payments because it wanted to **avoid perceptions of the private sector profiting at the expense of society**.⁴⁰ To avoid such perceptions, the commissioner also insisted that all investors had to be philanthropic or non-profit organisations.⁴¹ For the same reason, the interest rate of 3% was seen more as an “inflationary adjustment”⁴² rather than a return on investment.

Although the SIB brought together multiple investors, this was not the case for the commissioner. In theory, SIBs should avoid the so-called ‘**wrong pockets**’ challenge.⁴³ In the case of the Eleven Augsburg SIB, however, such considerations did not play a role. The scheme involved only one commissioner, which also served as the ultimate funder. One would expect the commissioner to be the government agency that incurs the greatest costs associated with youth unemployment, because it stands to gain the most from a successful intervention. However, the commissioner in the Eleven Augsburg SIB – the Free State of Bavaria – was not responsible for the target group, and the costs associated with this target group were largely incurred by the federal government or the municipalities. Nevertheless, the Bavarian State took up the project because it wanted to experiment with the SIB funding model and because the Benckiser Foundation was actively engaged. Therefore, while in theory a SIB should seek to solve the “wrong pockets” challenge, in practice other considerations drove the intervention.

With regard to other lessons learned about the SIB's design features, multiple stakeholders believed that the **overhead costs associated with the SIB should be integrated into the total volume of investment**. This would make it easier to judge whether or not the SIB was efficient. Furthermore, the commissioner noted that it would be important to monitor participants not only until they find employment, but also to **continue providing support once participants are in employment**—as was done in the subsequent AMA-Zukunft project. Lastly, the service provider interviewed said they would have preferred to negotiate directly with the commissioner without the involvement of the **intermediary**, because the provider believed they could better represent their own interests.

Impact on third-sector organisations and social enterprises

The role of service providers—all of which were NGOs in the Eleven Augsburg SIB—is also interesting to consider. Overall, **the service providers seem to have benefitted from their experience in the SIB**. A representative of one provider reported that, due to the SIB, the organisation had become part of a new impact investment community, which expanded its **network** and **future opportunities**. The organisation was invited to participate in various research projects, some of which also benefitted the project beneficiaries. According to the interviewee, “for me personally, [the SIB] provided a lot, but for the beneficiaries it didn't.”⁴⁴ The interviewee was convinced that the beneficiaries receive the same kind of social work under any funding scheme.

Despite the ultimate benefits of the project for the NGOs, the service providers initially expected that the SIB would unlock continuous funding without the need for repeated applications by applying the ‘**revolving fund**’ principle, as this was how the intermediary also initially envisaged the project. With a revolving fund, the investors would have invested their funds into the next project cycle in the event that the SIB was successful, with no need for the service provider to re-apply for funding. Furthermore, the commissioner would only pay the financial incentive (3% interest rate), and not the whole investment sum plus the

⁴⁰ Jebabli (2016), 41.

⁴¹ Jebabli (2016), 41.

⁴² Juvat (2016), 1.

⁴³ Kohli, J., D. J. Besharov, and K. Costa (2012). Social Impact Bonds 101: Defining an Innovative New Financing Tool for Social Programs. Retrieved from: <https://www.americanprogress.org/issues/general/news/2012/03/22/11238/social-impact-bonds-101/> for American Progress Center

⁴⁴ Interview with the service provider.

interest. Nevertheless, this situation was not ultimately realised. While it is not clear exactly why, it is possible that investors were hesitant to commit to long-term funding when the project might not align with their priorities in the future.

1.2. Perspektive:Arbeit (Austria)

1.2.1. Background

In September 2015, the Austrian Federal Ministry of Labour, Social Affairs and Consumer Protection, in cooperation with the state of Upper Austria and the Austrian Federal Ministry of Education and Women's Affairs, launched the first SIB in Austria, which focused on **domestic violence** in Upper Austria⁴⁵. The pilot SIB aimed to economically empower women affected by violence, so that they could permanently leave abusive relationships.

The main **objective** of the project was to place at least 75 survivors of domestic violence in a job with at least 20 hours of employment per week, for at least 12 months during the term of the project. The job should be subject to social insurance contributions and pay a living wage (i.e. EUR 19,500 gross salary). Women could join the project on a voluntary basis if they:

- had been in contact with a women's shelter or a Center for Protection Against Violence (Gewaltschutzzentrum) in Upper Austria within the last 24 months;
- were legal residents of Austria, of working age, and had valid work permits;
- did not earn a living wage or were at risk of losing their jobs.

The commissioner's **initial motivation** for the scheme was the desire to pilot the SIB instrument in an Austrian context, in order to test outcomes-based commissioning. While various social areas were considered (including persons not in education, employment or training), survivors of domestic violence were ultimately chosen as the focus for the SIB because this target group would not otherwise have received any new services, based on the annual budget allocated. Hence, it was an opportunity to gather additional financial support and pilot new services for this underserved group.

Project **activities** included:

- developing and providing individual counselling and assistance;
- offering protection, housing and supporting mobility;
- support for childcare;
- ensuring safe and sensible structures in the workplace;
- facilitating education and training, career guidance and job placement beyond existing services.

The counselling and placement process was tailor-made on an individual basis, whereby the appropriate type of support was determined in collaboration with existing services and programmes offered by the Austrian Public Employment Service (Arbeitsmarktservice – AMS) and women's organisations. Social service providers also collaborated directly with companies to place women on the primary job market, create new job opportunities, and raise awareness within the business community about domestic violence.

⁴⁵ Unless stated otherwise, the information in this section is taken from Millner, R., Üstün, S., Moder, C. & Meyer, M. (2019). Social Impact Bond Pilotprojekt "Perspektive:Arbeit". Wien: WU Wien.

Various organisations, including multiple investors and service providers as well as an intermediary, worked together to deliver the project. These are summarised in the table below.

Table 5. Summary of Perspektive:Arbeit SIB

Personally targeted social service	Labour market-related
Target population	Unemployed youth that are difficult to reach, in the region of Augsburg, Bavaria
SOC scheme type⁴⁶	Intermediated SIB
Payment model⁴⁷	Binary
Cost	EUR 250,000 ⁴⁸
Commissioner(s)	Bavarian State Ministry of Labour, Social and Family Affairs and Integration
Social service provider(s)	Main local social service providers: <ul style="list-style-type: none"> • Regionales Bildungszentrum Eckert gGmbH (Eckert Schulen) (NGO) • Kinder-, Jugend und Familienhilfe Hochzoll (NGO) <p>“Two external organisations specialising in youth and educational support”⁴⁹:</p> <ul style="list-style-type: none"> • apeiros e. V. (NGO) • Joblinge gAG München (NGO)
Investor(s)	<ul style="list-style-type: none"> • BHF-BANK-Stiftung (philanthropic organisation) • Eberhard von Kuenheim Stiftung (philanthropic organisation) • BMW Stiftung Herbert Quandt (philanthropic organisation) • BonVenture gGmbH (private investor)
Intermediary	Juvat gGmbH, “a non-profit subsidiary of the Benckiser Foundation Future” ⁵⁰ (non-profit organisation)
Evaluator(s)	<ul style="list-style-type: none"> • Evaluation of success: Law firm Spiegel RA WP StB Partnerschaft mbB (contract take-over from law firm Dr. Mohren & Partner, which was dissolved in 2015) • Accompanying evaluation: University of Hamburg, Prof. Dr. Barbara Scheck

Juvat gGmbH—the intermediary involved in the delivery of the Eleven Augsburg SIB (see Section 1.1) – also helped implement the Austria SIB. The intermediary recruited the

⁴⁶ Three types of SIB can be distinguished. In a **direct SIB**, a delivery contract is signed between the outcomes-payer and service provider. The service provider carries the risk if results are not achieved, as the investor lends funds and is repaid first with a loan secured against the enterprise’s assets. **Intermediated SIBs** also involve an intermediary to liaise between investors, service providers and the commissioner. The service delivery contract is signed between the outcomes payer and an investor-owned special purpose vehicle (SPV), which contracts the service provider, supports the performance management process and refines the financial model. The risk is shared within the SPV among the investors, and repayment depends on the SIB’s structure. A **managed SIB** is signed between the outcomes-payer and the prime contractor (usually an intermediary), who usually manages the entire process. The investors carry the risk but at the same time, they obtain profits if the results are achieved or exceeded. Please find more information on the conceptual framework developed in the Inception report.

⁴⁷ In **binary** models, the provider has to achieve an absolute target, and no payment is granted for achieving lesser results. This can be contrasted with **frequency schemes**, in which rewards are staggered according to the agreed frequency of results, with payments increasing as results increase. Another model is a hybrid grant, in which the cost of delivering a service is funded, but additional payments are awarded as bonuses if ‘additional’ impacts are demonstrated at the end of a programme. Please find more information on the conceptual framework developed in the Inception report.

⁴⁸ Juvat (2016), p. 2.

⁴⁹ Jebabli (2016), 27.

⁵⁰ Scheuerle & Nieweler (2017), 3.

investors and social service providers, and allocated the budget. **Payments** to service providers were transferred in quarterly instalments, based on quarterly progress reports and quarterly budget forecasts. The latter took into account necessary financial resources for the measures and activities implemented by the social service providers.

By the end of the project in September 2018, 311 women had benefitted from the intervention, with 182 women finding primary employment and an additional 16 starting a second job⁵¹. Nevertheless, its main target (75 women, working at least 20 hours a week for 12 months and earning at least EUR 19,500 per year) was not met because the employment acquired either did not last 12 months, or was remunerated at a level below the required threshold of EUR 19,500. As a result, the Ministry of Labour, Social Affairs and Consumer Protection did not issue any payments for the project's costs to the investors⁵².

Despite its failure to meet the targets, the intervention was subsequently continued using traditional financing (see the following section). This was because the SIB was perceived as a success by all stakeholders: the mere fact that they managed to make it happen in Austria in the first place was regarded as an accomplishment. The intermediary also saw success in being able to replicate the SIB model in Austria, after it had piloted a technically similar SIB in Germany. Even the investors – who lost their investment – saw value in supporting and showcasing impact-oriented SIBs for other stakeholders who had not previously worked with impact investing. Furthermore, the investors saw success in the continuation of the measure after the SIB. For the social service provider, the scheme was regarded as a success because otherwise there would have been no such service delivered to the target group. Impact was created at various levels and for a number of people, despite failing to achieve the number of successful cases required by the contract. Finally, the stakeholders claimed that the SIB provided important lessons learned: for example, that participants' qualifications are key to meeting the required targets. Hence, if more time had been available for the programme, participants could have acquired more qualifications and the targets could have been achieved.

1.2.2. Equivalent TF scheme

The initiative described above was extended after the completion of the SIB in 2018, and serves as the equivalent TF scheme for the purposes of this study. According to the Ministry of Labour, Social Affairs and Consumer Protection⁵³, the intervention has been transferred into the regular service portfolio of the Austrian Labour Market Agency (Arbeitsmarktservice, AMS) in Upper Austria. A follow-up phase of three months (September 2018 to November 2018) was funded by the intermediary to finalise all project-related work. Since then, the social intervention has been funded by the Ministry of Labour for a further 14 months (referred to as “phase II” in this document, lasting from December 2018 to January 2020) in the form of public subsidies. From February 2020 onwards, the social measure has been commissioned by the Austrian Public Employment Service (AMS, “Arbeitsmarktservice”) for another 14 months (referred to as “phase III” in this document).

The contracts for phase II and phase III were concluded with the Centre for Protection Against Violence in Upper Austria only, even though Women's Shelter Linz remained a cooperation partner. The number of subcontractors (who had formerly delivered activities to enhance qualifications and coaching services within the SIB scheme) has been reduced from two to one. Otherwise, the core characteristics of the social measure remain the same.

⁵¹ Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz - Eingang Sozialministerium (2019). Social Impact Bond. Retrieved from: <https://www.sozialministerium.at/Themen/Soziales/Soziale-Themen/Soziale-Innovation/Social-Impact-Bond.html>

⁵² Ibid. If the SIB had reached its targets, this amount would have been paid to the intermediary, which would have redistributed the relevant funding amount to each investor. Or, to put it differently: the commissioner established one (central) contract with the intermediary, who in turn contracted the investors and the social service providers within the total SIB arrangement.

⁵³ Phone call, 14 April 2020.

New clients have been enrolled into the programme, and existing clients were contacted again. The contract for phase II was based on a public subsidies agreement roughly one-third the size and scope of the three-year SIB. Hence, the duration of phase II was one year, and it had a budget equivalent to one-third of that for the SIB. The service provider indicated the number of envisaged performance criteria in its proposal for the commissioner. The contractual agreement did not explicitly specify performance targets due to the nature of the funding being public subsidies, which require a specified effort from the recipient, but do not involve specified targets. Otherwise, such an agreement would fall under procurement law.

Within phase III, similar services were delivered to those in the SIB. A number of changes were also implemented. These included: the removal of the minimum income clause; additional indicators such as job retention, and a number of measures to enhance participants' qualifications. Phase III is based on a contractual agreement between the AMS and the service provider – an arrangement that allows for performance targets.

Table 6. Comparability of Phases II and III of Perspektive:Arbeit with the SIB

	Same as SOC	Comments
Personally targeted social service	Yes	Crisis and emergency Labour market-related
Target population	Yes	Survivors of domestic violence
Number of participants	Unknown	Number of TF participants is unknown
Location	Yes	Upper Austria
Cost	Yes	Roughly the same annual cost as the SIB, amounting to roughly EUR 267,000
Commissioner(s)	Yes for Phase II No for Phase III	Austrian Federal Ministry of Labour, Social Affairs and Consumer Protection from December 2018 to January 2020 (Phase II); Austrian Public Employment Service (Arbeitsmarktservice – AMS) from February 2020 onwards (Phase III)
Social service provider(s)	Yes/No	Centre for Protection Against Violence in Upper Austria (Gewaltschutzzentrum Oberösterreich) remained a provider, but the Women's Shelter Linz was not formally included as a service provider

Unfortunately, no external evaluation took place during phase II. Although the commissioners have access to information on how phase III of the intervention is performing in relation to its performance criteria, these indicators were not disclosed during the research process. Therefore, the effectiveness and cost comparisons for this scheme rely mostly on information available through desk research and interviews.

1.2.3. Outcomes measurement

Data on the **outcomes upon which payment was conditioned** were extracted from employment contracts, social security system contributions and work records. All data were

collected and reported by the social service providers (Centre for Protection Against Violence in Upper Austria [Gewaltschutzzentrum Oberösterreich] – NPO; Women’s Shelter Linz [Frauenhaus Linz] – NPO). Data were verified by the auditing firm Ernst & Young (Wirtschaftsprüfungsgesellschaft m.b.H. Ernst & Young).

Qualitative evaluations were also performed by the **process evaluator** (Competence Center for Nonprofit Organizations and Social Entrepreneurship at the Vienna University of Economics and Business [NPO & SE Kompetenzzentrum, WU Wien]) and the **accompanying evaluator** (Institute of Conflict Research (Institut für Konfliktforschung), which performed the impact assessment⁵⁴).

Measurement of outcomes relating to payment mechanisms

To receive payment, social service providers were required to self-report how many participants had obtained employment that met the criteria specified in the contract (see the following section). These claims were then verified by the auditing firm Ernst & Young, who audited employment contracts, social security system contributions and work records.

The key outcome concerned the employment of women who had been subject to domestic violence. For payment to be triggered, at least 75 women from the target group had to be placed in jobs that met the following criteria:

- subject to social insurance contributions;
- paid a living wage (i.e. at least EUR 19,500 gross salary);
- lasted at least 20 hours per week;
- continued for at least 12 months during the term of the project.

The threshold of EUR 19,500 gross salary was demanded by the commissioner, and was based on:

- estimates of potential savings within the public sector (e.g. expenses saved within the needs-based minimum benefit system), and additional earnings for the general public (e.g. taxes or contributions to the social security system) in comparison to
- expenses within the needs-based minimum benefit system (hitherto ‘social assistance’) that would otherwise have been provided to the women.

Other evaluations

In addition to the process used to verify that the agreed outcomes had been achieved, the commissioner funded two additional evaluations (the cost of which was not included): an impact evaluation and a process evaluation.

Impact evaluation - qualitative analysis

A qualitative analysis was conducted as part of the impact evaluation of the scheme. This involved 17 qualitative interviews with women from the target group (16 individuals; one woman was interviewed twice). Interviews focused on three areas:

- investigating the status quo of the women in terms of their employment track record, job search experiences and the like;

⁵⁴ Haller and Amesberger (2018). Qualitative Evaluierung des Social Impact Bond PERSPEKTIVE: ARBEIT – Ökonomisches und soziales Empowerment von gewaltbetroffenen Frauen, Institut für Konfliktforschung, Wien.

- analysing individual conditions such as the level of education, skills, and social networks;
- evaluating the support received within the SIB-measures.

Potential outcomes were touched upon in an anecdotal and exploratory manner and not systematically. For instance, the interviews covered topics such as increased self-esteem, absence of violence, self-reported improvements in the mental state of children or improved school achievements, better health conditions, and increased mobility in terms of passing a driving test or purchasing a car.

Process evaluation - qualitative analysis

Twenty-seven qualitative interviews (carried out in four rounds) were conducted with representatives of all partners involved, to ascertain their current perception of the various phases of the SIB (formation of the SIB, negotiating responsibilities, implementation, termination and reflection). Key themes included the motivation to take part in the SIB, as well as the stakeholder's individual assessment of the instrument, the forms and effects of communication, the collaboration between the partners involved, areas of conflict as well as potential changes during the project's implementation.

Barriers to and enablers of the measurement process

The main strength of the measurement process that emerged from desk research is a **'dry run' of case checks**, requested by the evaluator in the middle of the project. This enabled stakeholders to better assess the data needed to verify outcomes. Social service providers learned in greater detail which documents were required and would be accepted as proof of having achieved the targets. Stakeholders were also generally willing to give interviews regarding the process evaluation, with only a few exceptions among the investors.

With regard to barriers, providers reported sometimes having difficulty in accessing the **required certificates and employment contracts**, because in some cases no written labour contracts were available (in Austria, a labour contract does not have to be in a written form). **Accessing beneficiaries for interviews** for the impact evaluation was complex in some cases. Some beneficiaries required translation services and could not be interviewed for this reason. Ultimately, given that none of the evaluations were of an experimental or quasi-experimental nature, the intervention's effectiveness (and hence the causality of impact) could also not be determined.

Table 7. Summary of the measurement process in Perspektive:Arbeit

Measurement methodology	
Methods relating to payment mechanisms	Data self-reported by service providers and verified by E&Y
Other evaluations	Impact evaluation – qualitative Process evaluation – qualitative
Methods using experimental or quasi-experimental design	No
Control group	No
Causality of impact	No

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

Enabling factors and strengths of the evaluation process	'Dry run' to check what proof would be required to verify outcomes Willingness of stakeholders to engage
Barriers to and weaknesses of the evaluation process	No experimental method, no control group Sometimes difficult to gain access to contracts to prove employment and stakeholders for interviews
Governance of measurement	
Independence of evaluator	Yes
Evaluator	<ul style="list-style-type: none"> • Evaluation of success: auditing firm Ernst & Young (Wirtschaftsprüfungsgesellschaft m.b.H. Ernst & Young) • Process evaluation: Competence Centre for Nonprofit Organizations and Social Entrepreneurship at the Vienna University of Economics and Business, (NPO & SE Kompetenzzentrum, WU Wien) <p>Accompanying evaluation: Institute of Conflict Research (Institut für Konfliktforschung)</p>

1.2.4. Effectiveness

Achievement of outcomes

As mentioned above, the target of the SIB was not achieved, since fewer than 75 women from the target group found jobs that satisfied the agreed criteria. However, the SIB allowed contact to be established with 417 women; 311 women and 430 children took part in the intervention, and 30 women received educational and training activities. In terms of outcomes, 182 women found employment in the primary labour market, and 16 women found jobs in the secondary labour market. Only 52 women represent successful cases, according to the agreed targets.

The qualitative evaluation of the SIB reveals that participants faced difficulties in acquiring or maintaining the specified employment for one of the following reasons:

Many participants were underqualified for local job opportunities.

The three-year duration of the project was insufficient to substantially improve participants' qualifications in a way that would allow them to compete for jobs paying the necessary target income.

The sociodemographic characteristics of the target group and, more specifically, of those participating were unfavourable to job acquisition, including factors such as:

- the number of children in need of care per beneficiary, in the context of a lack of childcare infrastructure in the region;
- geographical distance between their place of residence and the available jobs;
- lack of German language skills.⁵⁵

The lack childcare infrastructure in the region and the lack of public transport represent two structural factors hindered the ability of women to find and/or maintain a job, thus negatively influencing the results of the scheme,. In many cases, women with more favourable

⁵⁵ Haller and Amesberger (2018).

preconditions in terms of language skills, personal social networks or other sociodemographic characteristics did not enter into the programme.

Benefits and drawbacks of SOC compared with TF

Concerning its drawbacks, stakeholders agreed that the SIB was more intense and time-consuming with regard to the **management process** (e.g. planning and budget meetings). Nevertheless, interviewees argued that **additional meetings enhanced the quality of exchanges**, allowing different perspectives to be considered, or enabling the establishment of a joint project with aligned interests. While regular public subsidies still require administrative efforts in relation to commissioning, they do not entail such heavy communication efforts on the part of the commissioner.

With regard to the main benefits of the SIB, stakeholders noted that compared to TF models, the SIB allowed greater **flexibility** over spending. For instance, under the SIB it was possible to fund a driving licence for a participant to reach their workplace by car. In this sense, the SIB allowed providers to be more creative in designing tailor-made measures that are not very expensive, but have a potentially great impact on beneficiaries.

Another important benefit of the SIB was that it **unlocked new funding** for the specific target group, namely women who were victims of domestic violence. In fact, some investors were involved due to their interest in the instrument rather than the policy area or type of intervention. For instance, the ERSTE Foundation was interested in experimenting with SIBs to identify and test innovative social measures, despite providing services to women who are victims of domestic violence being outside the usual scope of the foundation's work.

Finally, an important benefit of the SIB was that it allowed the **testing of a new service that had not previously been provided** in the region. One of the reasons for the development of the SIB was, in fact, that the specific target group would otherwise not have received any special support. In this sense, the SIB addressed a gap in traditionally financed services, and allowed a new social service to be tested.

Scalability and replicability

As previously noted, following the SIB, the intervention was funded traditionally in other Austrian states. Replication in subsequent phases was easier than the establishment of the SIB because the project was known among the partners in the Public Employment Service (AMS). To comply with the rules for publicly funded arrangements, the services delivered had to be more strictly defined, which was seen as a drawback by the service providers.

1.2.5. Efficiency

Intervention costs

The total cost to the investors of the Perspektive:Arbeit SIB was EUR 804,688⁵⁶, or EUR 2,587 per programme participant. Given that the SIB helped 52 women find jobs, the cost per participant who acquired a job is equal to EUR 15,475. Given that the programme failed to achieve its goals, the investors were not repaid, so the cost to the taxpayer (or to the commissioner) is equivalent to zero.

A cost comparison of this scheme with the subsequent TF programme cannot be conducted because performance indicators for the equivalent TF intervention were not shared with the

⁵⁶ Millner et al (2019).

research team. When asked to compare the efficiency of SOC schemes and traditional financing more generally, the stakeholders argued that **efficiency is not the focus of SIBs**, at least during the first pilot projects. In Austria, the SIB was implemented to experiment with a new programme for an underserved group. Stakeholders therefore argued that it should be possible to “waste” money on new approaches in order to find out whether or not they work (e.g. by paying for women’s driving licences). Given that the same programme was later extended under traditional financing, at least part of any cost-savings—if they exist—could be attributed to the SIB experience, as it allowed stakeholders to understand which measures are effective.

Operational costs

The set-up of phases II and III was easier than that of the SIB: partners were already available, the programme’s activities were established, project structures had been created and staff were already trained and experienced. Phase II (the TF instrument) can therefore be considered more efficient and effective, having been able to build on previous work carried out within the SIB. By comparison, the SIB was more experimental, offering opportunities for trial and error that spurred learning effects.

With regard to **evaluation costs**, the service provider noted that delivering the necessary documents to prove that participants had found jobs required a lot of effort. Women were often employed in a sector for unskilled labour, in which employment documents were neither available nor provided.⁵⁷ This culminated in a great deal of work for the service provider at the end of the SIB. Due to the specifics of the target group, the provider refrained from getting in direct contact with employers to acquire these documents. **Incorporating lessons learned from the SIB, outcome tracking was handled differently in phase II, under which those documents were provided by the AMS. This made documentation easier and more efficient.**

Although overhead costs typically appear to have been greater in the initial SIB compared with the subsequent phases of the programme, the commissioner argued **the workload for financial settlement is higher in the TF intervention than in the SIB**. On the one hand, the commissioner plays a lesser role in tracking financial accounting within the SIB. This is partly because it is outcomes – rather than outputs – that matter the most in the SIB, and because of the involvement of the intermediary. On the other hand, once the measure is directly funded via public subsidies/grants, a stricter regime applies in terms of a) which expenses are eligible for funding, and b) what checks and verifications are required as to how and for what purpose funds are spent.

1.2.6. Design of the scheme

Design features

A number of design features were perceived to have made the project more effective and/or efficient. Stakeholders generally agreed on the following:

- It was crucial to gather **as much information as possible about the target group** and share it with all parties involved prior to the intervention.
- **A ‘dry run’ in the middle of the project** was helpful in identifying what information would be needed by the evaluator to confirm employment outcomes.

⁵⁷ In Austria, written work contracts are not mandated by law.

- **The SIB Forum**, organised twice a year, also helped to bring together all in the stakeholders SIB for exchange and oversight.
- **The social service provider's expertise** in the subject area was a key factor in attracting investors to participate.

Furthermore, the stakeholders noted a number of lessons learned that were later incorporated in the later phases of the project, or were utilised when launching other SIBs in Austria:

- **The three-year contract term was perceived as being too short.** Given that many women were underqualified for the employment that was available where they lived, stakeholders argued that more time was needed for them to acquire the necessary qualifications.
- Wherever possible, **commissioners should draw on existing documentation from public authorities or administrative data to demonstrate outcomes.** This enhances the efficiency of the intervention by allowing the service provider to focus on participants rather than document trails.
- **Contracts and respective templates should be part of the tender materials.**
- **Evaluation results should be provided to each and every partner of the SIB.**
- **All costs, including overheads, should be incorporated into the programme costs.**

In addition to these design features and lessons learned, the stakeholders involved in the Austria SIB made conscious decisions in order to minimise the possibility for the private sector to profit. As in Germany, the Ministry of Labour, Social Affairs and Consumer Protection limited **the pool of investors to charitable organisations** only, and an **interest rate of 1% was set to account for inflation** as opposed to providing monetary incentives for investors to participate. Furthermore, to address the '**wrong pockets**' challenge, the Austrian Federal Ministry of Labour, Social Affairs and Consumer Protection cooperated with the state of Upper Austria and the Austrian Federal Ministry of Education and Women's Affairs.

Another element of the scheme's design revolved around its choice to **treat the investment as a loan to the intermediary**. This consideration was closely related to the national legislation and related constraints mentioned above.

Impact on third-sector organisations and social enterprises

While the SIB unlocked new sources of funding for the service provider, and provided the flexibility to customise social services for the target group, the service provider interviewed nevertheless claimed that traditional funding is preferable. This is because during the SIB, staff members felt insecure about their jobs once the SIB project finished, or in the event that it would be terminated earlier than expected due to disappointing results. Some of these concerns could nevertheless be mitigated if the SIB contract were to last longer than three years.

1.3. DUO for a JOB (Belgium)

1.3.1. Background

DUO for a JOB (or DUO in short) is a non-profit organisation created in November 2012 that provides an intergenerational mentorship programme. The programme pairs **young immigrants** with Belgian mentors aged 50 and older for a period of six months, with the

aim of helping immigrant youth to find jobs. Young immigrants receive insights into life and work in Belgium from people with deep knowledge about the country, providing both the mentor and the mentee with valuable intergenerational and intercultural interactions.

Between 2014 and 2016, the mentorship programme was partly funded through a SIB, which is the main focus of this analysis. The SIB emerged in response to a wide gap in labour market participation between those born in Belgium and non-European immigrants (in 2016, the unemployment rates among the two groups were 6.1% and 21.2%, respectively) and a high youth unemployment rate in general (in Brussels, 17.7% of 15-24-year-olds in 2017 were not in education, employment or training [NEET])⁵⁸. Acknowledging a lack of change in this situation despite multiple initiatives in the past, as well as the absence of a specific employment policy targeting non-European Union (EU) youth in Brussels and budgetary pressures, Gregor Chapelle, the head of Brussels agency for employment Actiris, was seeking **alternative ways to fund targeted employment policies**. Chapelle contacted Thomas Dermine, a Belgian consultant who had worked on SIBs in Europe for his Master's thesis, to conduct a feasibility study for a SIB in Brussels⁵⁹.

After the feasibility study showed positive results, Actiris defined the following target population: i) youth under 30 years old; ii) immigrants from outside the EU, the US, Canada and Switzerland; and iii) registered with Actiris (i.e. residents of the Brussels-Capital Region). Actiris then set up a selection procedure for all service providers who were not eligible for traditional funding. Selection criteria were defined, including the innovativeness of the proposed methodology⁶⁰.

At the time, DUO for a JOB was facing financial issues. The organisation had initially been funded by its two cofounders' personal savings. By mid-2013, their available capital had dwindled. The lack of a proven track record for the programme was a key hurdle to gaining public subsidies or other types of funding (e.g. large donations by corporations). The organisation therefore welcomed the invitation to apply for SIB funding,⁶¹ which it subsequently acquired.

Other **parties involved** in the project included HNWI (investor), KOIS (intermediary) and two evaluators: Observatoire Bruxellois de l'Emploi (public agency, part of Actiris) and the Centre for Social Economy at the University of Liège. All parties are summarised in the table below.

Table 8. Summary of the DUO for a JOB SIB (2014-2016)

Personally targeted social service	Labour market-related
Target population	Unemployed non-EU youth in Brussels
SOC scheme type	Intermediated SIB
Payment model	Frequency
Cost	EUR 347,000
Commissioner(s)	Actiris (Brussels employment agency)
Social service provider(s)	DUO for a JOB asbl (NGO)
Investor(s)	Undisclosed foundation

⁵⁸ KOIS (n.d.). DUO for a JOB: the first Social Impact Bond in continental Europe. Retrieved from: <https://www.koisinvest.com/post/duo-for-a-job-the-first-social-impact-bond>

⁵⁹ Smans, V. (2015). Social impact bonds (SIB) et mesure d'impact social. Etude de cas du SIB belge en collaboration avec Actiris et DUO for a JOB (Unpublished Master's thesis), ICHEC, Brussels, Belgium. Retrieved from: https://issuu.com/fqfffg/docs/mtafinance2016-smans_tfe, 47. Also, personal communication with the service provider, (2020).

⁶⁰ Smans (2015), 48-49.

⁶¹ Dermine, T., Le Grelle, M., & Simonart, F. (2016). Social Impact Bonding. Stanford Social Innovation Review, 2016(Fall), 62. Also, personal communication with the service provider, (2020).

Intermediary	KOIS (for-profit organisation)
Evaluator(s)	Observatoire Bruxellois de l'Emploi (public agency, part of Actiris) Centre for Social Economy at the University of Liège (university research centres)
Other parties involved	A mediator was designated in case of conflict between two or more parties to the contract.

The DUO SIB included the following **activities**:

- Collective information sessions to present the project, clarify roles and answer possible questions;
- Individual interviews with candidate beneficiaries (mentors as well as mentees);
- Initial training session of four days for mentors;
- Matching mentors with mentees based on their respective needs and skills;
- Organising the pair's first meeting;
- Signing an agreement between the mentor and the mentee;
- Mentoring for six months;
- Evaluation by members of the DUO team as to whether the objectives had been achieved.

The main **objective** of the SIB was to achieve an employment rate among the DUO for a JOB group that was 10% higher than that of the control group. The control group consisted of migrants with similar demographic characteristics, also registered with Actiris, but who did not enrol in the DUO for a JOB programme. The employment rate was measured one year after the end of the intervention. The individual metric was set equal to 1 if the participant had been employed for at least 90 working days within the preceding 12 months.⁶²

In every year of the programme's duration, the **objective was achieved**. In total, 322 mentee-mentor duos were established, which resulted in 133 job placements for DUO mentees.

With regard to **funding arrangements**, the contract specified interim payments of EUR 58,500 to service providers every six months. This was so that the NGO could cope with the inflow of cash, and, on the other hand, so that DUO would not have to rely on credit to fund the programme's activities⁶³. Furthermore, for every year the objective was fulfilled, investors earned 4% interest on their investment.

According to the intermediary organisation KOIS, the programme generated EUR 760,000 in total savings, or 2.1 times the cost of the programme⁶⁴.

1.3.2. Equivalent TF scheme

Following the SIB, DUO was allocated a first 'bridging' subsidy for two years (2017-2019) by Actiris on the basis of preliminary results of the SIB. This was done because DUO had delivered good results, and the organisation had to wait for two years to apply to the quinquennial call for tenders (for traditional procurements). In 2019, DUO acquired funding from Actiris for an additional five years (until 2024) after a regular tendering procedure⁶⁵. In

⁶² Dermine et al (2016), 62.

⁶³ Ibid, 49

⁶⁴ KOIS (n.d.).

⁶⁵ Personal communication with the service provider (2020).

this study, we therefore compare the different time periods during which the DUO mentorship programme was implemented (2014-2016 vs 2017 onward), reflecting the different funding streams it received.

As demonstrated in the table below, after the bridging subsidy, the project was scaled up. During the SIB, DUO for a JOB was granted relatively little funding because it was the commissioner’s first experience with a SIB. Please note that during both time periods (2014-2016 and 2017 onwards), DUO for a JOB also received additional funding from donations, grants and other sources (for its full operating budgets, please see Table 12). Using these funds, DUO has expanded to Antwerp and Liège.

Table 9. Comparability of the SIB with subsequent Actiris-funded DUO for a JOB programme

	Same as SOC	Comments
Personally targeted social service	Yes	Labour market-related
Target population	Yes	Unemployed non-EU youth
Number of participants	Unknown	It is not possible to break down the number of participants by funding stream
Location	Yes/No	Brussels and other cities in Belgium
Cost	Yes for the bridging subsidy; No for subsequent funding	a) Bridging subsidy for 2017-2018 cost EUR 234,000 (EUR 117,000 per year – same as during SIB) b) Funding for 2019-2024: EUR 3.75 million (EUR 750,000 per year)
Commissioner(s)	Yes	Actiris (Brussels employment agency)
Social service provider(s)	Yes	DUO for a JOB asbl (NGO)

1.3.3. Outcomes measurement

A simple comparative design was used to assess the achievement of the outcomes upon which payment was conditioned. This assessment was carried out by the Observatoire Bruxellois de l’Emploi (a public agency), and later validated by an independent evaluator (Centre for Social Economy at the University of Liège)⁶⁶.

Measurement of outcomes relating to payment mechanisms

The achievement of the outcomes upon which payment was conditioned was assessed through a simple comparative design: namely a pre-post comparison of key outcomes in the treatment and control groups. The evaluation relied on administrative data (specifically, social security data). For each individual in either the DUO beneficiaries group or the control group, the number of days worked during a pre-defined period was collected from social security data based on employers’ declarations of employment. For each cohort of the treatment group, the control group included individuals who fitted the criteria for DUO’s target group (i.e. unemployed non-EU immigrants under 30 years old, living in Brussels), but were not taking part in DUO’s mentoring programme, and were registered with Actiris

⁶⁶ Unless stated otherwise, the information in this section is taken from interviews with the evaluators.

(the Brussel Agency for Employment, the programme's commissioner) during the same year as the cohort.

It is important to note that all individuals who were registered with Actiris that fit the criteria of DUO's target group were informed by Actiris of the DUO programme and collective information sessions: this means that the control group consisted of people who either decided not to follow the programme or (in very rare cases) were not selected by DUO during the second stage of the programme. This could happen when a candidate showed a mismatch with DUO's philosophy (e.g. an unwillingness to work with a mentor of a specific gender). After a recommendation by the independent evaluator (Centre for Social Economy, University of Liège), the control group was adjusted to match the demographic characteristics of DUO's beneficiaries, following a stratified sampling method, with individuals randomly selected within each demographic category.

Such a procedure enabled the creation of a similar control group. Despite this, the fact that the control group is mainly made up of individuals who had decided not to take part to the programme supports the notion that treatment and control groups were inherently different. For example, those who decided to participate in DUO might have been more motivated to look for work. The possibility of inherent differences between the treatment and control groups undermines the findings of the evaluation because it is unclear whether the treatment group achieved better outcomes because the mentorship programme was effective, or because participants in the treatment group were different (e.g. more motivated to look for work).

The key payment-related indicator for success is the employment rate (ER), which is measured one year after the end of the intervention. The indicator is measured using a binary variable that is equal to 1 if the individual has been employed for at least 90 working days during the last 12 months.

The programme also set a specific target for payment. The fixed target is a 10% higher rate of employment among the treated group than in the control group. Please note that the difference is relative and is calculated using the following formula:

$$\text{Increase in job placement rate due to mentoring} = 1 - \text{DUO's ER/control group's ER}$$

Barriers to and enablers of the measurement process

A key enabler and strength of the measurement process that emerged from the desk research is the **involvement of service providers in the planning phases** of the evaluation. According to the service providers, certain decisions associated with the evaluation can impact the services delivered to participants, how those services are delivered, and even what kinds of people are involved in the programme. Some of these decisions can be contradictory to service providers' mission statements (for example, limiting services only to a particular group of people), so their involvement in the design of the evaluation is necessary to ensure the evaluation can subsequently be effectively evaluated. The service provider also noted that the evaluation was not time-consuming (providers only had to provide the names of the participants, and when they enrolled or left the programme).⁶⁷ Finally, the **control group** provided some rigour to the evaluation.

Barriers emerge regarding the **single-dimension assessment**. The evaluation only considered employment as an outcome, excluding other relevant dimensions of the intervention (e.g. social cohesion, integration, etc.), which social providers found important in their work as well. Issues also emerged between the commissioners and providers in relation to the misalignment in the choice of evaluated outcome. The time span of the measurement (a year after the last potential exit day of the cohort) was perceived as a

⁶⁷ Interview with the service provider.

burden by several stakeholders, but was highly valued by the service provider DUO in order to detect non-immediate effects of the programme. Finally, as discussed previously, participants enrolled in the programme voluntarily, so they might have had different unobserved characteristics from those in the control group, leading to biases that undermine robust comparisons in relation to outcomes measurement. While the Centre for Social Economy at the University of Liège attempted to correct for the self-selection bias by matching participants in the control group with those in the treatment group based on similar demographic characteristics, such a matching procedure nevertheless failed to account for unobserved characteristics such as motivation to work, childcare obligations, and other latent factors that might affect the outcomes measured.

Table 10. Summary of the measurement process in DUO for a JOB

Measurement methodology	
Methods relating to payment mechanisms	Pre-post comparison in treated and control groups
Other evaluations	N/A
Methods using experimental or quasi-experimental design	No
Control group	Yes
Causality of impact	No, due to the possibility of the self-selection bias with regard to programme participants and the control group
Enabling factors and strengths of the evaluation process	<ul style="list-style-type: none"> - Involvement of providers in the planning phases to prevent methods being designed that may lead to mission drift. - Control group Limited burden on the time of service provider
Barriers to and weaknesses of the evaluation process	<ul style="list-style-type: none"> - Misalignment between commissioners and providers in the choice of the dimension assessed - Time span of measurement perceived as a burden, but necessary to detect nonimmediate effects Possible self-selection bias
Governance of measurement	
Independence of evaluator	Yes
Evaluator	Observatoire Bruxellois de l'Emploi (Public Agency) Centre for Social Economy, University of Liège.

1.3.4. Effectiveness

Achievement of outcomes

The programme achieved the agreed target. Relevant statistics for each year's cohort are summarised in the table below. It is important to note that DUO participants joined the programme **voluntarily**. Therefore, it is possible that those who enrolled were more willing to find employment, resulting in better outcomes for the DUO group. In other words, better outcomes might have been achieved for DUO participants due to characteristics inherent

to those who decided to enrol, rather than the effectiveness of the mentorship programme itself.

Table 11. Annual employment outcomes of the DUO for a JOB SIB⁶⁸

	2014	2015	2016
DUO cohorts	42.9%	43.6%	38.7%
Control group	33.7%	30.5%	33.4%
Number of duos	70	110	142
Increase in job placement rate due to mentoring (= 1-DUO's rate/control group's rate)	27%	42%	16%

The TF intervention also had positive outcomes: 55% of mentees found a job within the 12 months following the mentoring programme during the period 2017-2018 (no results are yet available for the period 2019-2024).

These data show that the 2017-2018 TF intervention had better results than the SOC scheme. However, no evaluation is available to compare the results of the SOC with those of the TF intervention. With regard the comparative effectiveness of the SOC and TF models, the stakeholders interviewed mentioned that since the TF model came at a later stage in DUO's development, they had already had a chance to identify what worked and what could be made more effective (and efficient). However, they did not relate effectiveness to a specific funding model. It is possible that the positive results of the TF intervention were facilitated by the previous work developed under the SOC scheme, but this has not been proven by a rigorous impact evaluation that takes into account all appropriate intervening factors.

Benefits and drawbacks of SOC compared with TF

All stakeholders reported a positive overall perception of the scheme and recognised its added value compared with previous interventions. In fact, there have been previous attempts to establish mentoring programmes, including by Actiris itself, but these were not successful. These attempts differed from the mentoring programme of DUO in two key respects:

- they did not take an intercultural and intergenerational approach;
- they were organised as a side activity, whereas in DUO mentoring was the core activity.

Due to these previous experiences, Actiris had been reluctant to fund mentoring activities. In this sense, the investor interviewed mentioned that the SIB **allowed a service innovation to be funded** that was perceived as risky by the commissioner, and to demonstrate value for the public authority by measuring outcomes in a rigorous way. The SIB allowed funding to be unlocked in order to experiment with a new approach to mentoring in a context in which the commissioner would not have directly funded a similar intervention. The SIB also allowed **the effectiveness of the intervention model to be proven**, which

⁶⁸ DUO for a JOB (2019). Rapport d'expérience. Le mentoring, la clé de l'insertion socio-professionnelle des jeunes issus de l'immigration ? Brussels. Retrieved from: <https://www.calameo.com/read/005093166cbb79844724e?authid=FYcbV2sqy2wo>, 72.

may also benefit other organisations that will provide mentoring services in future under the TF model.

The SIB also brought some benefits to the provider, since it ensured **financial stability in the medium term**, and acted as leverage to raise additional funds for a service that would not otherwise have been funded. Moreover, the provider also noted that the SIB increased the visibility of the organisation in the media and in conferences⁶⁹.

According to DUO, the **measurement process** was another benefit of the SIB, because it enabled the organisation to enhance the effectiveness of the intervention, and to prove it to the investors and the commissioner. The provider also stated that the effectiveness of other services – including traditionally funded ones – would also increase if they were provided with the same impact measurement evaluation, in particular the provision of a control group.

The provider perceived the presence of strict **eligibility criteria** that excluded potential users of their service as the main drawback of the SIB, in comparison to TF models. The SIB put a very strict threshold (max 30-year-old beneficiaries, only first-generation immigrants) due to data availability and the need to compute precise and controllable indicators for the reimbursement of the investors. Another drawback was the **more complex management** process of the SIB, compared with TF (see Section 1.3.5). Finally, given that DUO was the first SIB in Belgium, it generated **a lot of media attention**. While such attention allowed the service provider to acquire more funding for the intervention, it also placed extra pressure on all parties to ensure the programme succeeded. For this reason, the target thresholds may potentially have been set quite low, because all stakeholders had a lot to lose in terms of reputation. If the SIB failed, it would have reflected poorly on the commissioner's ability to manage such novel funding mechanisms; it could potentially have closed the door to future SIBs from the perspective of the intermediary or investors; and failure could have caused future funding for the service provider to dry up.

Scalability and replicability

As mentioned above, the same intervention provided under the SIB was scaled using a TF model. The replication of the SIB in other countries is currently under discussion. According to the interviews, opportunities may exist to replicate the SIB in order to implement similar interventions in Paris and Rotterdam. The SIB model is seen as a way to test the same intervention and demonstrate its impact in a different institutional context. The contract for the DUO SIB was also used as the basis for the future SIBs launched by Actiris, especially with regard to governance structure and the evaluation process.

1.3.5. Efficiency

Intervention costs

It is difficult to carry out a cost comparison between the DUO for a JOB SIB and the organisation's subsequent activities (funded via traditional financing) because the SIB funding was injected into the total pool of resources available to the organisation, and the results were tracked for the organisation as a whole rather than separately for the SIB. For example, in 2015, the total operating budget for DUO for a JOB was EUR 620,681⁷⁰, out of which approximately EUR 116,000⁷¹ came from the SIB investors. While we know that

⁶⁹ DUO for a JOB (2015). Rapport Annuel 2014. Retrieved from <https://en.calameo.com/read/00509316677e7560696d6?authid=W3c7leRHP2iV>

⁷⁰ DUO for a JOB (2017). Rapport Annuel 2016. Retrieved from: <https://www.duoforajob.be/wp-content/uploads/2017/06/Rapport-Annuel-DUO-for-a-JOB-2016.pdf>, 58.

⁷¹ This figure is approximate, given that a EUR 347,000 investment was spread over the course of three years.

between 2013 and 2016, 176 mentees found jobs,⁷² we cannot separate these results according to the funding stream used to achieve them.

To tackle these issues, we compared DUO's cost per mentee who found a job in two different time periods: 2013-2016 and 2017-2018. If the SIB had a notable effect on cost-effectiveness, we would expect to see a lower cost per participant who found a job during the 2013-2016 period, relative to the 2017-2018 period. In fact, **the costs per outcome achieved turned out to be almost identical for both time periods** (see the table below). On the one hand, this could be because outcomes-based funding comprised a relatively small share (19%) of all funding between 2013 and 2016. On the other hand, it could mean that the organisation's working approach is the same regardless of the funding model. An interview with the service provider supports the latter explanation: the provider noted that they do not see any difference in terms of programme effectiveness depending on whether they are funded via SIB or traditionally, because their approach to working with the beneficiaries is identical.

Table 12. Cost comparison between DUO for a JOB outcomes in 2013-2016 and 2017-2018 (EUR)

	SIB and TF funding 2013-2016	TF funding 2017-2018
Duos	332 ⁷³	1,248 ⁷⁴
Jobs	176 ⁷⁵	458 ⁷⁶
Revenue	1,667,524 ⁷⁷	4,305,777 ⁷⁸
Cost per duo	5,023	3,450
Cost per participant who found a job	9,475	9,401

Source: compiled by PPMI, based on DUO for a JOB Annual Reports for the years 2014-2018.

Note: the figures have not been adjusted for inflation. The figures include total DUO for a JOB operating budgets, including operational costs.

According to the commissioner, the primary goal of the SIB was not cost-savings, but rather to test an innovative intervention. For public authorities, it is difficult to take a risk and to engage in supporting new activities that have not yet been demonstrated to work. In such cases, the commissioner would recommend a SIB. If the intervention proves useful and has a positive impact (in comparison to its costs), then traditional financing will be made available for the intervention.

⁷² DUO for a JOB (2017), 44.

⁷³ DUO for a JOB (2017), 44.

⁷⁴ DUO for a JOB (2018). Rapport Annuel 2017. Brussels. Retrieved from: <https://en.calameo.com/read/00509316698795b998034?authid=VjT1JtxlfY8u&page=51>, 48, states that 1,580 duos have participated in the programme from the start until the end of 2018. DUO for a JOB (2017), p. 44 states that 332 duos had participated in the programme by the end of 2016. Therefore, the resulting number of duos for 2017-2018 is equivalent to 1,580-332=1,248.

⁷⁵ Compiled by PPMI, based on DUO for a JOB (2017), 44.

⁷⁶ DUO for a JOB (2019a). Rapport Annuel 2018. Brussels. Retrieved from: <https://www.bonnescauses.be/media/43178/duoforajob-annual-report-2018-fr.pdf>, 48, states that 634 jobs have been achieved from the start of the programme until the end of 2018. Since the number of jobs achieved by the end of 2016 is equivalent to 176 (see footnote 41), the resulting number of jobs for 2017-2018 is equivalent to 634-176=458.

⁷⁷ Compiled by PPMI, based on DUO for a JOB (2015). Rapport Annuel 2014. Retrieved from: <https://en.calameo.com/read/00509316677e7560696d6?authid=W3c7leRHP2iV&page=37>, 34.

⁷⁸ DUO for a JOB (2019a), 71.

Operational costs

Overall, it is extremely difficult to calculate overhead costs for this SIB and compare them to those for the TF intervention, because such costs were not carefully tracked during the interventions. Nevertheless, all stakeholders agreed that the **process of overseeing the scheme** was less intensive under the TF model and that **the time required to set up the SIB and coordinate between the various parties was probably greater — and possibly much greater — than during the TF interventions**. Under the TF model there were no large steering committees, and most evaluation occurred at the end of each year of the intervention, when an annual report was delivered by DUO. The easier management process of the TF model resulted from the involvement of fewer stakeholders and the less risky nature of projects financed under TF. However, multiple interviewees also emphasised that the expertise shared during joint meetings benefitted either the quality of the intervention itself or the service provider more broadly.

Unlike some of the other SOC schemes analysed in this study, proving outcomes under the DUO SIB did not entail large costs to the service provider. In fact, the service provider estimates that the **evaluation burden under both the SIB and subsequent TF extensions of the project amounted to just a few minutes per mentee**. This is because the data to prove outcomes was accessed from administrative databases and the service provider did not have to provide proof of employment for each individual. Participants were only asked to fill out a survey evaluating the programme. For the commissioner too, the verification of outcomes did not involve a large burden, since the commissioner relied mostly on the conclusions of the evaluator.

1.3.6. Design of the scheme

Design features

Interviewees pointed to a number of elements that made the SIB effective and efficient. For example:

- Multiple interviewees emphasised the importance of the **preparatory phase**, especially with respect to the **evaluation method**, in making sure that the evaluation of results is transparent and examined by an external expert. The service provider noted that its involvement in this phase is crucial to prevent **mission drift**; for example, in case the SIB set-up would require them to refuse services to certain participants. The **commissioner appreciated the binary outcome metric** (job/no job) because it focused stakeholders on the objective, it was easy to compute, and straightforward to interpret. The service provider added that it was crucial to **include SIB recipients in the evaluation**: while beneficiaries' perceptions did not affect the ultimate outcome calculation, the provider collected feedback from participants via surveys to guide its work. Lastly, the **use of administrative data** made the evaluation more efficient. The only downside, perhaps, is that participants were recruited to the programme on a **voluntary basis**, which makes it difficult to establish the true impact of the programme. However, refusing services to some potential participants in order to design a robust control group would have gone against the service provider's mission.
- Multiple stakeholders also lauded the **presence of an intermediary**. Interviewees argued that although intermediary services might seem costly, intermediaries facilitate the alignment of interests and therefore ultimately save time, effort and costs. The service provider added that the presence of the intermediary freed up their time to focus more on the beneficiaries.

- Fluent **communication** among the actors also appears crucial. The SIB was implemented with the help of the **steering committee**, which provided the occasion to pool knowledge and improve the services provided.
- Finally, although no attempts were made to **co-finance** the DUO SIB (this was the first SIB for the commissioner, and involving additional actors would have made the set-up of the scheme more complex), there have been attempts to co-finance future SIBs. This comes in contrast to traditional financing: according to the commissioner, the institutional architecture of Belgium makes it difficult to co-finance traditional interventions, although it does happen in rare cases.

However, some of the stakeholders interviewed questioned the true motivation of this SIB (and others that followed). They claimed that the outcome of the SIB outcome was fairly predictable, meaning that the **risk transferred onto investors was quite limited**. Interviewees argued that the SIB was an opportunity to test out a new funding mechanism, but doubted that the SIB was designed effectively to test whether or not a novel intervention works.

Impact on third-sector organisations and social enterprises

All interviewees underlined the **SIB's benefits to the service provider**. Whereas DUO for a JOB faced financial issues prior to the SIB, the project ensured **financial stability** in the medium term and acted as **leverage to raise additional funds** once the intervention had been evaluated as successful. The SIB also served as a sort of **guarantee for the project's stakeholders** (volunteers, partners, etc.). It increased the **visibility** of the organisation in the media and at conferences. Through the meetings of the steering group, the SIB offered opportunities for recommendations and **advice**⁷⁹. Largely due to its participation in the SIB, DUO for a JOB was able to **scale** its programme: its revenue grew from EUR 179,000⁸⁰ in 2014 to EUR 2.9 million⁸¹ in 2019. However, the service provider interviewed also noted a number of risks associated with SIB participation, namely, the threat of **mission drift**, **pressure** on the team, and the risk of **bad press** in the event of a negative result⁸².

1.4. Buzinezzclub SIB (Netherlands)

1.4.1. Background

Between 2014 and 2016, the first SIB in the Netherlands was launched, with the aim of combating **youth unemployment**. The intervention focused on discovering and developing the talents of individuals aged 17 to 27⁸³ who had so far been unsuccessful in securing education or employment, often due to problematic family and neighbourhood contexts. The goal was to help the participants into a job or appropriate education programme, or to start a business.

The intervention was delivered by Buzinezzclub, a private for-profit company, in Rotterdam. **Outcomes-based funding was chosen because the municipality could not fund Buzinezzclub's activities before they were proven to be cost-effective**. The SIB model enabled the municipality to pay for the programme only if its costs were exceeded by savings in terms of reduced unemployment benefits.

⁷⁹ DUO for a JOB (2015), 35.

⁸⁰ DUO for a JOB (2015), 34.

⁸¹ DUO for a JOB (2020). Rapport annuel 2019. Retrieved from: <https://www.duoforajob.be/fr/brochures-et-rapports/>, 57.

⁸² Personal communication with the service provider (2020).

⁸³ Davelaar, M., Tan, S., & Spies, H. (2015). WP7 Case Study: Innovative Practice „Buzinezzclub Rotterdam“. Rotterdam. Verfügbar unter Retrieved from: http://www.citispyce.eu/sites/default/files/15%20Rotterdam_%20Buzinezzclub.pdf, 13.

Buzinezzclub delivered 10 weeks of training sessions and six months of coaching to programme participants. Where it was relevant at intake or within the first few weeks, the programme also tackled personal problems such as debt. While coaches advised the youth on more manageable issues, for complex problems such as large debts and insecure housing, specialised public services were consulted, or participants were sent back to the Youth Counter for specialised help and care. The programme then officially began with a kick-off event at which new participants were introduced to the programme, to the coaches, and to former participants. During the six months of the programme, participants attended four days a week to work on the following:

- writing a business plan and a personal development plan;
- attending workshops by young, successful entrepreneurs and experts to learn about entrepreneurial attitude, personal branding and presentation, coping with difficulties, project management and time management;
- engaging in individual- and group coaching;
- completing an internship in the sector he/she wish to have a job or business;
- attending networking events in order to make education, job or business a success.

Upon completion of the programme, Buzinezzclub awarded certificates to those participants who achieved the transition into either education or employment. On request, participants could also be linked with a coach ('buddy') for up to two years.⁸⁴

The SIB initially **cost** EUR 680,000, which comprised costs of roughly EUR 4,000 per individual, plus an additional EUR 40,000 for an intermediary fee.⁸⁵ However, the investors – which included a philanthropic investor called the Start Foundation and a private investor ABN AMRO – were keen for the social enterprise to have a financial stake in the programme, so Buzinezzclub also invested EUR 450,000 on the same terms as the other investors.⁸⁶ According to the terms of the contract, investors would only be repaid if participants came off benefits at least 210 days faster than predicted without support (measured as a median for the entire participant group), over the two years the programme lasted.⁸⁷ The contract also provided for a 12% rate of return in the event that the main objective was successfully achieved. In the event of poorer outcomes, investors would be repaid less, to account for the smaller amount of money saved by the municipality. Outcomes were measured and payments issued on an annual basis.

In addition, investors and the service provider were rewarded for a reduction in the 'backsliding' rate, which is the rate of relapse into unemployment. During times of crisis, investors would receive EUR 9,915 for each percentage-point reduction.⁸⁸ For more information on how the backsliding rate was calculated and what targets were set, please see Section 1.4.3.

Importantly, the municipality was contractually obliged to refer participants to Buzinezzclub. Failure to recruit a sufficient number of participants would mean that the investors and service providers would still be repaid. Participants were chosen by the job coaches of the municipality of Rotterdam. Individuals were only rejected from participating in the programme if they failed to attend multiple meetings, or if their availability over six months could not be guaranteed due to, for example, pregnancy.⁸⁹

⁸⁴ Davelaar et al. (2015), 9.

⁸⁵ Van Es, B. (2015). Social impact bonds. Een nieuw instrument om jeugdwerkloosheid te bestrijden. Hogeschool Rotterdam.

⁸⁶ Eurocities (2016). Rotterdam Buzinezzclub Social Impact Bond. Retrieved from: http://nws.eurocities.eu/MediaShell/media/Feb16_BuzinezzclubRotterdam.pdf, 2.

⁸⁷ Eurocities (2016).

⁸⁸ Deloitte, terugvalrapportages

⁸⁹ Personal communication with the commissioner (2020); Van Es (2015).

While the programme lasted two and a half years, results are only available for the first two years. During that time, 91 out of 160 participants stopped receiving unemployment benefits after the training. The median unemployment benefit pay-out duration of the group was significantly lower than the expected pay-out duration. In the first year, the median reduction was 248 days; in the second year it was 324 days (compared with a minimum target of 210 days).⁹⁰ The project realised a reduction in backsliding of 8.4% in Year 1 and 3.8% in Year 2.⁹¹ The municipality estimated that a 1% reduction in relapse results in savings of EUR 9,915.

Overall, the SIB was considered a **success** and investors were paid in full. In the first two years, the municipality paid EUR 560,246 per year to the intermediary Stichting Social Impact Bonds Rotterdam, which then redistributed the money among the investors.⁹² For a summary of the parties involved in Buzinezzclub SIB, please see the table below.

Table 13. Summary of the Buzinezzclub SIB

Personally targeted social service	Labour market-related
Target population	Unemployed youth
SOC scheme type	Intermediated SIB
Payment model	Frequency
Cost	EUR 680,000, extended to EUR 1,130,000 ⁹³
Commissioner(s)	Municipality of Rotterdam
Social service provider(s)	Buzinezzclub (for-profit social enterprise)
Investor(s)	Start Foundation (philanthropy organisation) ABN AMRO (private investor) Buzinezzclub (for-profit enterprise)
Intermediary	Society Impact (VanDoorne lawyers took a supervisory role)
Evaluator(s)	Deloitte (Consulting firm)

1.4.2. Equivalent TF scheme

After the SIB was completed and a proof of concept was established, Buzinezzclub won funding for the project via public procurement. Its activities remained the same, as well as the contractor and the service provider. Therefore, we compare the same scheme when it was delivered through a SIB model and a TF model. It should be noted, however, that the difference in timing is important. First, lessons learned during the SIB were taken on board in the later, traditionally financed scheme. Second, macro-economic conditions were more favourable during the period of the traditionally financed scheme.

Table 14. Comparability of the Buzinezzclub SIB with subsequent traditional funding

	Same as SOC	Comments
Personally targeted social service	Yes	Labour market related
Target population	Yes	Unemployed youth

⁹⁰ Deloitte (2016), 3.

⁹¹ Deloitte, rapportage terugval

⁹² Deloitte (2016), 3.

⁹³ Van Es, (2015)

	Same as SOC	Comments
Number of participants	Unknown	The target of 400 for the number of participants was set in the contract obtained under procurement, which lasts from 2018 to 2021, but the detailed breakdown of participants actually recruited for the programme is not available. ⁹⁴
Location	Yes	Rotterdam
Cost	Unknown	The cost of the TF scheme is not known
Commissioner(s)	Yes	Municipality of Rotterdam
Social service provider(s)	Yes	Buzinezzclub (for-profit social enterprise)

1.4.3. Outcomes measurement

Measurement of outcomes relating to payment mechanisms

Participant results were compared against the benchmark using propensity score matching⁹⁵. The evaluation was carried out by a private firm, Deloitte. On the basis of data collected between 2012 and 2016, the expected chance of employment and the expected duration of benefit payments until employment was calculated for each participant. This calculation was based on the characteristics of age, gender, household situation, ethnicity, level of education, residential area, starting date of benefits, trajectory, and cyclical characteristics. It was important to take the individual's residential area into consideration, so that those living within the city of Rotterdam did not 'drive up' the benchmark, given that they had more opportunities to work. Individuals were then compared against the benchmarks calculated for them.

Essentially, the baseline was the expected proportion of young people who would successfully move off employment benefits (to work or study) without participating in the Buzinezzclub, and how quickly they were expected to do so.⁹⁶

The main outcome of interest was the average number of days for which unemployed youth tended to receive benefits. The goal was to reduce this average by 210 days compared with the median expected days of payment (748.3 days)⁹⁷. Other indicators related to transitions into jobs, education or self-employment, and to possible backsliding into allowances. As for backsliding, participants were compared with the expected backsliding rate, calculated for times of recession and non-recession times. No other outcomes/impacts were measured or discussed in the contract, even though it was highlighted that the intervention was thought to enhance empowerment, entrepreneurial skills and self-efficacy.

⁹⁴ Personal communication with the service provider (2020).

⁹⁵ Deloitte, statusrapportage 8, p. 25

⁹⁶ Interview with the investor (2020).

⁹⁷ Deloitte (2016), 3.

Barriers to and enablers of the measurement process

The main barriers to the measurement process concerned the selection of outcomes and metrics. For instance, there was a lack of consideration for wider outcomes such as improvements in health, reductions in debts or increases in wellbeing, or for potential side-effects of the intervention. In addition, the metrics selected should have considered participants' personal situation aside from their employment status, for instance including wage and salary jumps, and participants' satisfaction regarding the service and the job. Moreover, it was unclear how to evaluate outcomes for those participants who slid back into claiming unemployment benefits for just a short time⁹⁸.

Table 15. Summary of the measurement process in Buzinezzclub SIB

Measurement methodology	
Methods relating to payment mechanisms	Propensity score matching
Other evaluations	N/A
Methods using experimental or quasi-experimental design	Quasi-experimental
Control group	No
Causality of impact	No
Enabling factors and strengths of the evaluation process	All parties were involved in the selection of indicators
Barriers to and weaknesses of the evaluation process	Lack of emphasis on wider outcomes and side-effects; Lack of data from municipalities about some participants.
Governance of measurement	
Independence of evaluator	Yes
Evaluator	Deloitte

1.4.4. Effectiveness

Achievement of outcomes

The targets for the SIB were achieved. However, it was not possible to compare the SIB and the TF schemes because at the time of writing, the TF scheme had not yet been evaluated.

The problem of **creaming** emerged in the Buzinezzclub SIB, associated with the scheme's eligibility criteria. Only individuals registered as unemployed in the municipality could participate in the programme offered through the SIB in order to generate measurable savings (a pre-condition set out by the commissioner). However, adolescents who were not

⁹⁸ Dekker & Verhoeven (2019).

registered were automatically excluded from the programme as potential savings could be measured.

One long-term impact considered by the scheme was the sustainment of participants' occupations, thus preventing backsliding into unemployment benefits. A reduction in the rate of participants sliding back into unemployment benefits (compared with that forecast by the econometric model) was achieved.⁹⁹ Specifically, the backsliding rate was reduced by 8.4% in Year 1 and 3.8% in Year 2.¹⁰⁰

Benefits and drawbacks of SOC compared with TF

Before the implementation of the SIB, the municipality already provided social services to support unemployed adolescents, but the SIB intervention differed from the previous programmes in its long-term perspective. **The SIB therefore enabled a new intervention with uncertain outcomes to become a proven intervention**, when it would have been too risky for a TF model.

Investors claimed that the SIB is an interesting financial instrument for initiating change when it is used for as yet **unproven interventions** and around themes in which social benefits fall to different parties. However, after having experienced the use of a SIB, they admitted to **doubting the actual contribution of SIBs to innovation in the social domain**. The commissioner also agreed that SIBs are a useful way of testing new interventions, because they reduce risks and offer access to external finance. The commissioner also thought that once the impact of the programme is proved using the SIB, the municipality should fund the intervention traditionally.

Partnership was mentioned as a positive aspect of the scheme, as it appears the various parties involved in the SIB (the municipality, financial investors and social entrepreneurs) learnt to appreciate each other's positions by holding regular meetings. According to the service provider, one advantage of the SIB's design compared with 'traditional' commissioning was the **service-supplier relationship**, under which the municipality was contractually obliged to recruit participants. This made the municipality a "trustworthy partner"¹⁰¹. On the downside, however, the municipality had too little manpower available to support the execution of the SIB, and the turnover of the contact person was high. This appeared to be problematic when working with a fairly complex contract.

Another benefit of the SIB was to enable a **diversification of finance** within the municipality.¹⁰² In addition, the binding conditions of the SIB forced the municipality to meet concrete targets, and stimulated it to **implement more efficient supervision**¹⁰³. In this respect, the evaluator also reported that the use of a SIB forced the municipality to look more closely at the outcomes of the intervention¹⁰⁴ and to favour the implementation of an impact-assessment.

As a result of the SIB, the service provider was able to **access new sources of funding**¹⁰⁵, **scale its promising initiative** and replicate it in another municipality as a 'proven concept'.¹⁰⁶

⁹⁹ Deloitte, rapportage terugval

¹⁰⁰ Deloitte, rapportage terugval

¹⁰¹ Davelaar et al (2015).

¹⁰² Van Es (2015).

¹⁰³ Holvast, L. (2018). Social Impact Bonds: The Holy grail by doing good and making profit? Realistic Evaluation of Social Impact Bonds in the Netherlands. Wageningen University.

¹⁰⁴ Personal correspondence with the evaluator (2020 06 30).

¹⁰⁵ Van Es (2015)

¹⁰⁶ Holvast (2018)

Nevertheless, investors expressed concerns that **the use of SIBs had become a policy goal in itself**. Investors argued that “now SIBs seem more likely to be used on the basis of a political conviction that the market is better able to make a difference than public bodies...[Investors] therefore advocate continuing experimentation with SIBs, but only after a serious **ex-ante assessment** (i.e. research prior to a policy decision) in which the expected advantages and disadvantages of a SIB are determined on a case-by-case basis and alternatives are carefully weighed. Only then will this win-win situation be created through public-private partnerships in the social domain.”¹⁰⁷

Scalability and replicability

The SIB was initially intended to last two years, but was later extended to a third year. After its completion, Buzinezzclub then won a contract via public procurement. The programme was also replicated, through a SIB, in the cities of Utrecht and Eindhoven.

The service provider reported that some lessons learnt from the piloting experience in Rotterdam it applied in replicating the programme in other cities: more attention was paid to cooperation with the municipality, more time was devoted to linking together all stakeholders, and greater effort was made to increase the enthusiasm of provider's employees.

According to the investors, scaling up is only possible if the procedures relating to the SIB design and implementation are shared and acknowledged at all managerial and operational levels within the commissioning public institution, rather than being supported by a few civil servants.¹⁰⁸

¹⁰⁷ Dekker, F. & Verhoeven, J. (2019). Vijf Jaar Social Impact Bonds: De Balans Opgemaakt. Start Foundation.

¹⁰⁸ Scholten, P. (2015). Social impact bonds. Van fenomeen naar praktijk. Start Foundation.

1.4.5. Efficiency

Intervention costs

A cost comparison cannot be conducted for this scheme because the performance indicators for the equivalent TF intervention were not shared with the research team. Nevertheless, we estimate the cost per participant for the first two years of the Buzinezzclub SIB to be equivalent to EUR 7,003, and EUR 12,313 for each participant who found a job. This is based on the fact that in each of the first two years of the programme, the municipality issued a payment of EUR 560,246 to Stichting Social Impact Bonds Rotterdam (which then redistributed the money among investors).¹⁰⁹ During that time, 160 participants were recruited to the programme; of these, 91 moved off unemployment benefits.¹¹⁰

The commissioner interviewed noted that **traditionally funded programmes are usually cheaper** because no interest is paid, and they are easier to manage. Even so, the commissioner argued that **SIBs are a useful way to test new interventions** because they reduce risks and provide access to external finance. Once the impact of the programme is clear and it is working efficiently, the commissioner said, the municipality should fund the intervention traditionally to reduce costs.

The service provider added that SIBs enhance the efficiency of a programme because **investors identify dysfunctional procedures** in order to protect their investment. Furthermore, new programmes funded through SIBs create **competition with the services of the municipality**, motivating municipal services to be delivered more efficiently as well.

Operational costs

In addition to the costs of the SIB programme (see Table 13), the programme involved overhead costs for all parties of between EUR 50,000 and EUR 100,000.¹¹¹ An additional EUR 40,000 was also included in the SIB budget to pay for the intermediary. Equivalent costs for the TF scheme were not available to the research team, but the interviewees noted that the SIB involved more stakeholders, so communication between them was more time-consuming than in TF programmes. Nevertheless, the SIB was evaluated as being fairly efficient, because it relied on administrative records that were available to the municipality.

1.4.6. Design of the scheme

Design features

The Buzinezzclub SIB included the following design features that enhanced the programme's efficiency:

- **Payments to investors were tied directly to savings for the municipality.** In 2013, the average allowance for an individual receiving unemployment benefits was EUR 39.70 per day. This information was used when determining payments to investors, to ensure that they were not compensated beyond the level of any savings to the municipality.¹¹²

¹⁰⁹ Deloitte (2016), 3.

¹¹⁰ Deloitte, rendement rapportage

¹¹¹ Van Es (2015)

¹¹² Van Es (2015); Start Foundation

- **Job sustainability was incorporated into the payment mechanism:** the SIB not only measured how many people found jobs (and thus, for how many days they claimed benefits), but also if and when they started to claim benefits again (the backsliding rate). Investors were compensated separately for both outcomes achieved.¹¹³
- To account for fluctuations in the labour market, **different outcome targets were specified for times of crisis and non-crisis times.**
- **The municipality was legally bound to deliver a certain number of participants to the programme.** This led to a certain amount of stability for the service provider, and enabled Buzinezzclub to fulfil its promises by delivering the agreed results.¹¹⁴
- **The service providers appreciated the presence of an intermediary.** Intermediaries can make sure that investors do not take over the supervisory role of the municipality and interfere, for fear of losing the investment.¹¹⁵

Nevertheless, the SIB also included some design elements that were perceived as more controversial:

- Multiple stakeholders noted that two years was too short a time to foster innovation. Therefore, they argued that **longer contracts** should be in place for SIBs.
- **The service provider contributed 40% of the total investment.** While this was intended to motivate Buzinezzclub to work efficiently, as well as potentially providing the social enterprise with greater financial independence in the event of success, investments from service providers are uncommon in SIBs because they can decrease the innovativeness or flexibility of the intervention. In other words, if service providers have more to lose, they take safer approaches, which contradicts the purpose of testing out new interventions.¹¹⁶
- **Enthusiasm for the SIB was not shared equally by all members of the municipal administration.** The organisation of new finance models is time-intensive for civil servants.¹¹⁷ It is therefore important to make sure that not only upper, but also middle management within the municipality are on board with regard to moving forward with the SIB, because they ultimately carry out the bulk of the work.

Impact on third-sector organisations and social enterprises

While Buzinezzclub is a for-profit company rather than an NGO, it is nevertheless a social enterprise, so we also consider the impacts of the SIB on Buzinezzclub. According to a Buzinezzclub representative, the SIB offered clear advantages compared with a traditional commissioner-service supplier relationship, since it is an agreement under which the municipality is contractually obliged to recruit participants. This makes the municipality “a trustworthy partner, and not one that can choose to send you first zero and the next time 100 people.”¹¹⁸

Furthermore, Buzinezzclub used to be financed by private funds or its own resources. Through the SIB, Buzinezzclub was able to access new sources of funding.¹¹⁹ Lastly, the

¹¹³ Deloitte, Terugvalrapport section 1; 2

¹¹⁴ Personal communication with the investor (2020).

¹¹⁵ Boukema, R. (2017). Roadmapping A New Social Impact Bond In Rotterdam. Leiden University.

¹¹⁶ Boukema (2017)

¹¹⁷ Start Foundation (2015)

¹¹⁸ Davelaar et al (2015).

¹¹⁹ Van Es (2015).

SIB enabled the organisation to scale its initiative to another municipality, because it provided evidence that the approach worked.¹²⁰ Overall, the SIB therefore appears to have had a positive impact on the service provider.

1.5. BOAS Werkt (Netherlands and Germany)

1.5.1. Background

When the BOAS Werkt scheme began in 2016, the city of Enschede (close to the German border) had an **unemployment** rate of 8.5%. By contrast, the German region of North Rhine-Westphalia, just over the border, had a very low unemployment rate of around 3%. The number of unfilled vacancies in North Rhine-Westphalia was expected to rise, as large sections of the working population were reaching retirement age. However, few unemployed jobseekers from Enschede seemed to find their way to the German jobs market unaided.

This scheme was an attempt by the municipality of Enschede, investors Start Foundation, ABN AMRO and service provider BOAS Werkt to address unemployment in Enschede, making use of a SIB as a financial instrument. Outcomes-based funding enabled the municipality of Enschede to collaborate with an organisation such as BOAS Werkt. The promise of only paying for successful transitions away from unemployment benefits was seen as a risk-reducing factor, and the investors in the scheme would come up with the financial investment required to get the scheme up and running. The SIB targeted the 7,000 households in Enschede receiving unemployment or social assistance benefits.

Jobseekers who took part in the programme spent four weeks of training at Kreishandwerkerschaft Steinfurt-Warendorf, a German partner of BOAS Werkt. This training consisted of activities in various trades including construction, metalworking, plumbing and welding. The training focused on general technical skills, safety, German language and culture, differences between the German and Dutch working cultures, and a visit to a German employer. The training was followed by a practical placement and, if both parties agreed, an employment contract for at least six months with that German employer. The specific **objective** was to get at least 75% of the 138 participants into jobs. These individuals came from among the 7,000 households receiving benefits in Enschede. Of those who gained a job, 90% were supposed to find sustainable employment, meaning that they should acquire a job contract of no less than 12 months.¹²¹ BOAS Werkt only selected employers who paid at least EUR 12/hr.¹²²

The main objective of the programme was not achieved, mostly due to lower than anticipated enrolment in the programme. Due to low take-up, the programme was discontinued at the end of 2017, despite being scheduled to last until 2018. In total, 19 people found jobs, although nine of whom claimed unemployment benefits again within the next two years.¹²³ However, in comparison to a control group, tracked over a period of two years, participants in the scheme claimed unemployment benefits significantly less frequently. The estimated financial benefit of the scheme (in comparison to control group participants) is estimated at EUR 9,960 per participant¹²⁴, or EUR 300,000 in total.

The programme was supposed to **cost** EUR 1.1 million¹²⁵ (of which the service provider contributed 5%¹²⁶) over its intended 30-month duration between 2016 and 2018. The

¹²⁰ Holvast (2018).

¹²¹ Scheltinga, P. & Nijhof, B. (2019). Concept Eindrapportage 'Monitoring Social Impact Bond Werken in Duitsland'. Kennispunt Twente, p. 1

¹²² Municipality of Enschede (n.d.) Factsheet Social Impact Bond 'Werken in Duitsland' in het kort. Retrieved from: <https://www.enschede.nl/sites/default/files/factheet-sib-werken-in-duitsland--online-versie-gem-enschede.pdf>

¹²³ Scheltinga & Nijhof (2019), p. 6

¹²⁴ Scheltinga & Nijhof (2019), p. 8

¹²⁵ Municipality of Enschede (n.d.)

¹²⁶ Municipality of Enschede (n.d.)

contract specified interim payments by the investors to the service providers in 10 instalments (each instalment covering a quarter of a year and targeting 13-14 participants). In the event of success, investors were to be repaid their investment, plus an additional % interest rate of 10. Financial returns based on savings (calculated on the basis of the evaluation results, but also including a fixed amount) were to be paid by the municipality every six months, starting from month 24. Payments were issued to and from a foundation set up for this purpose.¹²⁷ Payments were issued according to the contract; however, they stopped after the contract was discontinued. The final financial settlement is currently under negotiation.

For a summary of the scheme, please see the table below.

Table 16. Summary of the Boas Werkt SIB

Personally targeted social service	Labour market-related
Target population	Unemployment benefit claimants in the municipality of Enschede
SOC scheme type	Direct SIB
Payment model	Binary
Cost	EUR 1.1 million ¹²⁸
Commissioner(s)	Municipality of Enschede
Social service provider(s)	BOAS Werkt (for-profit enterprise)
Investor(s)	Start Foundation (social philanthropic organisation) ABN AMRO (private investor)
Intermediary	N/A
Evaluator(s)	Kennispunt Twente

1.5.2. Equivalent TF scheme

The Province of Drenthe, UWV (Employee Insurance Agency), and the municipalities of Borger-Odoorn, Coevorden, Emmen, Hardenberg and Eems Dollard Region are all working together in the context of cross-border mobility. Under the umbrella of 'Work in Germany'¹²⁹, they involve German partners and various municipalities in the Dutch provinces of Groningen and Drenthe.

Within this broader initiative, we focus on efforts within the Dutch municipality of Emmen, which launched a so-called 'Transfer Point'¹³⁰ initiative in 2019 that includes various activities to support the cross-border mobility of workers from the Netherlands to Germany.

The target group is roughly the same as in Enschede, and the two cities are roughly comparable in terms of size. Both cities are also located along the border with Germany. Despite these clear similarities, it should be noted that the approach taken by BOAS Werkt was somewhat different, in particular with regard to its dedicated practical training programmes. Whereas BOAS Werkt paid ample attention to job activities, Transfer Point Emmen focuses more on providing generic support concerning red tape, customised to the needs of individual users living in the Netherlands but working in Germany. Furthermore, instead of actively searching for participants like BOAS Werkt, Transfer Point Emmen provides support for individuals who express interest. Therefore, comparisons can only be made in terms of the advantages and disadvantages of the two programmes separately, rather than in relation to one another.

¹²⁷ Personal communication with the service provider (2020).

¹²⁸ Municipality of Enschede, Factsheet SIB, via enschede.nl

¹²⁹ Werken in Duitsland (2020). Retrieved from: <https://www.werkeninduitsland.nl/>

¹³⁰ Werken in Duitsland. (2020). Persbericht. Retrieved from: <http://persberichten.deperslijst.com/118466/uwv-persbericht-uwv-werkbedrijf-emmen-maakt-werken-over-de-grens-eenvoudig-met-transferpunt-werken-in-duitsland.html>

Table 17. Comparability of the Boas Werkt SIB with Transfer Point

	Same as SOC	Comments
Personally targeted social service	Yes	Labour market-related
Target population	Yes	Unemployment benefit claimants
Number of participants	No	SOC: 19 people found jobs in Germany during the 1.5 years of the SIB ¹³¹ TF: 13 people found jobs under the TF programme in 2019-2020 ¹³²
Location	No	SOC: Enschede TF: Emmen
Cost	No	SOC: EUR 1.1 million, or roughly EUR 440,000 per year ¹³³ TF: roughly EUR 180,000 per year ¹³⁴
Commissioner(s)	No	SOC: Municipality of Enschede TF: UWV Emmen in collaboration with the Municipality of Emmen
Social service provider(s)	No	SOC: BOAS Werkt (for-profit enterprise) TF: UWV branch Emmen; the UWV (Employee Insurance Agency) is an autonomous administrative authority (ZBO) and is commissioned by the Ministry of Social Affairs and Employment (SZW) to implement employee insurance and provide labour market and data services.

1.5.3. Outcomes measurement

No evaluation details are available with regard to **payment-related outcomes** for the BOAS Werkt SIB. Its **outcome** target was to achieve successful transitions from unemployment benefits to jobs for 75% of programme participants, out of which 90% were to be in sustainable employment (meaning that they should acquire a job contract of no less than 12 months).¹³⁵ Due to lower than anticipated enrolment in the programme, it was clear that the SIB would not meet its objective of getting at least 138 individuals off benefits and into work. It is therefore possible that a formal evaluation of whether the outcome was achieved was never carried out.

Nevertheless, a separate impact assessment was conducted by Kennispunt Twente, a government-related non-profit agency. Although the work it carries out is government-related, the organisation claims to be based within Regio Twente, an independent

¹³¹ Scheltinga & Nijhof (2019), p. 6.

¹³² Personal correspondence with the TF commissioner (2020).

¹³³ Municipality of Enschede, Factsheet SIB, via enschede.nl

¹³⁴ Estimate based on 2.5 FTE, gross salary EUR 4,500 and 33% additional premiums

¹³⁵ Scheltinga & Nijhof (2019), p. 1

government unit with its own management and budget.¹³⁶ Below we briefly describe the methodology used for this assessment. Given that this assessment did not determine whether or not investors were repaid, we classify it as ‘Other evaluation’.

Other evaluations

The **impact assessment** carried out by Kennispunt Twente focused on public cost savings. The municipality and BOAS Werkt provided Kennispunt Twente with information about the 33 participants who enrolled in the programme. The evaluator then identified 96 Enschede residents to serve as the control group. The two groups were matched on the basis of gender, education, occupational history, living situation, age group, and the number of months they had already been on benefits. The resulting groups were largely comparable on the basis of these characteristics, with the exception of education and sectors of occupation. The control group appeared to be more educated. However, this was not a significant issue because it implied that the estimated public savings among the control group would be lower-bound.

Employment records and benefit claims were gathered for all participants over the period of two years. Based on this information, the evaluator determined how many days each person in the treatment and control group claimed unemployment benefits, which were equal to EUR 38.01 per day. In the control group, the total number of days for which group members claimed benefits was divided by 2.9, reflecting the fact that the control group was almost three times larger than the treatment group. The resulting number of days for which unemployment benefits were claimed was multiplied by EUR 38.01 to see how much each group had claimed in unemployment benefits over the two years of the programme.¹³⁷

Barriers to and enablers of the measurement process

One weakness of the measurement process was that the metrics focused on job acquisition, without considering other outcomes. Furthermore, the investors reported that the main weakness of the evaluation was that **user satisfaction did not play a role in evaluating the project’s success.**

Measurement methodology	
Methods related to payment mechanisms	No evaluation available related to payment-related outcomes
Other evaluations	Impact assessment
Methods using experimental or quasi-experimental design	No
Control group	Yes
Causality of impact	Partial (control group was robust, but participants were not randomly allocated to the programme, so the evaluation cannot establish the impact taking unobserved characteristics into account)

¹³⁶ Kennispunt Twente. (n.d.). Over Kennispunt Twente. Retrieved from: <https://www.kennispunttwente.nl/over-ons/over-kennispunt-twente>

¹³⁷ Dekker & Verhoeven (2019).

Enabling factors and strengths of the evaluation process	Checks on intermediate results
Barriers to and weaknesses of the evaluation process	The focus of the evaluation was on jobs acquired and savings generated for the municipality. No wider impacts on beneficiaries (such as job satisfaction) were assessed.
Governance of measurement	
Independence of evaluator	Yes
Evaluator	Kennispunt Twente

1.5.4. Effectiveness

Achievement of outcomes

As mentioned above, the target (75% off benefits and into jobs) was not achieved, because only 58% of people trained found a job. Nonetheless, the evaluation still reports overall savings. The main reason for the programme's failure to achieve its target was a lower than expected participation rate. While stakeholders expected to recruit at least 55 participants into the programme each year,¹³⁸ the true take-up rate was roughly half that. Some stakeholders ascribed the lower participation to improving economic conditions, which had led to lower unemployment rates in the region. However, the service provider noted that the municipality had initiated a new programme for the unemployed at the same time as the SIB scheme began, so the entire target population was brought under the new programme. This resulted in lower participation in the SIB scheme.

Moreover, the obligation of the municipality to refer a set number of participants to the programme may have resulted in **cherry picking**. According to one interviewee, the commissioner referred some participants that were the "easiest to help", meaning that they could have been able to overcome their problems without taking part in the programme.

The impact assessment carried out by Kennispunt Twente revealed that **participants in the SIB programme were half as likely to claim benefits as the control group**, demonstrating that the programme had an impact despite its early discontinuation.

It was not possible to compare the SIB and TF schemes in terms of outcome achievement, because at the time of writing the TF Transfer Point scheme had not yet been evaluated. Between January 2019 and July 2020, Transfer Point helped had 13 beneficiaries to find a job in Germany.

Benefits and drawbacks of SOC compared with TF

The SIB model was useful for service providers because it offered clarity as to the size and scope of a project in terms of providing a good estimate of the costs involved and the potential returns it may yield. Moreover, the SIB enabled service providers to implement programmes on a larger scale than interventions that were financed traditionally. The investors in BOAS Werkt saw the main benefit of using a SIB model in the ability to experiment with a new intervention that had uncertain outcomes. This would not have been possible under a TF scheme, due to the level of risk involved. It was also mentioned by the

¹³⁸ Scheltinga & Nijhof (2019), p. 2

service provider that without the SOC there would probably not have been a comparable TF project.

Perceptions of the effectiveness of this mechanism were mixed. The first problem was that expectations in terms of the number of participants to enrol in the program were too ambitious, given the improving economic climate and the fact that a similar TF programme was launched by the municipality. Providers also reported some difficulties in cooperating with the municipality because the SIB was not broadly accepted and rooted within local government.

Scalability and replicability

The programme was originally intended to last until 2018, but was discontinued at the end of 2017. Despite talks between the service provider and various interested local governments, nothing materialized. The reasons for this are probably the failure of the project, and the fact that unemployment decreased due to improved macro-economic conditions.

1.5.5. Efficiency

Intervention costs

Given that the BOAS Werkt SIB was discontinued early due to lower than anticipated take-up, the stakeholders are still in the process of negotiating a financial settlement. It is therefore not possible to estimate the cost to the taxpayer of the jobs acquired. Nevertheless, we can assess to some extent the cost to the investor of the jobs acquired, relative to the costs thus far of the similar programme in Emmen, although this assessment is based on a number of assumptions.

The total budget of the SIB was EUR 1.1 million for a total of 30 months between 2016 and 2018.¹³⁹ The contract specified interim payments by investors to the service providers in 10 instalments, with each instalment covering a quarter of a year and targeting 13-14 participants.¹⁴⁰ The programme was discontinued in 2017, 15 months after its launch, so the investor is likely to have issued five instalments to the service provider before the programme was terminated. Assuming these instalments were paid in equal amounts, that would bring the total investment to EUR 550,000. During the programme, 33 participants received training and 19 of them found jobs,¹⁴¹ so the cost per participant comes to EUR 16,667, and the cost per job acquired is equivalent to EUR 28,947. Please note that we do not include the 10% rate of return for investors in these calculations because, given the SIB's early termination, we assume the interest was not/will not be paid, although this is still the subject of negotiations between the municipality and the investor.

The costs of the Transfer Point are even more difficult to quantify. According to the commissioner, staff salaries constitute the main cost and amount to 2.5 full-time equivalents. Though compensation varies between employees, the commissioner estimated that a typical employee at UWV Emmen earns roughly EUR 4,500 in gross monthly wages, and is entitled to approximately 33% in premiums. Given that the Transfer Point was initiated in 2019 and outcomes were reported to the research team in July 2020, we estimate that in roughly 1.5 years, the programme has cost EUR 269,325. In total, 13

¹³⁹ Municipality of Enschede (n.d.)

¹⁴⁰ Personal communication with the service provider (2020).

¹⁴¹ Scheltinga & Nijhof (2019), 1 & 6.

people have acquired jobs since the programme launched in 2019¹⁴², which brings the cost per participant who found a job to roughly EUR 20,717.

Table 18. Cost comparison between the BOAS Werkt SIB and Transfer Point (EUR)

	Boas Werkt SIB 2016-2017	Transfer Point (TF) 2019-July 2020
Participants who found jobs	19	13
Total cost	550,000	269,325
Cost per participant who found a job	28,947	20,717

Source: elaborated by PPMI, based on the available information.

Note: costs have not been adjusted for inflation. This adjustment would further increase the costs of the SIB relative to those of Transfer Point. Transfer Point's only costs involve staff salaries. The costs of BOAS Werkt only include costs incurred by the investors, not the commissioner.

While the cost per outcome achieved appears to be substantially lower for the Transfer Point programme than for the BOAS Werkt SIB, **it is important to qualify these findings**. First, the cost of Transfer Point disregards the rent, utilities, etc., required to run the UWV branch. These costs could substantially increase our estimate. Second, the nature of the programmes is different: whereas the SIB funded training, Transfer Point is designed mostly to assist interested persons in preparing the administrative documents required to work in Germany. Therefore, **the different costs more likely reflect the different nature of the programmes rather than the effectiveness of the funding models (SIB vs TF)**. Finally, the cost per job in the SIB is high not because the programme failed to help its participants, but because too few participants were recruited to the programme.

Given the substantial limitations involved in drawing a cost comparison, the BOAS Werkt case is perhaps more useful for investigating the way in which investors or service providers should be compensated if the number of participants turns out to be substantially lower than was assumed during the planning of the intervention. In other words, **who bears the responsibility for recruiting a sufficient number of participants?** In the BOAS Werkt SIB, the municipality and BOAS Werkt agreed that they would aim to recruit 55 participants per year, with a commitment from the municipality to recruit 40 participants and BOAS obliged to recruit 15.¹⁴³ Therefore, when few participants joined the programme, the municipality was required to pay penalties to the investor. The investor highlighted this as one of the most important **contractual provisions**, because investors would have otherwise have lost their investment. Failure to recruit participants resulted not only in financial penalties, but inefficiencies more broadly. According to one municipality officer not directly involved in the SIB, the municipality referred some participants to BOAS Werkt that the municipality could have placed into jobs itself, solely in order to comply with the recruitment requirements set out in the contract.

Overall, this SIB raises questions regarding the extent to which commissioners can transfer the risk associated with an SOC scheme. While the risk associated with the effectiveness of the programme was largely transferred to investors, the municipality also took on risk by committing to recruit participants – a risk that ultimately materialised.

¹⁴² Personal correspondence with the commissioner (2020).

¹⁴³ Scheltinga & Nijhof (2019), p. 2

Operational costs

It is not possible to compare the overhead costs of BOAS Werkt SIB and the Transfer Point programme due to limited information about the latter.

Regarding the **setting-up of the SIB**, it took about 18 months to settle the contract between the parties concerned. There is no account of the number of hours and costs involved in this process, but the service provider mentioned that the number of hours was very significant, also considering the cross-border nature of this scheme: it was also necessary to acquire and negotiate with German stakeholders. Since the SOC scheme is not embedded within the workflow of the municipality, extensive communication was needed to align individuals within the municipality too.¹⁴⁴

Conversely, the **evaluation** method adopted did not take much time because it was conducted by an external evaluator, drawing on administrative data.

1.5.6. Design of the scheme

Design features

Although the SIB failed, it nevertheless offers important lessons learned about effective and efficient design features of SIBs:

- The first lesson learned addresses **participant recruitment**, given the voluntary nature of the programme. Many programmes require a sufficient scale to operate cost-effectively and to demonstrate whether the intervention is effective. The BOAS Werkt SIB recognised this, and therefore **minimum required participation levels were specified in the contract**. The SIB designers went a step further and specified **which parties were expected to recruit participants, how many, and the penalties associated with failure to do so**. These are important contractual provisions for any programmes in which participation is voluntary. Reflecting on the process, however, it seems crucial to conduct a **pre-assessment to investigate the extent to which a legally binding supply of participants is feasible**.¹⁴⁵ Furthermore, the contract should address how all stakeholders – not just the investors – will be compensated if one of the parties withdraws from the contract.
- One of the reasons why few participants enrolled in the programme was an improvement in macro-economic conditions. This change was not expected at the time the SIB was set up, which led to SIB being, at least in part, superfluous. The **contract should therefore consider whether the intervention can continue in the event of changing macro-economic conditions and, if so, whether targets should be adjusted**.
- The launch of a new programme for the unemployed may also have reduced enrolment in the SIB programme. Therefore, the **SIB should be well-rooted within the municipality** to make sure that two competing programmes are not launched simultaneously. This is also important because the SIB demands a lot of communication across all layers of the municipality. The stakeholders noted that when a civil servant changed her or his position, a significant amount of knowledge and commitment was lost in the process.¹⁴⁶

¹⁴⁴ Start Foundation (2015)

¹⁴⁵ Start Foundation (2015)

¹⁴⁶ Personal correspondence with the investor (2020).

1.1.1. Impact on third-sector organisations and social enterprises

The failure of the SIB had a **negative impact** on the service provider: although the provider had talks with other local governments about replicating the same intervention, nothing materialised from these talks. This was, however, also partly due to the improving economic conditions.

1.6. Adolescent Behavioral Learning Experience (ABLE) Program at Rikers Island (USA)

1.6.1. Background

Between 2013 and 2015, the New York City Department of Correction (DOC), in collaboration with the Mayor's Office and a number of non-profit and private sector partners, designed and implemented a large-scale programme, Adolescent Behavioral Learning Experience (ABLE), for **adolescents aged 16 to 18 detained at the city's Rikers Island jail (Rikers)**. The ABLE SIB was developed on the basis of some of the work New York City had been doing to understand why young men of colour fail to achieve educational and social outcomes. The city investigated recidivism rates among young adolescents detained in jail (either serving a short-term sentence or not yet convicted but unable to make bail), and found that over 50% would return within one year. The DOC and the city's Office of Management and Budget had little (if any) discretionary funding for innovation and experimentation. Funding a programme that might reduce the likelihood of future incarceration would have been a stretch for an agency consumed by day-to-day crises. Like many resource-constrained government agencies, the DOC struggled to effectively manage daily life at the jail — from securing, feeding, clothing and housing inmates to educating them. External resources provided under the ABLE programme made it possible for the DOC to try out a new way of preventing recidivism. If the SIB had not provided the necessary 'risk capital', the DOC would have been unlikely to experiment with cognitive behavioural therapy for the adolescents it housed. The SIB was conceived as a strategy to attract private investment for preventive services, particularly important in a time of constrained government spending.

The ABLE programme aimed to break the cycle of reincarceration for adolescents in jail using **Moral Reconnection Therapy (MRT)**, an evidence-based intervention that focuses on improving decision making.¹⁴⁷ MRT was designed to equip adolescents with social and decision-making skills in order to help them to take responsibility for their actions, avoid rash and impulsive reactions, make more constructive life choices, and, ultimately, avoid a costly return to Rikers. It involves the delivery of a 12-step curriculum by a trained facilitator in 50-minute group sessions every day.¹⁴⁸ The SIB targeted the 3,000 adolescents incarcerated at Rikers Island each year, with the intention of getting as many participants as possible through the 12-stage MRT process.

Cognitive behavioural therapy using MRT had been provided before in jails and prisons across the US. Indeed, the choice of the intervention delivered under the ABLE SIB was based on previous results achieved by this type of intervention. However, delivering such a programme to the **specific population** targeted by the ABLE SIB – namely, youth between the ages of 16 and 18 in jail – was an innovation. Moreover, MRT has been used in a variety of settings, including jails, but never on this scale. Thus, the demands of this SIB required the programme to operate on an **unprecedented scale** within the service

¹⁴⁷ Parsons, J., Weiss, C. & Wei, Q. (2016). Impact evaluation of the Adolescent Behavioral Learning Experience (ABLE) Program at Rikers Island. New York: Vera Institute of Justice. Retrieved from: <https://www.vera.org/publications/rikersadolescent-behavioral-learning-experience-evaluation>.

¹⁴⁸ Berlin, G.L. (2016). Learning from Experience: A Guide to Social Impact Bond Investing. MDRC. p. 3.

environment. Each day, the DOC housed approximately 600 adolescents aged 16 to 18, making ABLÉ the largest MRT programme to date.¹⁴⁹

The SIB was structured as follows¹⁵⁰. Goldman Sachs funded the project’s delivery and operations through a USD 9.6 million loan to the intermediary MDRC, paid out quarterly over four years¹⁵¹. Bloomberg Philanthropies provided a USD 7.2 million grant to the MDRC, funded annually over four years, to guarantee a portion of the loan, reducing the lender’s risk. MDRC, through a contract with the City, oversaw the day-to-day implementation of the project, managed two service providers (Osborne Association and Friends of Island Academy), and was responsible for repayments to the private investor. The Vera Institute of Justice, an independent evaluator, determined whether the project achieved the targeted reductions in reincarceration.

The Department of Correction **paid** MDRC on the basis of reduced re-admissions and the associated cost savings. Repayment to Goldman Sachs was determined on the basis of two factors: the impact on readmission bed days (RBDs, also called “future days in jail”) in NYC Department of Correction (DOC) custody during the two years following the release of each young person, and the number of participants served by the scheme over a four-year period. If recidivism rates fell by 10% relative to a comparison group, the city would pay back Goldman Sachs in full; if the programme reduced recidivism by more than that amount, the city would pay an additional return according to a capped, sliding scale, as presented in Figure 1. As shown in the figure, the deal also included the possibility of a partial payment if recidivism was reduced by 8.5%.¹⁵²

Figure 1. Success payment schedule for the Rikers Island SIB

Reduction in Re-Admission Rate	Projected Long Term City	
	Net Savings (\$)	City Payment to MDCR
≥20.0%	\$20,500,000	\$11,712,000
≥16.0%	\$11,700,000	\$10,944,000
≥13.0%	\$7,200,000	\$10,368,000
≥12.5%	\$6,400,000	\$10,272,000
≥12.0%	\$5,600,000	\$10,176,000
≥11.0%	\$1,700,000	\$10,080,000
≥10.0% (<i>breakeven</i>)	\$ ≥ 1,000,000	\$9,600,000
≥8.5%	\$ ≥ 1,000,000	\$4,800,000

Source: Berlin (2016), 4.

For the purposes of repayments under the programme, “success” or “failure” is defined narrowly in terms of taxpayer savings, which are in turn based purely on operational savings within the jail system. The calculation of savings involves the following steps: i) estimating the number of programme participants; ii) estimating the baseline Bed Days Per Participant by using the DOC’s historical cohort analysis; iii) calculating the effect of the ABLÉ Program on Projected Bed Usage (calculating projected number of beds used with and without ABLÉ); iv) turning Bed Day Savings into Cost Savings (see Figure 2).

To trigger a complete repayment, a minimum of 9,240 participants had to be served during the first four years of the programme’s operation.¹⁵³

¹⁴⁹ Rudd, T., Nicoletti, E., Misner, K. & Bonsu, J. (2013). Financing Promising Evidence-Based Programs: Early Lessons from the New York City Social Impact Bond. MDRC, 41.

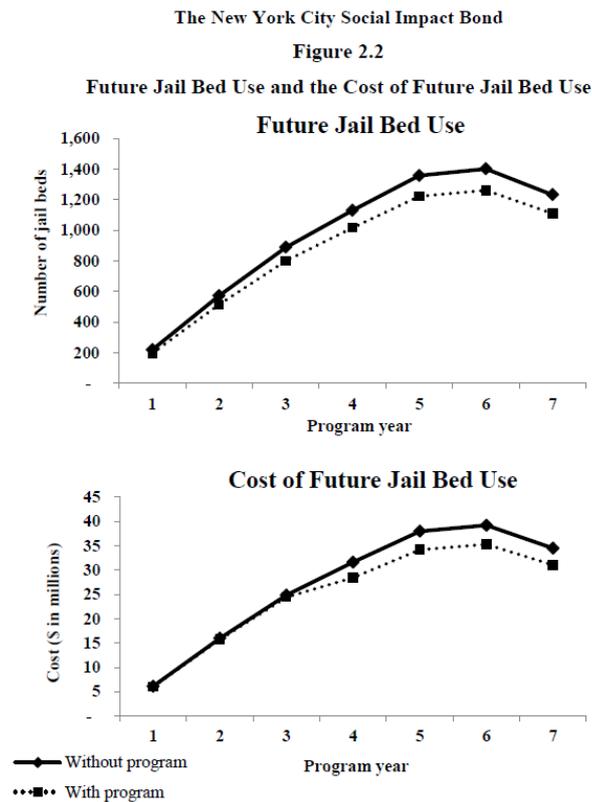
¹⁵⁰ Rudd et al. (2013).

¹⁵¹ Founded in 1974 as the Manpower Demonstration Research Corporation, “MDRC” became the official name of the organisation in 2003.

¹⁵² Berlin (2016), p. 4.

¹⁵³ 1,440 participants in Year 1; 2,500 in Year 2; 2,500 in Year 3; 2,400 in Year 4, plus an additional 400 participants who could be engaged at any time during these four years. Rudd et al. (2013).

Figure 2. Future jail bed use and the associated cost



Source: Rudd et al. (2013), 14.

At the three-year point in the four-year project, the partners reviewed one-year follow-up data. If these early results had shown that recidivism had been reduced by more than 9%, then the programme would have continued to be funded for an additional year, and the city would have made a USD 2.4 million success payment to the investors.

However, the evaluation performed by Vera Institute for Justice determined that the ABLÉ programme **did not reduce recidivism** among adolescents aged 16 to 18 after one year of follow-up.

Since the targets were not met, the investors ended the programme and the city did not pay anything. By the time the scheme ended, Goldman Sachs had invested USD 7.2 million, USD 6 million of which was guaranteed by Bloomberg Philanthropies, leaving Goldman Sachs with a loss of USD 1.2 million.¹⁵⁴

While not entirely due to the ABLÉ SIB, in 2016 the administration of Mayor Bill de Blasio, under intense pressure to improve conditions in the Rikers Island jail complex, developed a plan to move 16- and 17-year-olds from Rikers Island to a dedicated jail for adolescents in the Bronx. Indeed, New York was the only US state other than North Carolina that prosecuted all children aged 16 and over as adults when accused of a crime. In 2017, New York passed the Raise the Age legislation, which raised the age of criminal responsibility to 18 years by October 2019.

Table 19. Summary of the ABLÉ programme

Personally targeted social service	Social exclusion
Target population	Adolescent prisoners

¹⁵⁴ Berlin (2016), p. 8.

SOC scheme type	Managed SIB
Payment model	Frequency
Cost	USD 9.6 million
Commissioner(s)	New York City Mayor's Office New York City Department of Correction
Social service provider(s)	The Osborne Association, non-profit organisation Friends of Island Academy, non-profit organisation
Investor(s)	Goldman Sachs, private investor (provided loan) Bloomberg Philanthropies, philanthropic investor (guaranteed loan)
Intermediary	MDRC, non-profit organisation
Evaluator(s)	Vera Institute of Justice, non-profit research and policy organisation

1.6.2. Equivalent TF scheme

Given the impossibility of reaching any interviewees other than the investor, our analysis of comparable TF schemes is based on desk research. We chose the TF scheme on the basis of similarities between the two interventions (SIB and TF), as well as the availability of data to substantiate findings.

The comparable TF programme chosen is RESTART. RESTART is a jail-based transition and re-entry programme designed to reduce recidivism. It is a collaborative, evidence-based initiative involving the following parties: the Dutchess County Jail, the Office of Probation and Community Corrections, the Department of Behavioral and Community Health, Family Services, Inc. (formerly Hudson Valley Mental Health, Inc.), and Project Model Offender Reintegration Experience (Project MORE). Project MORE is a not-for-profit corporation based in Connecticut that specialises in behaviour reform and alternatives to incarceration. It provides several community correctional services for Dutchess County. RESTART is funded by appropriations through Dutchess County Jail. This programme is designed for individuals at higher risk of reoffending, as determined by a risk assessment and an interview process. The duration of the jail-based portion of the programme is six to eight weeks, followed by an ongoing community-based programme tailored to each individual.

The TF programme and the SIB share some comparable features: first, they both aimed to reduce recidivism, and included interventions based on cognitive behavioural therapy through Moral Reconciliation Therapy (MRT) delivered in a jail setting. Moreover, the region in which they were implemented was fairly comparable, with both programmes having taken place in New York State, even if the specific prison settings of the two schemes differ. However, the RESTART programme specifically targets individuals at higher risk of reoffending, which is not the case for the SIB, and unlike the ABLE SIB, is not specifically targeted at adolescents (see table below).¹⁵⁵

Table 20. Comparability of ABLE programme with RESTART

	Same as SOC	Comments
Personally targeted social service	Yes	Reoffending

¹⁵⁵ Lois, R. L. (2019). Audit report Project M.O.R.E., Inc. Office of the Dutchess County Comptroller.

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

	Same as SOC	Comments
Target population	No	SOC: Adolescents aged 16 to 18 detained at the city's Rikers Island jail TF: Male and female inmates housed in Dutchess County Jail (DCJ) who have a high risk for reoffending.
Number of participants	No	SOC: 1,470 adolescents participated in at least one ABLE session during 2013 ¹⁵⁶ TF: between 240 and 400 participants were admitted to the programme per year ¹⁵⁷
Location	Yes	New York State
Cost	No	SOC: USD 9.6 million ¹⁵⁸ over four years, roughly USD 2.4 million per year TF: USD 250,000 per year ¹⁵⁹
Commissioner(s)	No	SOC: New York City Mayor's Office New York City Department of Correction TF: Dutchess County Sheriff's Office, Corrections Division
Social service provider(s)	No	SOC: The Osborne Association, non-profit organisation Friends of Island Academy, non-profit organisation TF: Project Model Offender Reintegration Experience (Project MORE), non-profit corporation

1.6.3. Outcomes measurement

Measurement of outcomes relating to payment mechanisms

The ABLE evaluation used a **quasi-experimental design**, namely a propensity-score matching approach. Data were based on administrative records and programme participation data provided by NYC government agencies and the agencies responsible for implementing the ABLE programme (The Osborne Association and Friends of Island Academy).¹⁶⁰

The study cohort was defined as 16- to 18-year-old youth who were admitted to Rikers between 1 January and 31 December 2013. To allow for a one-year post-release tracking period, only those young people who were released on or before 31 March 2014 were included in the cohort. In addition, only those young people who were held in the jail for seven days or longer were included in the cohort.¹⁶¹

¹⁵⁶ Parsons et al. (2016), 10.

¹⁵⁷ Lois (2019), 12.

¹⁵⁸ Phillips, A., & Olson, J. (2013). Rikers Island: the first social impact bond in the United States. Community Development Investment Review, 097-101.

¹⁵⁹ Lois (2019), 5.

¹⁶⁰ Parsons et al. (2016), 9.

¹⁶¹ Parsons et al. (2016), 10.

The control groups were:

- A “historical’ control group made up of 16- to 18-year-olds who were incarcerated at Rikers between 2006 and 2010. To ensure comparability, this group was identified using the same criteria as the study cohort (held for seven or more days and released by 31 March of the year following admission.) This was used to control for individual-level factors.
- A control group of 19-year-olds incarcerated at Rikers over the same time period of the SIB intervention. Since ABLE services were provided to all 16- to 18-year-olds, this ensured that 19-year-olds included in the analysis could not have participated in the ABLE programme. To ensure comparability, this group was identified using the same criteria as the study cohort (admitted to Rikers in 2013, held for seven or more days, released by 31 March 2014). This group was used to control for system-level factors.¹⁶²

Vera researchers compared the change in RBD among the ABLE study cohort against the matched historical comparison group. The difference in outcomes between these two groups represented a 13.4% increase in recidivism among the ABLE cohort, without controlling for system-level factors that may impact recidivism. To assess the impact of system-level factors on recidivism, researchers compared RBDs with the matched samples of 19-year-olds. The analysis assumed that, absent of any programme effect, recidivism among 16- to 18-year-olds would have increased by the same amount as among 19-year-olds¹⁶³.

The outcomes related to the payment mechanism concern the rate of recidivism among the target group in the intervention. For this outcome, the specific indicator identified is defined as follows: percentage change in recidivism bed days (change in the average number of days these adolescents spent in jail) following an arrest on a new charge in the 12-month period following their initial release.

Other evaluations

The desk analysis did not find evidence on any other measurement process.

Barriers to and enablers of the measurement process

The main enablers and strengths of the measurement process that emerged from the desk research concern the **good level of internal validity of the propensity-score matching methodology**, which allowed historical matches to be identified, in order to account for individual-level changes. Another strength is that the evaluation design entailed the **use of a control group** of 19-year-olds incarcerated at the same time as the SIB was implemented, to account for system-level changes.

We did not identify any specific barriers to or weaknesses in the measurement process. However, it is worth pointing out that factors relating to the operation of New York City jails precluded the possibility of a randomised experimental design. During the evaluation’s pilot phase, between February and June 2012, Vera conducted a test of a possible measurement methodology using a randomised control trial by randomly assigning youth to a ‘treatment’ housing unit (who received ABLE services) or a ‘control’ housing unit (who did not receive ABLE). However, since adolescents were moved frequently between housing units during

¹⁶² Parsons et al. (2016), 12.

¹⁶³ Rudd et al. (2013), 35.

this period due to security and space issues, Vera researchers concluded that an experimental design could not be implemented effectively at Rikers. A preliminary analysis of DOC data collected during the pilot phase revealed that approximately 40% of the adolescents being followed were relocated from their randomly-assigned housing unit. This led to a high degree of contamination between the proposed treatment and control groups, severely compromising the ability to assess the programme's impact.

Table 21. Summary of the measurement process in the ABLÉ programme

Measurement methodology	
Methods related to payment mechanisms	Outcomes evaluation: propensity score matching
Other evaluations	n/a
Methods using experimental or quasi-experimental design	Yes
Control group	Yes
Causality of impact	Causality of impact was assessed using a quasi-experimental design. However, the evaluation results found a difference (increase) in recidivism that was not statistically significant, and is therefore equivalent to a zero programme effect.
Enabling factors and strengths of the evaluation process	Good level of internal validity of propensity-score matching methodology
Barriers to and weaknesses of the evaluation process	One single metric is too restrictive to represent the actual intervention's impact.
Governance of measurement	
Independence of evaluator	Yes
Evaluator	Vera Institute of Justice

1.6.4. Effectiveness

Achievement of outcomes

Payments were not issued by the commissioner to the investors because the **target outcome was not achieved**. In this case, an approach with a strong evidence base for successfully reducing recidivism in other settings did not work for the young people at Rikers.¹⁶⁴

In terms of outputs, in the first year, 1,255 young people participated in at least one ABLÉ session between 1 January 2013 and 31 March 2015. The programme was delivered to almost all of the target population, but the intensity of MRT was lower than planned. While 40% of participants attended the number of sessions associated with positive outcomes in

¹⁶⁴ Anderson, J. & Phillips, A. (2016). What we learned from the nation's first social impact bond. The Huffington Post.

earlier studies, only 11% completed all 12 stages of the programme¹⁶⁵. The expectation at the beginning of the SIB was that 30% of participants would complete the programme.

Vera's analysis showed a general upward trend in RBDs among both the study and comparison cohorts, indicating that changes in recidivism were the result of factors other than the ABLE programme.¹⁶⁶ More specifically:

- RBDs among 16- to 18-year-olds (ABLE cohort) increased by 13.4% between 2006-2010 and 2013;
- RBDs among 19-year-olds (system-level factors) increased by 8.0% between 2006-2010 and 2013.

The evaluation for the comparable TF scheme is not available. However, it must be noted that the set of interventions provided under the RESTART goes well beyond MRT alone, also providing employment training, counselling, case management, etc. In the comparable scheme, an average of 53% of participants successfully completed the programme.

One of the possible explanations for the lower percentage of completions in the ABLE SIB is that the number of sessions participants received was directly related to the **instability of the jail and the inmates' indeterminate length of stay**. Rikers experiences security-related events such as lockdowns and alarms on a daily basis, which create extensive delays and cancellations for scheduled activities such as school classes, MRT sessions and recreation. When fights occur, inmates are often moved to different housing areas, which further disrupts the continuity of MRT classes. Because the length of stay for young people awaiting trial is unpredictable, ABLE offered MRT at several locations in the community for adolescents who left Rikers without completing the programme. However, few of those young people attended, as adolescents leaving Rikers face a host of competing obligations, including school, court-mandated activities, and sometimes jobs.

Finally, adolescent inmates experienced an enormous amount of stress during their time at Rikers, limiting their ability to internalise and practise a promising therapeutic programme. The well-documented culture of violence and intimidation at Rikers, exacerbated by the strong presence of gangs, means that adolescents sometimes feel compelled to align themselves with other inmates who engage in negative behaviours. This perceived compulsion conflicts with the main goals of MRT, which include promoting personal responsibility and sound decision making.¹⁶⁷

Benefits and drawbacks of SOC compared with TF

The main **benefits** of the ABLE SIB scheme noted by the stakeholders were as follows:

- **A collective focus on data and accountability**, compared with similar programmes that were funded on a cost-reimbursement basis.¹⁶⁸ Investors brought an important level of accountability to the table and involved others into this conversation. Moreover, the embeddedness of the data collection and measurement infrastructure into the operations was considered to have helped to improve government efficiency, because inefficient programmes often continue due to a lack of outcomes tracking.¹⁶⁹

¹⁶⁵ Parsons et al. (2016), p. 14.

¹⁶⁶ Parsons et al. (2016), p. 20.

¹⁶⁷ Berlin (2016), p. 13

¹⁶⁸ Interview with the investor.

¹⁶⁹ Interview with the investor.

- **The clear definition of the outcome** allowed the offering of capital from impact investors to be aligned with the needs and priorities of the local authorities and thus the community.¹⁷⁰
- **The search for strong evidence** (quasi-experimental design to emphasise impact instead of just outcomes) **to assess success**.

Also, the SIB provided a way to **unlock funding** for the specific intervention. MDRC, the intermediary, worked with the City of New York to structure a contract wherein the Department of Correction agreed to pay for the programme. The SIB was seen as a strategy to **attract private investment for preventive services** in times of constrained government spending, which limit the ability to experiment with cognitive behavioural therapy for adolescents.¹⁷¹

Indeed, the Department of Correction struggled to effectively manage daily life in jail, and the city's Office of Management and Budget had little if any discretionary funding for innovation and experimentation. Funding a programme that might reduce the likelihood of future incarceration would have been a stretch for an agency consumed by day-to-day crises.

Furthermore, the city did not know whether cognitive behavioural therapy could be delivered effectively in jail, nor whether it would actually reduce recidivism among the population of adolescents housed there (even though there was good evidence demonstrating its effects elsewhere). The SIB provided the 'risk capital' necessary to experiment with the programme and collect evidence.¹⁷²

In terms of **drawbacks**, a very complex operating environment such as Rikers Island was not considered appropriate for the implementation of a SIB because the service providers did not have **full control over the intervention model and the recruitment process for participants**.¹⁷³ Furthermore, the process was more **difficult to manage** than a traditional programme.¹⁷⁴

Scalability and replicability

The programme was discontinued in August 2015 at the end of its third year, after the results of the evaluation proved that it was not effective at reducing recidivism.¹⁷⁵

1.6.5. Efficiency

Intervention costs

Table 22 below compares the intervention costs of the ABLE SIB with those of the RESTART programme. Please note that we do not control for the participants' unique characteristics in the two programmes or other factors (for example, location, service provider, etc.) which probably better explain these differences than the funding model alone (SIB vs TF). We include two different columns on the ABLE SIB: one outlining the costs incurred by the investors, while the other specifies the costs to the taxpayer. Although the

¹⁷⁰ Ibid.

¹⁷¹ Berlin (2016), p. 8

¹⁷² Berlin (2016), p. 8

¹⁷³ Berlin (2016), p. 8

¹⁷⁴ Misner, K. (2013). An interview with Kristin Misner of the NYC Mayor's Office about social impact bonds and the city's ABLE project. Think Justice Blog.

¹⁷⁵ Interview with the investor.

investors spent a total of over USD 2.4 million in the first year of the programme, the intervention cost nothing to the taxpayer because the targeted outcomes were not achieved, the programme was shut down and investors were not repaid. Therefore, **while the costs per admission and per participant who completed each programme favour the RESTART programme, perhaps the greatest advantage of the SIB is that it was possible to easily shut it down when it failed to deliver value for money.** This point was reiterated by the investor, which claimed that government programmes often continue regardless of their effectiveness or efficiency (unless the commissioning body experiences a budget cut), because there is no tracking of outcomes.

Table 22. ABLE SIB costs to taxpayers and to investors, compared with RESTART costs (USD)

	ABLE SIB 2013 Cost to investor	ABLE SIB 2013 Cost to taxpayers	RESTART 2017-2018
Admissions ¹⁷⁶	1,470 ¹⁷⁷	1,470	688 ¹⁷⁸
Programme completions	162 ¹⁷⁹	162	363 ¹⁸⁰
Total cost	2,400,000 ¹⁸¹	0	518,134 ¹⁸²
Cost per admission	1,633 ¹⁸³	0	753 ¹⁸⁴
Cost per programme completion	14,814 ¹⁸⁵	0	1,427 ¹⁸⁶

Source: Derived by PPMI from the available information.

Note: the costs have not been adjusted for inflation. This adjustment would further increase the ABLE SIB's costs in relation to RESTART.

Operational costs

The operational costs incurred by the commissioners of the ABLE SIB and the RESTART programme are not available. While information about the set-up and implementation of the latter programme is particularly limited, the **set-up** of the ABLE SIB was mostly carried out by the intermediary, MDRC:

MDRC worked with the City of New York to structure a contract wherein the Department of Correction agreed to pay for the programme if it achieved clear and measurable goals

¹⁷⁶ Youth in the study cohort; 1,255 youth in the study cohort who attended at least one ABLE session.

¹⁷⁷ Parsons et al. (2016), p. 14.

¹⁷⁸ Lois, R. L. and Schlegel, K. G. (2019). Audit report Project M.O.R.E., Inc. Office of the Dutchess County Comptroller. Retrieved from: <https://www.dutchessny.gov/Audit-Report-Project-MORE-Inc.pdf>. Outcomes were added up for 2017 and 2018.

¹⁷⁹ Derived by PPMI from the fact that 11% of participants successfully completed the programme (Parsons et al. (2016), p. 17).

¹⁸⁰ Lois and Schlegel (2019), 12. Outcomes were added up for 2017 and 2018.

¹⁸¹ Derived by PPMI from the fact that total investment in the programme amounted to USD 9.6 million, paid out quarterly over four years (over 16 quarters in total). Given that the information on outcomes is only available for 2013, the total cost is therefore equivalent to $9.6 \text{ million} / 16 \times 4 = 2.4 \text{ million}$.

¹⁸² Lois and Schlegel (2019), p. 5. Costs were added up for 2017 and 2018.

¹⁸³ Derived by PPMI from the total cost and the number of admissions.

¹⁸⁴ Derived by PPMI from the total cost and the number of admissions.

¹⁸⁵ Derived by PPMI from the total cost and the number of programme completions.

¹⁸⁶ Derived by PPMI from the total cost and the number of programme completions.

relating to recidivism. MDRC used the DOC's promise to pay to secure a loan from Goldman Sachs, guaranteed by a grant from Bloomberg Philanthropies.

In February 2012, MDRC issued a selective Request for Quotation, open to five of the most respected community-based organisations in the city that worked with young people involved with the justice system. MDRC selected the Osborne Association (Osborne) in collaboration with Friends of Island Academy (Friends). MDRC began contract negotiations with the Osborne Association in April 2012, and Friends entered into a subcontract with Osborne shortly thereafter. Almost immediately following the contracting process, MDRC worked with Osborne and Friends to hire and train the staff they needed. MDRC arranged for Correctional Counseling, the MRT developer, to schedule a series of MRT training sessions for new staff members as they were hired.

Before going full-scale, the Adolescent Behavioral Learning Experience (ABLE) project was **pilot-tested** to allow the project partners to identify and address any start-up challenges together. In January 2012, Correctional Counseling trained teachers from the Rikers school in MRT. From February until the end of June, these teachers provided MRT as part of the school day to approximately half of the adolescents in Rikers. In July, the non-profit service providers took over service delivery.¹⁸⁷

MDRC's intermediary work was funded through a separate grant from Bloomberg Philanthropies,¹⁸⁸ but the exact costs were not disclosed to the research team.

1.6.6. Design of the scheme

Design features

The ABLE SIB included a number of effective and efficient design features, including:

- **explicit agreement** between multiple parties concerning the time periods, benchmarks, and standards of evidence required to trigger payments;
- insightful and thorough **calculation of cost savings**, which subsequently informed payments;
- **a pilot-test** of the programme which allowed for some operational changes;
- provision of a **loan guarantee** from a philanthropic institution, which reduced the investor's risk;
- the **search for strong evidence** (quasi-experimental design to emphasise impact, instead of just outcomes) **to assess success**;
- **selection of the intervention on the basis of a strong base of evidence**;
- presence of an **intermediary** which helped to set up and manage a relatively complex contract;
- **structuring of the investment through a loan** (such as the one with MDRC) **with an outcomes-based forgiveness clause attached**. Accordingly, the investor forgave the loan since the outcome was not achieved: this was easy to understand and easy to document from a legal perspective;
- **an early exit option**, based on programme performance, which was taken up.

Nevertheless, there were other features about which stakeholders disagreed:

¹⁸⁷ Rudd et al. (2013), p. 33

¹⁸⁸ Rudd et al. (2013), p. 37

- relying on a **single outcome metric** might help to increase cost-savings, but it probably fails to capture the multifaceted impact of the programme;
- repayments to investors were structured in such a way that investors received nothing if the re-admission rate decreased by less than 8.5% (although staggered payments would have been made had the programme exceeded the breakeven target of 10%, see Section 1.6.1). While this reduced the risk for the commissioner, the investor argued that such a **sharp drop-off point** makes it difficult to attract investors. If a government entity truly believes that small changes in effect size represent an absolute difference between success and failure, then it may be necessary for benevolent funders to step in to ‘smooth the curve’ for more traditional investors.¹⁸⁹

Impact on third-sector organisations and social enterprises

The two service providers involved in the programme – the Osborne Association and Friends of Island academy – were both non-profits. Even though the project transferred the risk to investors rather than the service providers, the scale of the project made it too daunting for either agency to attempt to bid on the solicitation independently. Nevertheless, together Osborne and Friends brought a well-balanced and capable service provider team to the project. Osborne had extensive experience in running cognitive behavioural therapy programs, and had worked in prisons for more than 30 years. Friends had worked with young people in New York City for two decades, much of it with young men and women in the public school that operates on Rikers, East River Academy.¹⁹⁰ While the service providers could not be contacted for this study, it appears that **the project might have given an opportunity for these non-profits to acquire funding** that would otherwise have been out of reach. Nevertheless, the **SIB’s failure may also have resulted in negative publicity**.

1.7. The Benevolent Society Social Benefit Bond (Australia)

1.7.1. Background

Resilient Families is a family preservation programme developed by The Benevolent Society (TBS) as one of Australia’s first social benefit bonds (SBBs) between 2013 and 2018. It provided intensive family support services to **families whose children were at risk of being placed into out-of-home care (OOHC) (or ‘foster care’)** by the Government of New South Wales (NSW).

The programme was based on TBS’s ‘Resilience Practice Framework’, and was informed by TBS’s extensive experience in working with at-risk families, as well as international evidence-based programmes. The Resilience Practice Framework’ focuses on family resilience and promotes 42 Evidence-Informed Practices (EIPs) within the following five outcomes domains: secure and stable relationships, increased safety, improved coping/self-regulation, increased efficacy, and increased empathy. This intervention model takes inspiration from Homebuilders, a model of support that was developed in the 1970s in the US. Homebuilders’ services target families within the child protection system who are at a point where OOHC is likely without a significant change in both parental behaviour and the safety of the environment for the children. This model brings a strengths-based and

¹⁸⁹ Rudd et al. (2013), p. 52

¹⁹⁰ Rudd et al. (2013), p. 32

holistic approach to a structured assessment of needs, problem behaviours, and other safety and wellbeing concerns.¹⁹¹

The idea of using a Social Benefit Bond came about in 2011, when the Treasury of the NSW Government wished to undertake an SBB trial in the areas of out-of-home care and criminal recidivism. In 2009, the NSW Government had introduced a five-year action plan, *Keep Them Safe*,¹⁹² which aimed to reshape the way family and community services were delivered in NSW and to improve the safety and wellbeing of children and young people. The funds provided under the SBB were intended to expand social investment in prevention and early intervention approaches that might otherwise not have received sufficient funding.

In NSW, the number of children in OOHC had risen from 12,700 in 2007 to 18,000 in 2012, equating to 1.13% of all children in NSW. Less than one in ten of the highest-risk families were offered intensive support services. As a result, in excess of 3,000 children in NSW per annum were removed from their families and placed in OOHC.¹⁹³ While the purpose of the statutory care system was to prevent or minimise the impact that neglect and abuse have on children's development, it was costly to provide, and there was evidence that children and young people in OOHC generally had higher rates of physical, developmental and emotional problems, and lower rates of education than others. The Benevolent Society SBB **aimed** to allow up to 400 additional families to benefit from intensive family support services.¹⁹⁴

The main **reasons** why a social benefit bond (SBB) was chosen were:

- to demonstrate that the NSW Government, community and investors could all benefit from SBB initiatives;
- to mobilise new sources of mainstream capital to tackle social problems, expanding the level of upfront investment available for prevention and early intervention activities, and thus freeing up government funds to be used in other areas;
- to establish an investment model with potential for scaling investment into other social impact services;¹⁹⁵
- to help reduce the demand for government expenditure on acute crisis services and tertiary, curative interventions;¹⁹⁶
- to provide an opportunity to try out new approaches to working between the NSW Department of Family and Community Services (FACS) and the non-government sector: the shift away from funding tied to defined service specifications was intended to encourage the responsiveness and adaptation of the service as lessons emerged.¹⁹⁷

The Resilient Family service was provided in central and south-west Sydney, particularly in the areas of Rosebery, Liverpool and Campbelltown, which covered two FACS regions. The families identified for participation in the scheme had the following characteristics:

- at least one child younger than six years old living at home; and

¹⁹¹ ARTD Consultants (2016). Evaluation of the Resilient Families service – Interim report. NSW Treasury, 4.

¹⁹² ARTD Consultants (2014b). Evaluation of the Resilient Families Service – Preliminary report. NSW Treasury, 1.

¹⁹³ Benevolent Society (2013). Information memorandum, 4.

¹⁹⁴ Benevolent Society (2016). The Benevolent Society Social Benefit Bond – Investor report.

¹⁹⁵ Benevolent Society (2013), 4.

¹⁹⁶ ARTD Consultants (2014a). Evaluation of The Benevolent Society's resilient families service – Evaluation plan, 2.

¹⁹⁷ ARTD Consultants (2014b), 20.

- a FACS assessment of all the known dangers, current protective abilities, safety interventions and any other information available that indicated the child was at risk of serious harm but ‘safe with plan’.¹⁹⁸

The Resilient Families service provided for an initial intensive intervention of six to 12 weeks, designed to address crises and build client trust. During this period, TBS’s Resilient Families service staff comprehensively assessed the family’s situation, and worked with the family to identify both immediate and long-term goals.¹⁹⁹

From the start of the contact, TBS’s Resilient Families service staff aimed to stabilise the family environment, address housing and debt problems (including through the use of discretionary funding), to secure income support, and provide safety planning where domestic violence was an issue. Once the crisis period had passed, the Resilient Families service targeted improvements in the family’s functioning and support networks and arranged for children to receive any specialist support they needed. The Resilient Families service then provided for up to nine months of additional support in order to embed sustained changes in behaviour, build family capacity to cope with future challenges, and to connect the family with the wider network of community services. The final component of the Resilient Families service involved a schedule of check-ins after the family exited the programme, offering the family an opportunity to renew contact if problems re-emerge, and informing the family that TBS will contact them again if advised by Department of FACS about renewed child protection concerns.²⁰⁰

In total, the programme cost roughly AUD 10 million. The **return on the investment** depended on the degree of improvement in social outcomes and the type of bond. Two types of bonds were used:

- Class P Bonds – so-called senior, performance-based secured bonds; or
- Class E Bonds – so-called subordinated, principal ‘at risk’, performance-based secured bonds.

Each type of bond was associated with a different level of return, as shown in the table below.

Table 23. Returns on Class P and Class E bonds

Performance Percentage	Class P Bond	Class E Bond
Fail (< 5%)	0%	0%
Baseline (≥5% <15%)	5%	8%
Good 1 (≥15% <20%)	6%	10.5%
Good 2 (≥20% <25%)	7%	15%
Good 3 (≥25% <35%)	8%	20%
Good 4 (≥35% <40%)	9%	25%
Out-Performance (≥40%)	10%	30%

Source: The Benevolent Society Social Benefit Bond –Investor Report for the nine months ending 30 June 2014, p. 3.

¹⁹⁸ This indicates that there are one or more dangers present for the child concerned, and that without effective preventive services, the planned arrangement for the child/young person will be out-of-home care. The child is able to remain in the home as long as the safety interventions outlined in their Plan mitigate the identified danger(s). ARTD Consultants (2014a).

¹⁹⁹ ARTD Consultants (2017). Evaluation of the Resilient Families service (social benefit bond pilot) – Evaluation stage 2 progress report. NSW Treasury, 12.

²⁰⁰ Benevolent Society (2013), 4.

Investors were rewarded on the basis of a performance percentage, which was calculated at the measurement date and certified by an independent third party (the ‘independent certifier’). The measurement date was set four years and nine months after the date of the first referral. The data on which the independent certifier based its calculation was provided by the Department of FACS. Performance was based on the outcomes for the intervention group versus a control group over the period of the trial. The key performance measures for The Benevolent Society SBB referred to the numbers of: entries into OOHc per child (‘entries’); helpline reports (‘reports’); and SARAs (Safety and Risk Assessments) (‘assessments’).²⁰¹

Concerning the **payment schedule**, The Benevolent Society, within eight weeks of the measurement date, provided the Department of FACS with a valid invoice of the total amount payable to the organisation and a copy of the independent certifier’s report. The Department paid this amount payable within four weeks of receiving the invoice. Payments were made to investors at the end of the trial period, looking in particular at the performance percentage at the measurement date of 3 July 2018. The performance percentage also took into account the improvement percentage (weighting 94%) for each annual cohort. The improvement percentage considered the results of each annual cohort during the four years of the project’s implementation. Payment was issued according to the contract.

The programme achieved a performance percentage of 16% at the measurement date of 3 July 2018, triggering returns of 6% to Protected Class investors and 10.5% to Equity Class investors. The Resilient Families programme achieved an 86% preservation rate for the families that were referred to it. Out of the families referred to the Resilient Families scheme, 32% fewer children entered out-of-home care than children from the matched control group of families (i.e. those who received a ‘business as usual’ service). This shows that Resilient Families was a highly effective model in preventing entries into out-of-home-care.²⁰²

Aside from achieving its payment-related outcomes, the overall evaluation of the SBB was positive, showing the scheme’s added value in relation to long-term outcomes. Specifically, an average of 0.2 more children showed an increase in secure and stable relationship; 0.7 more children showed an increase in safety; 0.8 more children showed an increase in coping/self-regulation; and 0.3 more children showed an increase in efficacy.²⁰³ Children involved in the programme also performed better than the control group in terms of their number of entries into out-of-home care (14% of index children²⁰⁴ compared with 19% in the control group), the number of helpline reports (reports were filed for 65.3% of index children compared with 69.3% in the control group), and the number of safety and risk assessments (38.6% vs 42.9%).²⁰⁵

Table 24. Summary of the Benevolent Society Social Benefit Bond

Personally targeted social service	Caring obligation
Target population	390 families with resident children aged five years or younger (including unborn children) at risk of significant harm over four years
SOC scheme type	Managed SIB
Payment model	Frequency
Cost	AUD 10 million (Class P AUD 7.5 million and Class E AUD 2.5 million) Class P Bonds: Senior, performance-based secured bonds; and Class E Bonds: Subordinated, principal ‘at risk’, performance-based secured bonds.

²⁰¹ Benevolent Society (2013), 6.

²⁰² Benevolent Society (2018). Resilient Families – Impact report 2013-2018, 4.

²⁰³ NSW office of Social Impact Investment (2020). Evaluation of the Resilient Families service – final report, p. 77

²⁰⁴ Youngest child in each participating family.

²⁰⁵ NSW office of Social Impact Investment (2020), p. 79.

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

Commissioner(s)	NSW Government (NSW Treasury)
Social service provider(s)	The Benevolent Society Social Benefit Trust No.1 in partnership with the Department of FACS
Investor(s)	High net worth individuals, self-managed super funds, a trust, foundations and institutions (the identities of investors were not available)
Intermediary	Westpac Banking Corporation, Commonwealth Bank of Australia, The Benevolent Society
Evaluator(s)	The consultancy firm, ARTD Pty. Limited, conducted an evaluation of the programme. The results of the SBB were independently certified by Deloitte Touche Tohmatsu.

1.7.2. Equivalent TF scheme

The comparable TF scheme chosen is the Intensive Family Preservation (IFP) service. It is provided by the NSW Government in collaboration with the Department of FACS. It is based on the understanding that it is in the child's best interests to remain in the care of their family, wherever this is a safe option. It focuses on improving children's safety, the permanency of their placement, and their wellbeing. The service targets families with children aged from birth to 18 years old, who meet the eligibility criteria. The service model is informed by Homebuilders[®]. It provides a three-month period of intensive support, including 24-hour access to a caseworker. This is followed by a less intensive level of individually tailored and multi-faceted services that include advice and referral, assessment and case planning, counselling, family-focused casework, home visiting, parent support groups, and skills-focused groups.²⁰⁶

The intervention is comparable because it is provided by the same commissioner in the same area (Sydney), using the same intervention model, Homebuilders[®], and addressing the same outcomes. However, the target population differs slightly (TBS included only families with younger children).

Table 25. Comparability of IFP with TBS

	Same as SOC	Comments
Personally targeted social service	Yes	Caring obligation
Target population	No	SOC: families with resident children aged five years or younger (including unborn children) at risk of significant harm TF: Families with children aged 0-17 years
Number of participants	Yes	SOC: annualised target: 80 families per year ²⁰⁷ TF: annualised target: 98 families per year ²⁰⁸
Location	Yes	New South Wales
Cost	No	SOC: annualised budget AUD 2 million ²⁰⁹ TF: annualised budget AUD 3,980,443 ²¹⁰

²⁰⁶ Policy, Programs and Strategy Statutory Child Protection (2014). Intensive Family Preservation Program Guidelines, 5.

²⁰⁷ ARTD Consultants (2014), 19.

²⁰⁸ ARTD Consultants (2014), 19.

²⁰⁹ ARTD Consultants (2014), 19.

²¹⁰ ARTD Consultants (2014), 19.

	Same as SOC	Comments
Commissioner(s)	Yes/No	SOC: NSW Treasury TF: NSW Department of Family and Community Services (FACS)
Social service provider(s)	No	SOC: The Benevolent Society Social Benefit Trust TF: Non-governmental service providers

1.7.3. Outcomes measurement

The scheme's evaluation design aims to assess the implementation, effectiveness and cost-effectiveness of the service. It involves outcomes, process, and economic evaluations. The outcomes, process and cost evaluations were carried out by ARTD Consultants. The specific section of the outcomes evaluation relating to the bond payment was verified and certified by Deloitte Touche Tohmatsu.

Measurement of outcomes relating to payment mechanisms

The performance of the bond, and therefore the level of returns to investors, was measured by comparing the children who receive the Resilient Families service against a control group of children who did not receive the TBS service. Families in the control group were not aware that they had been selected and the FACS services they received did not change. The following three measures were monitored for both groups:

- Number of entries into out-of-home care;
- Number of Safety And Risk Assessments (SARAs) commenced;
- Number of reports to the Child Protection Helpline.²¹¹

The intervention group involved families that received the service through the three TBS sites (Rosebery, Campbelltown and Liverpool) during the period October 2013 to the end of June 2017, and consented to being involved in the evaluation.

The intervention group comprised the youngest children in each participating family (the index child). Each index child was matched with a child from a non-participating family according to specific criteria, in order to form the study control group. The matching criteria were the following: child's age, family size, indigenous status, out-of-home care/SARA history of the mother. The children in the intervention group and the matched children in the control group were allocated to an annual cohort. There were four annual cohorts, one for each year of referrals. No new referrals were accepted after year four²¹².

The main long-term outcome of the SBB was the safety of the family, which meant reduced contact with the child protection system. This was measured by taking into account three main indicators²¹³:

1. Fewer helpline reports ('Reports') – calls made by people in the community or service providers to the helpline to report a concern about the children's safety.
2. Fewer SARAs – investigations conducted by the Department of FACS.

²¹¹ Benevolent Society (2013), p. 6.

²¹² Benevolent Society (2013), 6.

²¹³ Benevolent Society (2013).

3. Fewer OOHC (out-of-home care) placements ('entries') – statutory removals of children.

These three measures covered a wide spectrum of child safety indicators, and together determined the improvement percentage. The annual improvement percentage was averaged over the term of the trial, using the annual improvement percentages of each succeeding year's annual cohorts of families.

The improvement percentage for an annual cohort was calculated by the independent certifier by applying the following formula:

Improvement percentage = (66% x entries) + (17% x reports) + (17% x assessments)

In the formula, entries, reports and assessments were each calculated as the aggregate frequency of the measure for the control group minus the aggregate frequency of the measure for the intervention group, expressed as a percentage of the aggregate frequency of the measure for the control group (subject, among other things, to conditions for annual minimum numbers of entries to OOHC, helpline reports and SARAs on a state-wide basis, as well as exclusions, measurement periods and period adjustments)²¹⁴.

Ultimately, investor payments depended on the performance percentage, which was based on the following:

- Improvement percentage (weighting 94%);
- Unmatched children percentage – where children referred by FACS cannot be matched with a comparable control child, (weighting 1%); and
- Guaranteed referrals shortfall percentage – where FACS is unable to fill vacancies notified by The Benevolent Society within the agreed period of time, up to a guaranteed minimum (weighting 5%).

Other evaluations

Outcomes evaluation – Resilience Outcome Tool

The main tool for the outcomes evaluation was the TBS Resilience Outcomes Tool, which was used by SCFWs (Senior Child and Family Workers – TBS primary carers) to inform individual case planning and assessment²¹⁵. The tool comprised a range of validated scales or sub-scales:

- Strengths and difficulties questionnaire (SDQ): this assesses a child's social-emotional wellbeing and emerging behavioural problems. It consists of 25 items completed by parents/carers. Items fall under five scales: an emotional symptoms scale, conduct problems, a hyperactivity scale, a peer problems scale, and a pro-social scale. These scales measure outcomes related to:
 - the child's cognitive development;
 - the child's social competence;
 - the child's emotional regulation.

²¹⁴ Benevolent Society (2014). The Benevolent Society Social Benefit Bond –Investor Report for the nine months ended 30 June 2014.

²¹⁵ ARTD Consultants (2014a).

- Protective factors survey (PFS), which considers the following areas of problems: knowledge of parenting; nurturing and attachment; family functioning; social support; and concrete support. It is a pre–post tool, designed when working with caregivers who are receiving child abuse prevention services. The survey results are designed to provide a snapshot of families, changes in families’ protective factors, and areas in which workers can focus on increasing family protective factors.
- The Longitudinal Study of Australian Children, which includes data about parenting, family and relationships, community links, health and wellbeing. The study focuses on the following outcomes:
 - secure and stable relationship;
 - increasing safety;
 - increasing self-efficacy;
 - improving coping/self-regulation.
- Parenting Sense of Competence (PSOC) scale: this was created to assess perceived parental competence by measuring outcomes such as satisfaction with parenting and the perceived self-efficacy of parents. In this scale, satisfaction reflects the parental level of frustration, anxiety, and motivation, while efficacy reflects their level of competence, problem-solving ability and capabilities as a parent.
- Personal Wellbeing Index. This is a 10-item questionnaire intended to yield a global measure of distress based on questions about anxiety and depressive symptoms that a person has experienced in the most recent four-week period. Its target outcome is increasing safety.
- Home physical environment is a practitioner based observation of the home environment in order to better understand the outcome of increased safety.
- The K10 Scale is a standardised screening tool used for clinically significant depressive and anxiety disorders. It concerns improvements to coping/self-regulation outcomes.
- Family Resource Management relates to the financial circumstances of the family, in order to investigate the outcome of increased safety.

Process evaluation

The evaluation used a mixed-method design, drawing on primary and secondary data to understand the implementation of the Resilient Family service and its client characteristics²¹⁶. The evaluation considered the implementation of services in relation to the contextual factors of the implementation sites and the characteristics of the TBS structure under which the service was implemented.

The evaluation identified:

- Demographics of index children, their primary carers and other family characteristics such as family structure, housing.
- Service characteristics e.g. intensity, duration, EIPs, services to which families were referred.

²¹⁶ ARTD Consultants (2014a), p. 10-23.

- Effectiveness of processes for referral, joint working, and support for staff in delivering the Resilient Family (RF) programme.

To better understand the reasons behind the results achieved through RF, and to capture more qualitative results, **semi-structured exit interviews with primary carers** were conducted. These interviews covered:

- family context;
- perceptions about the service (when they were initially referred, and now that they have been involved for a while);
- relationship with the Senior Child and Family Worker;
- what aspects of the way the project worked were important to them;
- what they liked/did not like about the programme;
- what they found useful/not useful about the programme;
- suggestions for changes/improvements to the programme;
- impact (if any), increases in support networks, changes in feelings about accessing mainstream services/knowledge of which services to access and when, parenting skills, family functioning;
- service gaps—services that families needed, but which were not provided.

Another part of the process evaluation consisted of **group interviews with staff from each of the three TBS programme sites**. These interviews gathered structured information on implementation and how families reacted to/engaged with the project. These interviews covered:

- training and support;
- contextual and service system factors impacting the implementation;
- referral processes (from and to CSCs at the beginning and end of involvement with a family);
- families' engagement with the programme (characteristics of those who didn't engage vs those who did);
- working relationships with other relevant local services;
- early indications about how the programme was working for families;
- what aspects of the project families liked/found useful;
- suggestions for changes/improvements.

Finally, individual **interviews with key FACS** staff with knowledge or oversight of the TBS SBB pilot were conducted. These interviews covered:

- referral criteria and process;
- the approach to programme delivery;
- processes for joint working with TBS;
- perceived outcomes for clients;
- any lessons for future delivery;
- contextual factors affecting programme implementation (e.g. adverse events, high service demand, and lack of providers for particular services).

Cost evaluation

The cost analysis identified the costs of programme implementation and service delivery at the level of the programme, of the individual family unit. This analysis drew on two sources of data:

1. TBS data—RF costs, including caseworker salary information and an estimate of hourly costs for the team leader’s supervision.
2. Other data from the programme – comparison with the costs of similar programmes, taking into account the duration and intensity of service within these. Programmes funded by FACS (Brighter Futures, Intensive Family Preservation Services, Intensive Family-Based Services) were included where possible, as well as some from other jurisdictions. The consultants, ARTD, identified relevant programmes from other jurisdictions, and NSW Treasury approached relevant organisations to obtain cost and service data.

Barriers to and enablers of the measurement process

Strengths of the measurement process that emerged from the desk research relate to the development and use of a **mixed-method design** for the evaluation, which allowed for a holistic perspective on the programme, taking into consideration not only the outcomes of the programme but also its processes and costs. In addition, the presence of a comparison group can be considered to have strengthened the methodology.

A barrier recognised by both the commissioner and the provider concerns **data collection and management**. Indeed, the amount of primary and secondary data that needed to be collected required a huge effort from the stakeholders involved. The NSW Treasury was particularly interested in understanding the savings that the SBB could provide to the public budget, but other stakeholders that provided data were not so engaged with this objective, leading to difficulties with the availability of data.

One enabling factor mentioned by the stakeholders was the **relationships established** between the partners in the SBB – although the provider noted that the joint development phase had been quite tense. The SBB mechanism aimed to create a partnership model that was different from the relationships among stakeholders in other, more traditional programmes. The provider noted that “there was quite a cultural shift and it took a little bit of time to move over” since “everyone had their preferences and agenda”.

Table 26. Summary of the measurement process in the TBS SBB

Measurement methodology	
Methods relating to payment mechanisms	Outcomes performance percentage (Counterfactual analysis)
Other evaluations	Outcomes evaluation Process evaluation Cost evaluation
Methods using experimental or quasi-experimental design	Yes
Control group	Yes
Causality of impact	Causality of impact cannot be attributed based on the measurement methodology adopted.

Enabling factors and strengths of the evaluation process	Relationship among stakeholders
Barriers to and weaknesses of the evaluation process	Data management and collection
Governance of measurement	
Independence of evaluator	Yes
Evaluator	ARTD Consultants

1.7.4. Effectiveness

Achievement of outcomes

As mentioned in the background section, the programme achieved its targets, with a performance percentage of 16% at the measurement date of 3 July 2018.²¹⁷

A comparison with the IFP programme based on the same outcomes cannot be conducted, since the data are not available.

Benefits and drawbacks of SOC compared with TF

The stakeholders interviewed agreed that the evaluation process is a positive aspect of the scheme, compared with traditionally financed interventions. The evaluation process performed as part of the SBB provides investors with a strong evidence base and the government with locally relevant data for future social policy making.²¹⁸

The provider also noted that the scheme was beneficial for unlocking funding for the service, as well as opening up new relationships with banks and the NSW treasury.

The main drawback of the programme mentioned by stakeholders was that the SBB was situated outside of 'regular governmental programmes', and was not properly integrated. The commissioner noted that it was difficult to apply the lessons learned from this programme to other government interventions.

It is also worth mentioning that the programme initially faced **adverse public opinion** in the region. The Benevolent Society was seen as a private organisation profiting from children in need. Some of the staff in the commissioning body also had a negative opinion about the programme; however, according to the commissioner, they changed their minds when the positive results of the programme became clearer.

Scalability and replicability

The programme was extended for two years using a traditionally-financed model. There was no replication or scaling.

²¹⁷ Benevolent Society (2018).

²¹⁸ ARTD Consultants (2014b), p. 2.

1.7.5. Efficiency

Intervention costs

According to the preliminary evaluation of the Resilient Families service, its budgeted cost per family compared well with the Intensive Family Preservation programme. In the table below, we also include additional similar TF programmes.

Table 27. Comparison of Resilient Families funding with similar programs in NSW

Programme	Service level	Annualised budget	Annualised target number of families	Average funding per family
Resilient Families	Intensive	2,000,000	80	25,000
Intensive Family Preservation	Intensive	3,980,443	98	40,617
Intensive Family-Based Service	Intensive	3,200,000	88	36,364
Intensive Family Support	Intensive	6,113,027	170	35,959
Brighter Futures	Medium	58,300,000	3,124	18,662

Source: ARTD Consultants (2014), p. 19.

Given that true expenditure can diverge from budgeted costs, we also calculated the actual cost per family for the entire duration of the programme. Prior to the pay-out to investors, the SBB cost AUD 10 million: AUD 7.5 million was issued in Class P bonds whereas additional AUD 2.5 million was invested through Class E bonds. Since the programme achieved a 16% weighted average performance rate, Class P and Class E investors received a 6% and a 10.5% return on their investment, respectively. This brought the total cost of the SBB to AUD 10,712,500. The programme served 354 families between 2013 and 2018, so the cost per family is equivalent to AUD 30,261. This is slightly higher than the budgeted cost, but still compares favourably with the similar TF programmes. We have assumed that the budgeted and actual costs of the TF interventions are similar because no additional expenditures such as return on investment have to be paid out to investors.

Please note that a more recent cost analysis has been conducted that compares the Resilient Families service with other NSW programmes, but the results of this analysis are not available to the public.

Operational costs

While the costs of the Resilient Families intervention compare favourably with those of other services, interviewees argued that **the cost of operating the SBB was higher than that of traditional programmes**. Although the set up costs of the IFP service are not available for comparison, the setting-up of Resilient Families cost AUD 196,954 between June 2013 and September 2013.²¹⁹ A representative of one service provider claimed that the SBB was ten times more time-consuming than other contracts they had worked on. The pilot nature of the project, as well as the inclusion of a joint development phase, increased the costs of setting up the programme. The SBB generated media attention, so service providers also had to communicate with the public more frequently than they normally would. Interviewees also argued that the management and evaluation costs were substantial, particularly because the complex nature of the contract entailed legal fees, and because a control group was used to assess the programme's impact. The government agency responsible had to hire an additional staff member to oversee the extraction of the performance data.

1.7.6. Design of the scheme

Design features

The commissioner argued that although the inclusion of a **control group** was expensive, it was a worthwhile investment because it provided evidence that the programme works, which served as the justification to continue the programme and apply it elsewhere. Among other effective design features, the stakeholders mentioned the **annual review of the contract**, after which certain changes could be introduced, depending on how the programme was progressing. Another expert noted that it is important that participants are engaged on a **voluntary** basis, especially if a programme lasts for a prolonged amount of time.

The service provider nevertheless noted that **a start-up phase lasting more than six months** would have been helpful to better prepare for service delivery. Another interviewee also argued that **a longer contract** duration would have provided greater stability for the service provider. It would have been more difficult for the commissioner to budget for such a contract, however, and it would have entailed less flexibility because the government would have been locked into using the same service provider.

Impact on third-sector organisations and social enterprises

In the service provider's opinion, the SBB provided a number of advantages, including:

- The collection of rich data, which helped it to understand what works and what doesn't.
- Establishing a good reputation with banks and other private investors that could result in future funding opportunities.
- Establishing a relationship with the government, which could also help in expanding its business profile.
- Enhancing its reputation as an effective service provider.
- Serving families that would otherwise not have taken part in such a programme.

²¹⁹ ARTD Consultants (2014b), p.18.

The main disadvantages were that the programme was time-consuming and involved an expensive data collection system.

The commissioner mentioned that it was decided not to involve **small or young service providers**, because it would have been too hard for them to set up and conduct the programme. According to the commissioner, small or young providers could not assume all of the risks involved in this type of contract.

1.8. Mental Health and Employment Partnership (the UK)

1.8.1. Background

The Mental Health and Employment Partnership social impact bond (MHEP SIB) was set up in 2015 to drive a large-scale expansion of high-quality **employment programmes for people with mental health issues and other groups with health conditions and disabilities**. Specifically, MHEP SIB aimed to scale up Individual Placement and Support (IPS) – a service that had a track record of delivering outstanding job outcomes²²⁰. This service was designed to address the employment support needs of secondary health service users with severe mental health conditions. The IPS service is based on a ‘place first, then train’ model, which is thought to be more effective than traditional approaches such as vocational training and sheltered work. In this model, the person is immediately placed in a job position so they can experience the benefits as well as the challenges of these situations, and is then trained to successfully maintain this placement.

During the MHEP SIB, each service recipient worked with an employment advisor from same the team of mental health professionals who usually interacted with the service recipient. As the first step, an employment specialist carried out an assessment of the service user to identify their work preferences and capabilities. The employment specialist then co-produced a ‘return to work’ plan with the service user, using input from the treatment team as well as clinical records, family and previous employers as required. Within one month of creating the plan, the specialist identified potential jobs and supported the service user to apply for vacancies. Once a job placement had been made, support was provided to both the new employer and the employee, to ensure sustainment, help with any adaptation, and to build up mutual confidence between the two parties. Therefore, while people with mental health issues or other health conditions were the direct beneficiaries of the SIB, employers can also be considered intermediate beneficiaries. The specialist subsequently also provided in-work support and training, as well as additional non-work support such as signposting debt and benefits advice.²²¹

MHEP SIB combined outcomes-based **funding** from national (Commissioning Better Outcomes Fund [CBO], Social Outcomes Fund [SOF]) and local sources (from the counties where the intervention took place). It was also supported with capital from a socially minded investment company, Big Issue Invest. In April 2016, services were launched in Tower Hamlets, Haringey and Staffordshire. Our analysis focuses on these three cases because the schemes in other ones are still ongoing. Services were later expanded to cover Barnet, Enfield and Camden, and more recently to eight boroughs in West London, but these areas are not covered by our analysis. The initial services launched in 2016 lasted until 2019 (a period of 36 months).

²²⁰ Social Finance. (2019). Mental Health and Employment (MHEP). Retrieved from: <https://www.socialfinance.org.uk/projects/mental-health-and-employment-mhep>.

Health and Employment Partnership. (2015). Launching A Social Impact Bond In Mental Health And Employment.

²²¹ Bromage and Swersky (2018). Outcomes-Based Commissioning in Mental Health & Employment. Presented at the Innovation in Outcomes-Based Commissioning West Midlands Regional Conference.

The SIB was designed and set up by Social Finance using a special purpose vehicle (SPV), a limited company called Mental Health and Employment Partnership (MHEP)²²². Through MHEP, various commissioners were able to contract providers using individual contracts. For example, the contracts for Staffordshire and Tower Hamlets were different from each other, even though they delivered the same IPS service. MHEP provided ready-to-use and flexible contract templates that were used for different contractual arrangements according to each commissioner's needs. MHEP was established with an initial investment of GBP 400,000 from Big Issue Invest, which retains a 100% economic interest in the MHEP company.²²³

Although commissioners considered funding IPS through a fee-for-service type of contract, a SIB combined with fee-for-service funding was eventually the preferred choice. The main **reasons** for this included:

- flexibility in structuring the partnership between local and central government commissioners²²⁴;
- opportunity to combine large-scale national funding with local funding for mental health services²²⁵;
- extending the funding and increasing the scale of an existing service that had been proven to work²²⁶;
- the interest of the Department for Work and Pensions (DWP) in reviewing potential interventions for the 'hardest to help' referrals to the Work Programme.

In the MHEP SIB, local commissioners paid for the successful engagement of users (namely, after the completion of a vocational profile and an Action Plan), while CBO and SOF funding was used to pay for job outcomes. Payments were made six weeks after referral and six months after service recipients took up job placements. Specific per-user payments included²²⁷:

- successful engagement of users: GBP 790-1,000;
- job entry outcomes (<16 hours/week): GBP 700;
- job entry outcomes (>16 hours/week): GBP 1,350;
- job sustainment outcomes²²⁸ (<16 hours/week): GBP 1,400;
- job sustainment outcomes²²⁹ (>16 hours/week): GBP 1,650.

While different contracting models existed at local level, all contracts were characterised by both block funding for participant engagement and a per-annum amount linked to job outcomes, with different shares of each type of funding included in the contracts for each of

²²² Carè and Ferraro (2019). Funding Innovative Healthcare Programs Through Social Impact Bonds: Issues and Challenges. *China-USA Business Review*, 18(1), 1-15. doi: 10.17265/1537-1514/2019.01.001

²²³ Big Issue Invest (2019a). Fund Management Social Enterprise Investment Fund II L.P. Annual Report. Retrieved from: <https://images.bigissueinvest.com/2019/01/SEIF-II-2018-TO-PRINT-V2.pdf>

²²⁴ Government Outcomes Lab (2019a). Mental Health and Employment Partnership (Staffordshire, Haringey & Tower Hamlets). Retrieved from: <https://golab.bsg.ox.ac.uk/knowledge-bank/case-studies/mhep/>

²²⁵ Carè and De Lisa, 2019. Social Impact Bonds for a Sustainable Welfare State: The Role of Enabling Factors. *Sustainability*. 11(10), 2884. doi:10.3390/su11102884

²²⁶ CBO (2016). Mental Health and Employment Partnership (MHEP) Social Impact Bond – produced as part of the CBO Fund Evaluation.

²²⁷ Government Outcomes Lab (2019c). MHEP - Data Template. Retrieved from: <https://golab.bsg.ox.ac.uk/knowledge-bank/case-studies/mhep/>

²²⁸ This outcome is achieved when an IPS client sustains paid, competitive employment for at least 13 weeks during which they work for an average of less than 16 hours per week.

²²⁹ This outcome is achieved when an IPS client sustains paid, competitive employment for at least 13 weeks during which they work for an average of more than 16 hours per week.

the three areas. In each area, these payments were allocated differently between local government and the MHEP, as we outline below.

In Staffordshire, Staffordshire County Council paid for job entries and sustained jobs while MHEP covered the attachment fees. However, the service provider had only one contract and only invoiced the County Council, which, in turn, invoiced MHEP for its proportion of the payments. The original contract agreed by MHEP in negotiation with SOF blended block payment with a PbR model, which allowed for GBP 60,000 of block funding (provided by MHEP capital) and GBP 240,000 per annum linked to outcomes. This was changed to block funding of GBP 82,500 (provided by MHEP capital) and GBP 217,500 per annum on outcomes to adjust the funding balance and make the new staffing model possible.

In Tower Hamlets, both MHEP and the Clinical Commissioning Group paid providers directly for the outcomes. A memorandum of understanding between Tower Hamlets and MHEP set out this agreement. Technically, the provider had two contracts – one with MHEP and another with the borough’s Clinical Commissioning Group. In practice, the terms of these contracts were identical, so from the provider’s perspective they were effectively working under a single contract.

In Haringey, there was a commissioning agreement between Haringey council and MHEP. Formally, MHEP engaged the service provider, Twining, but in practice the contract was jointly managed by both MHEP and the borough. After some re-negotiation, half of the payments were made as block payments, while the other half were paid on the basis of outcomes. This balance was achieved by setting lower payment levels for each of the job start outcomes²³⁰.

Overall, the **targets** were²³¹:

- 550 users referred in one year;
- 2,800 service users referred in total;
- 2,624 service users actively engaged in total.

In addition, providers were given different targets:

- The Staffordshire Invitation to Tender specified the number of job outcomes expected.
- Under the Tower Hamlets contract, MHEP provided a ready-made model for the intervention’s outcome structure and the targets triggering payments; this was based on each customer engagement, together with fixed payment amounts according to separate outcomes.
- In Haringey, the outcome payments were all linked to the IPS standard performance reporting models rather than anything developed by or with Haringey.

Specific details about the targets are available only for Staffordshire²³² (see the table below).

Table 28. MHEP SIB targets in Staffordshire

	Target Outcomes <u>2016-2017</u>	Outcomes Achieved <u>2016-2017</u>
Successful engagement of users	450	559

²³⁰ CBO (2017). Midpoint in-depth review report produced as part of the CBO Fund Evaluation. Retrieved from: https://www.tnlcommunityfund.org.uk/media/research-documents/social-investment/comissioning_better_outcomes_in_depth_review_190320_122442.pdf?mtime=20190320122441

²³¹ Government Outcomes Lab (2019c).

²³² Government Outcomes Lab (2019c).

**STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE
PROVISION OF SOCIAL SERVICES AND INTERVENTIONS**

	<u>Target Outcomes 2016-2017</u>	<u>Outcomes Achieved 2016-2017</u>
Job entry outcomes (<16 hours/week)	80	NA
Job entry outcomes (>16 hours/week)	60	NA
Job started	NA	133
Job sustainment outcomes (<16 hours/week)	50	NA
Job sustainment outcomes (>16 hours/week)	40	NA
Job sustainment (six weeks)	NA	91
Job sustainment (six months)	NA	37

For a summary of stakeholders in the MHEP SIB, please see the table below.

Table 29. Summary of the MHEP SIB

Personally targeted social service	Labour market-related Disabilities
Target population	Adults registered with a general practitioner in Tower Hamlets, Haringey or Staffordshire, who had or were recovering from a mental health condition and were looking for work. Typically, these adults were diagnosed with psychosis, schizophrenia, bipolar disorder, severe depression or anxiety, and were in contact with statutory mental health services. Target service users referred: 2,800 Target service users actively engaged: 2,624
SOC scheme type	Intermediated SIB
Payment model	Binary
Cost	GBP 2.2 million outcomes-based funding raised from the Big Lottery Fund and Social Outcomes Fund GBP 400,000 social investment raised from Big Issue Invest GBP 2.9 million: maximum outcome payments
Commissioner(s)	Commissioning Better Outcomes Fund Social Outcomes Fund London Borough of Haringey Staffordshire County Council Tower Hamlets Clinical Commissioning Group
Social service provider(s)	Making Space (Staffordshire, charity) Twining (Haringey, charity) Working Well (Tower Hamlets, charity)
Investor(s)	Big Issue Invest, social impact equity-based fund
Intermediary	Social Finance UK Mental Health and Employment Partnership was designed and set up by Social Finance, as a social investment vehicle
Evaluator(s)	ATQ Consultants (private consulting firm)

	Ecorys UK (private international research and consultancy company)
	Behavioural Insights Team (private consulting firm in partnership with Cabinet Office)

1.8.2. Equivalent TF scheme

Camden Individual Placement Support (IPS) was a pilot employment support programmes designed to support people with mental health needs into employment, in line with the ambitions of the London Borough of Camden (LBC) to reduce unemployment and tackle health inequalities in the borough under its Camden Plan. The aim of the pilot was to assess the effect of specialist support for employment among Camden residents with severe and enduring mental health conditions (such as bipolar and schizophrenia). Camden IPS was co-commissioned by the Economic Development Team at LBC, the Integrated Commissioners at LBC, the Camden Clinical Commissioning Group, and Job Centre Plus. Camden Individual Placement Support (IPS) was delivered from January 2016 to May 2017 by Twining Enterprise, the same service provider engaged in the MHEP SIB in Haringey. Moreover, the evaluation of Camden Individual Placement Support (IPS) considered the same KPIs used in the three cases of the MHEP SIB analysed in the present report.²³³

In addition, this comparison appears suitable and interesting because the London Borough of Camden is one of the locations where the MHEP SIB, firstly implemented in Staffordshire, Haringey and Tower Hamlets, was later replicated.

Table 30. Comparability of MHEP with Camden Individual Placement Support (IPS)

	Same as SOC	Comments
Personally targeted social service	Yes	Disabilities Labour market-related
Target population	Yes	Unemployed and economically inactive clients 18+ who have a diagnosed severe and enduring mental health condition and are engaged with specialist mental health services.
Number of participants	No	SOC: 875 clients assessed per year in the SIB ²³⁴ TF: roughly 143 clients were referred in a year ²³⁵
Location	No	Different counties
Cost	No	SOC: The annualised cost of the SIB in Staffordshire and Tower Hamlets was GBP 900,000, and GBP 300,000 in Haringey ²³⁶ . TF: This compares with an annualised cost of GBP 150,000 in the Camden IPS (TF). ²³⁷
Commissioner(s)	Yes/No	All local authorities SOC: London Borough of Haringey

²³³ Begum, S. (2018). Transforming employment support for people with mental health conditions - An evaluation of two pilots in Camden.

²³⁴ CBO. (2017). In-Depth Review: Mental Health and Employment Partnership (MHEP), page 2.

²³⁵ Begum, S. (2018). Transforming employment support for people with mental health conditions - An evaluation of two pilots in Camden, p. 14.

²³⁶ ECORYS and ATQ. (2019). Commissioning Better Outcomes Fund Evaluation. 2nd Update Report Full Report.

²³⁷ Begum, S. (2018). Transforming employment support for people with mental health conditions - An evaluation of two pilots in Camden, p.7.

	Same as SOC	Comments
		Staffordshire County Council Tower Hamlets Clinical Commissioning Group TF: London Borough of Camden
Social service provider(s)	Yes	Twining Enterprise

1.8.3. Outcomes measurement

The MHEP SIB involved a quantitative evaluation based on secondary data collected by service providers to describe the performance of each service. A **descriptive study design** was chosen.

Parameters relating to job entry and sustainment have been used in employment support programmes in the UK since 1997, and are quite well established. They are also fairly easy for the provider to track: job entry can be validated with the independent employers, or tax checks can be used as a validation proxy; job sustainment can be checked via tax records or by contacting the employers²³⁸.

Moreover, a **process evaluation** was carried out that aimed to understand how the two variants of the SIB model work in practice, and to identify the barriers to and facilitators of successful commissioning and delivery of the model, from the perspective of commissioners, service providers, service users and MHEP. In this evaluation, qualitative methods were used to gain an in-depth understanding of the views and experiences of different stakeholder groups in relation to the commissioning and delivery of the model, as well as to inform its potential scaling-up²³⁹. Other evaluations were carried out using **qualitative analysis** of data collected through interviews and focus groups. The Behavioural Insights Team carried out the former, while ECORYS and ATQ Consultants conducted the latter.

Moreover, data on costs were collected directly from the providers and from the MHEP team via an interview. Out of the five IPS providers who participated in the evaluation, four provided cost data. Providers supplied information about the cost of their service overheads, staffing, equipment, and any unexpected costs generated during the delivery of the IPS service.

Measurement of outcomes relating to payment mechanisms

The evaluation of outcomes relating to employment was based on validated administrative data provided by the service providers. The providers formally reported on a quarterly basis, along with monthly updates and conference calls.

To determine whether payment should be issued, the outcome measured (in terms of the number of users who had achieved that outcome) was compared against the expected performance agreed at the outset with the commissioners, CBO and SOF, and stated in the contract bid.

²³⁸ Rephrased from interview with the evaluator.

²³⁹ Gadenne V., Maglicic M., Nolan D., Wright H. & Frerichs J. (2020). Individual Placement Support: A Social Impact Bond Model. The Behavioural Insights Team.

The provider was in charge of collecting the data and the evidence for each outcome. An independent third-party validator, The Behavioural Insights Team, validated the outcomes, while MHEP worked on the data cleaning.

The outcomes upon which payment was conditioned fall under the following domains: **initial engagement of the user, entry into employment, and sustainment of employment**. For each domain, specific indicators were identified, as in the following table.

Table 31. Indicators for each outcome domain

Outcome domain: initial engagement of the user <u>(Not Linked to Payments)</u>	
Successful engagement of users	Service user actively engaged, based on the completion of a vocational profile and an action plan
Outcome domain: Entry into Employment	
Job starts	Service user started a job
Outcome domain: sustainment of employment	
Job sustainment outcomes (<16 hours/week)	MHEP's target outcomes include improved wellbeing and self-reliance for the individuals concerned ²⁴⁰
Job sustainment outcomes (>16 hours/week)	NA
Job sustainment	Six weeks
Job sustainment	Six months

Other evaluations

Process evaluation²⁴¹

Data for the process evaluation were collected using semi-structured interviews and focus groups between August and September 2019. In all, 31 participants took part in the evaluation. These included: two commissioners; 12 employees of service providers; six service users; and 11 MHEP employees/members of the MHEP board. For service users, a convenience sample was used. Two researchers analysed the data using the **framework approach**, which is an established approach within applied policy evaluations.

Qualitative analysis

The first in-depth review of the Mental Health and Employment Partnership (MHEP) SIB was based on interviews carried out in 2016 with three public sector commissioning bodies – Staffordshire County Council (Staffordshire), the London Borough of Haringey (Haringey) and Tower Hamlets Clinical Commissioning Group (Tower Hamlets); and three providers of services – Making Space, Twining Enterprise and Working Well. A social investor was also

²⁴⁰ Big Issue Invest (2019b). "Mental Health and Employment Partnership case study", Webpage, <https://bigissueinvest.com/case-studies/mental-health-and-employment-partnership>; Big Issue Invest (2019a), p. 33.

²⁴¹ Gadenne et al. (2020).

interviewed – Big Issue Invest (BII), as well as and the MHEP contract manager (MHEP team).²⁴²

A second review focused on the delivery mid-way through the SIB. This was based on interviews with stakeholders between January and February 2018. A focus group was also carried out with Working Well Trust's IPS service team.²⁴³

The respondents (which were the parties involved in the implementation of the SIB and not the service users) were asked to score their overall impression of how the MHEP SIB had gone to date. Therefore, the method mostly assessed the effectiveness of the scheme rather than indicators relating to the intervention's outcomes.

Barriers to and enablers of the overall evaluation process

An experimental approach was not chosen because:

- Services **were commissioned and set up before the evaluation work began**, so it was not possible to integrate experimental procedures into the model.
- The selection of teams for IPS was subject to a number of factors, including the preferences of the National Health Service (NHS) Trust and the service commissioners, as well as practical staffing considerations. Moreover, participants were deliberately referred to IPS because of their interest in getting back to work, and were often transferred between teams as their needs changed. **These procedures could not be adapted for the purpose of the evaluation.**
- The **limited number of delivery sites**.

The main enablers and strengths of the measurement process that emerged from the desk research concern the **clear definition and understanding of outcomes**. The evaluation report also highlighted that, for the service provider, the **frequency of reviews** helped to address issues more quickly. Furthermore, agreement by all the parties on a **single and shared way of reporting data**, in terms of both form and content, was reported as being another strength of the evaluation process. Finally, including a '**ramp-up stage**' in the outcome targets for volumes may have been crucial, especially given that the service was not already in place, because it takes time to build up all the relationships needed to deliver the service.

With regard to the barriers to and weaknesses of the measurement process, service providers struggled with the pressure and additional scrutiny related to outcomes measurement. In particular, providers felt the **burden of additional reporting**, which led to a great amount of extra work in relation to data collection and reporting, sometimes involving multiple reporting lines.

Indicators based on the volumes of users referred, or those who might suffer from shortfalls in capacity (although payments are not linked to engagements) created a great deal of pressure on service providers. Indeed, to remain compliant with the outcome target set in the bid for the users engaged, **providers might be forced to increase the case load for each professional** and be constantly worried that a decrease in the volume, which is out of their control, might affect the achievement of targets.

Moreover, the configuration of the SIB's metrics did not include sufficient margin for error, thus providers struggled to overcome shortfalls in capacity.

²⁴² CBO (2016). Mental Health and Employment Partnership (MHEP) Social Impact Bond – produced as part of the CBO Fund Evaluation.

²⁴³ CBO (2017). Midpoint in-depth review report produced as part of the CBO Fund Evaluation. Retrieved from: https://www.tncommunityfund.org.uk/media/research-documents/social-investment/commissioning_better_outcomes_in_depth_review_190320_122442.pdf?mtime=20190320122441

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

Another challenge for the service providers was that the MHEP SIB imposed different targets from those recommended by the Centre for Mental Health for achieving Centre of Excellence status. Specifically, MHEP had a higher target for the successful engagement level. This, on top of consistent requests from the contract management for additional data that were not always related to the measurement of outcomes linked to payments, increased the **burden in terms of data collection** for the service provider.

Lastly, another issue was that outcomes triggered payment on the basis of a **binary metric** (yes / no). This limited the opportunity for performance analysis broken down by customer profile (e.g. minority ethnicity or educational attainment), which would have been helpful to commissioners in understanding the cohort's requirements.

Table 32. Summary of the measurement process in the MHEP SIB

Measurement methodology	
Methods relating to payment mechanisms	Quantitative evaluation based on secondary data collected by service providers
Other evaluations	Qualitative analysis of data collected through interviews and focus groups
Methods using experimental or quasi-experimental design	No
Control group	No
Causality of impact	No
Enabling factors and strengths of the evaluation process	<ul style="list-style-type: none"> - Simple and standardised data collection and reporting - Design of metrics allowing a 'ramp up' phase in the implementation of the project - Frequent reviews Unaffected by changes in volume
Barriers to and weaknesses of the evaluation process	<ul style="list-style-type: none"> - Data only focused on outcomes that trigger payment - Burden felt by service providers in relation to collecting additional data on top of the usual requirements and remaining compliant with the rigour required by the method - Pressure created by metrics linked to volume of referrals Binary metrics
Governance of measurement	
Independence of evaluator	Yes
Evaluator	The Behavioural Insights Team (consultancy) ECORYS (consultancy) ATQ Consultants (consultancy)

1.8.4. Effectiveness

Achievement of outcomes

Between January 2016 and March 2019, the MHEP SIB engaged 1,926 users, 590 of whom started work. This means that, on average, 31% of people engaged across all providers started work. Across the providers, 21.5% of all participants achieved a six-week outcome (so, roughly 70% of individuals who found a job sustained it to the six-week mark). In most cases, targets in terms of job starts and job sustainment at the six-week and six-month mark were not achieved.²⁴⁴ The evaluation report issued by The Behavioural Insights Team²⁴⁵ confirmed that all providers successfully engaged and supported individuals in finding employment to varying degrees, while long-term job sustainment remained a challenge.

The TF scheme in Camden exceeded its targets in terms of the number of clients referred and assessed (203 against a target of 192). However, the pilot did not engage as many clients or meet its target for sustained paid employment. The provider engaged 118 clients, of whom 30 clients acquired paid employment (25%). Lastly, only eight (26%) clients sustained employment after six months out of all total job outcomes (30).²⁴⁶

Benefits and drawbacks

The first benefit attributed to the SIB compared with the TF equivalent was its **focus on outcomes**. Team members felt more accountable to clients, as they wanted to achieve the targets. Constant monitoring of performance created a focus on the achievement of outcomes, collecting good-quality data, and on services being fixed promptly if something was not going according to plan.

Even though neither the TF scheme nor the SIB met their respective targets, users of both programmes appear to have been satisfied with the services. In the case of the Camden TF scheme, the evaluation included a number of interviews conducted with clients. Based on our analysis, most of these clients reported improved mental wellbeing, and the staff reported improvements with regard to mental health and employment. Clients found the practical support really useful, such as CV writing, interview practice, and guidance about the interview process and online job searches. One particularly positive element of the pilot for clients was the fact that support was personalised.²⁴⁷ The MHEP programme scored 3.5 out of 5 across all respondents who were asked for their overall impression of how the MHEP SIB performed (on a scale of 1 to 5, with 3 being Fair).

Job outcomes were seen as the main focus of the support service users received. However, service **users saw job outcomes as a wider category than simply getting a job**: they included completing up-to-date CVs, attending interviews, and gaining knowledge on how to look for employment. They viewed these outcomes as comparable to gaining employment because they had positively impacted their ability to find work. This employment support, while provided in an environment related to mental health, contributed to service users' trust that the service would be tailored to their situation. Alongside the feeling that the service was helping them move closer to employment, service users described feeling **increased confidence, a new-found sense of independence, or reduced anxiety** as a result of receiving this support. Receiving employment support from the staff, who were trained to work with mental health issues, meant that they could discuss

²⁴⁴ Gadenne et al. (2020).

²⁴⁵ Gadenne et al. (2020).

²⁴⁶ Begum, (2018).

²⁴⁷ Begum, (2018).

their mental health with their employment specialists if they wanted to, and **felt supported** when they did so.²⁴⁸

Service providers noted that the SIB allowed funding and financial support to be unlocked for a service which they felt would not otherwise be considered for **statutory funding**. More specifically, the SIB allowed the **combination of large-scale national funding (through the National Outcome Payer) with local funding** for mental health services. This contrasts with experience in the Camden IPS. One of the main barriers to uptake in this TF programme was the inability to combine health funding with outcomes-based employment funding. Indeed, despite the significant interrelationships between mental health and unemployment, a disconnect existed between large-scale national funding targeted at helping people secure employment, and locally-funded mental health support services.²⁴⁹

The funding provided through the SIB was also **more stable** for providers, which enabled **services to be run on a larger scale**. For example, in Staffordshire, the new contract extended the IPS service to the north of the county for the first time taking on the new outcomes payment basis.

The downside of the SIB compared with the traditionally financed model concerned the difficulty in forecasting the **volume of referrals**, which was required to set targets. This proved to be an issue because the service providers' bids set high expectations in terms of the number of participants they could engage. In practice, they were unable to meet these expectations during the implementation of the intervention.

Furthermore, the contracts **were vulnerable to policy changes** in community mental health services. While this risk exists for any type of contract, the agreed basis for the outcomes payments in the MHEP SIB meant that anything that affected a provider's ability to work and deliver as originally planned, also had a potential financial impact.

Finally, the **complexity of the SIB contracts** compared with more traditional funding models was seen to affect the delivery team's ability to provide the service. Providers themselves described having to spend an excessive amount of time on contract-related tasks, paperwork, and on making sense of the implications of the contracts.

Scalability and replicability

First, the SIB in Haringey was extended from April 2017 to April 2019 to engage up to 1,000 adults. Then, another SIB began in May 2019 which will last until April 2023. This SIB concerns the delivery of IPS services, and is supported by the Life Chances Fund.²⁵⁰

The Mental Health and Employment Partnership social impact bond has now been replicated in three London boroughs: Barnet, Camden and Enfield. This five-year social impact bond will support the delivery of Individual Placement Support services to assist up to 2,632 adults with severe and enduring mental health issues to gain and retain employment.

Our analysis suggests that the first three SIBs helped the commissioner to gain confidence from the fact that other such contracts were already in operation. In addition, thanks to the MHEP scheme, there was an available model including an Invitation to Tender specification and outcomes payments structure that was ready for use by other commissioners in the future, as well as a performance management infrastructure in place. However, according to the intermediary, it was still challenging to standardise the contract because the requirements of the commissioners differed between locations. When replicating the

²⁴⁸ Gadenne et al. (2020).

²⁴⁹ Bromage and Swersky (2018).

²⁵⁰ Government Outcomes Lab (2019b). Mental Health and Employment Social Impact Bond (Haringey & Barnet). Retrieved from: <https://golab.bsg.ox.ac.uk/knowledge-bank/indigo-data-and-visualisation/impact-bond-dataset-v2/INDIGO-POJ-0176/>

schemes, the intermediary decided to target areas in which the target population was 'harder to help'. This required some adaptation of the model.

Some other elements have also been changed in subsequent SIBs:

- longer start-up stage requirement;
- higher block funding or engagement payments in the revenue model, funded upfront by MHEP's investment funding.

1.8.5. Efficiency

Intervention costs

A cost comparison between the MHEP SIB and the Camden IPS is presented in the table below. Please note that this is a 'simple' comparison, meaning that we do not control for differences in participant characteristics and respective savings gained from their employment. Despite the differences between the two programmes (see Table 30 in Section 1.8.2), **the costs appear fairly similar**. The cost per programme participant equals GBP 1,200 and GBP 1,254 in the MHEP SIB and Camden IPS, respectively. The cost per job achieved in the MHEP SIB is 13% lower than in the Camden IPS, but this is not a substantial difference, given that the scale of the MHEP SIB was seven times greater (based on the total cost), which lends itself to cost-savings. There is also some information in the literature that assesses how effective the two programmes have been in terms of achieving sustained employment (please see Section 1.8.4). However, we do not include such analysis here because the available information is not comparable. The available cost per sustained job achieved for the MHEP SIB measures employment sustained for six weeks²⁵¹, whereas the Camden IPS counterpart measures employment sustained for six months²⁵².

Table 33. Cost comparison between the MHEP SIB and Camden IPS (GBP)

	MHEP SIB 2017 - 2019 ²⁵³	Camden IPS JAN 2016 - MAY 2017
Participants	N/A	118 ²⁵⁴
Jobs achieved	N/A	30 ²⁵⁵
Sustained jobs achieved	N/A	8 ²⁵⁶
Total cost	1.1 million	147,990 ²⁵⁷
Cost per participant	1,200	1,254 ²⁵⁸
Cost per job achieved	4,300	4,933 ²⁵⁹

²⁵¹ Gadenne et al. (2020), 29.

²⁵² Begum (2018), 28.

²⁵³ Gadenne et al. (2020), 27-31.

²⁵⁴ Begum (2018), 14.

²⁵⁵ Begum (2018), 14.

²⁵⁶ Begum (2018), 14.

²⁵⁷ Derived by PPMI from the cost per job achieved and the number of jobs achieved. The figure is consistent with information cited in Begum (2018), 28, which puts the cost of the programme at GBP 148,000.

²⁵⁸ Derived by PPMI from the cost per job achieved, the number of jobs achieved, and the number of participants engaged in the programme.

²⁵⁹ Begum (2018), 28.

Source: elaborated by PPMI, based on Gadenne et al. 2020 and Begum 2018.

Note: the figures quoted for the MHEP SIB are based on the funding received by providers, in order to make them comparable with the figures for the Camden IPS. The true costs of the MHEP SIB are slightly higher²⁶⁰ because some providers incurred losses, meaning that they cross-subsidised from other streams.

While the intervention costs appear similar, the commissioner noted that “PbR funding model has given greater sense of urgency and drive as a collective because [they] all want to see it work and want it to be financially viable for the provider.”

Operational costs

It is not possible to compare the MHEP SIB’s operational costs with those of the Camden IPS because limited information is available regarding the costs of the latter intervention. Therefore, we outline below only the information regarding the MHEP SIB’s operational costs and provider insights from interviews where relevant.

The SIB’s management costs for activities between 2017 and 2019 amounted to slightly less than a third of the funds spent by the service providers (see table below). Please note that management costs are not included in the funds spent by the SIB service providers to deliver the services; these costs are in addition to the intervention costs. Management costs in this case include: the cost of MHEP staff wages; audit, insurance and board costs; and the exceptional use of consultants for the special purpose vehicle (SPV); the cost of time for a specialist team member, providing operational advice and guidance to providers to improve performance; investor returns, debt interest payments and any additional project surplus; as well as evaluation costs.²⁶¹

While a clear comparison is not possible between the SIB’s management costs those of the Camden IPS or another TF scheme, **the commissioner noted that the MHEP SIB required significantly more of their time and resources compared with other contracts because contracts had to be re-negotiated**. Providers, too, described having to spend an excessive amount of time on contract-related tasks, paperwork, and on making sense of the implications of the contracts. Furthermore, they reported that they had to commit greater resources to data collection and reporting than in previous contracts. The evidence gathering and reporting burden took up a quarter of a day per week for each team member, and half a day per week of the team leader’s time. However, according to the evaluator, greater rigour resulted in better outcomes. Lastly, the intermediary noted that involving multiple investors increased the SIB’s management costs.

Table 34. SIB intervention costs compared with Management costs, 2017-2019 (GBP)

Funds spent by SIB service providers	1,142,264 ²⁶²
Management costs	319,000 ²⁶³
Operational costs, expressed as an equivalent percentage of funds spent by SIB providers	28% ²⁶⁴

Source: elaborated by PPMI, based on Gadenne et al. 2020.

Note: Operational costs are not included in the funds spent by SIB providers. In other words, these are additional costs.

²⁶⁰ See Gadenne et al. (2020), 28.

²⁶¹ Gadenne et al. (2020), 29-30.

²⁶² Gadenne et al. (2020), 31.

²⁶³ Gadenne et al. (2020), 30.

²⁶⁴ Derived by PPMI based on the information presented in the table.

Set-up costs are not included in the table above. As shown in the table below, **these amount to roughly 5% of total potential outcome payments.**²⁶⁵ By set-up costs, we mean the grant received by Social Finance from the DWP to find the commissioners, investors, set up MHEP, etc. Including the work of the commissioners involved would further increase this estimate. Nevertheless, according to the commissioner interviewed, **commissioning and bidding costs were not significantly different from normal.**

Table 35. MHEP SIB set-up costs in relation to total potential outcome payments (GBP)

SIB's total potential outcome payments	2.9 million ²⁶⁶
Set-up costs	150,000 ²⁶⁷
Operational costs, expressed as an equivalent percentage of funds spent by SIB providers	5% ²⁶⁸

Source: elaborated by PPMI, based on information in CBO 2016.

Desk research provided additional insights into the set-up process. It took over **21 months** between the inception of the SIB idea to contract delivery going live (summer 2014 to April 2016). As noted in the review of the SIB in 2014, "Social Finance was engaged in discussions with various parts of central and local government including the [DWP] about potential new areas where the application of social investment might be appropriate. DWP was itself interested in reviewing potential interventions for the 'hardest to help' referrals in the Work Programme and, particularly, those with health and mental health conditions... This thinking led to an expression of interest to the CBO/SOF by DWP. Only local commissioners and agents working with them were eligible for CBO development grant funding, so DWP appointed Social Finance to engage with commissioners, including local commissioners and to research the feasibility of using social investment to support a specialised mental health employment intervention. A development grant of GBP 150,000 was awarded to Social Finance for this work."²⁶⁹

Social Finance subsequently researched effective interventions for mental health patients, and came across IPS services. Based on the available evidence, Social Finance approached both central government and the NHS as well as other stakeholders. In late 2014, the DWP was not in a position to commission another programme, so the focus remained on local commissioners and government agencies. To allow multiple stakeholders to invest, Social Finance established MHEP, which took until the third quarter of 2013. SOF was the first investor to commit funds to the SIB (amounting to GBP 1 million). This was instrumental in convincing other commissioners to join, along with the work Social Finance had done to gather information on the effectiveness of IPS.

Procurement for the programme started in the last quarter of 2015. At the beginning of 2016, evaluation of provider bids and the awarding of IPS provider contracts was completed jointly by MHEP and the respective commissioning authorities. MHEP's role included supporting commissioners with the ITT, the specification of the IPS service, and defining outcomes and outcome payment structures. Each commissioner ran its own selection process and awarded the contracts to the successful providers.²⁷⁰

²⁶⁵ Comparison with the total potential outcome payments here is more appropriate than with the funds spent by SIB service providers because set-up costs are related to the number of investors and the overall funds that are made available for the SIB, rather than the funds actually spent. This comparison might also be more informative for commissioners designing other SIBs, because it is difficult to predict the actual outcomes and hence funds spent on the intervention during the design stage.

²⁶⁶ CBO (2016), 5.

²⁶⁷ CBO (2016), 7.

²⁶⁸ Derived by PPMI from the information presented in the table.

²⁶⁹ CBO (2016), 7.

²⁷⁰ CBO (2016), 7-9.

1.8.6. Scheme design

Design features

The set-up process described in the previous section illustrates the importance of having an **intermediary**. Social Finance found the commissioners as well as the investors, and set up the MHEP. The MHEP contract managers in turn were essential to the measurement process, especially in the supervision of data collection. Furthermore, according to the investor, MHEP put resources in place to support the providers and commissioners with both operational issues and contract re-negotiations. The support of the MHEP contract management team also meant that commissioners were open to contracts being re-negotiated. MHEP's structure was flexible enough to run different arrangements in each location. As a result of this flexibility, the models could develop and evolve, at the same time contributing to best-practice recommendations for the expansion of the IPS services. Lastly, MHEP's legal structure can be readily used by other commissioners in the future, with a ready-made ITT specification and outcomes payments structure for an IPS service, as well as a performance management infrastructure.

In addition to emphasising the importance of an intermediary and the added value provided by MHEP's structure, interviewees underlined **good communication among the stakeholders**. As a requirement partly from the social investors, quarterly meetings took place to which providers, commissioners, MHEP and MHEP board representatives were invited. These meetings were felt to be constructive and helpful in pulling everyone together towards a common goal. The evaluation report suggested that "for such a model to evolve further, the relationships between partners are of as much importance as the contract structure."²⁷¹

In addition, **combining block payment with PbR (with an emphasis on block funding) helped to solve performance challenges and issues relating to the unpredictable volume of users**. Block funding was especially important, because the service had not already been in place and needed to get off the ground; once this had been achieved, it was possible to move to outcomes-based payments.

Despite the design features that stakeholders generally agreed were positive, there were other elements that stakeholders believed could have worked better:

- The need for **multiple reporting lines** added a layer of administrative burden on providers. One provider explained that they had to create different reports for MHEP and the NHS, which necessitated spending a significant amount of time training their teams to report in such a way. Another provider identified three stakeholders to whom they reported: the local authority, the Clinical Commissioning Group commissioners, and Social Finance as the MHEP contract manager, who in turn reported to BII as the investor.
- **More time would have been appreciated for the mobilisation and start-up phase**. The transition in the contracts from block payments to a PbR model left little time to enhance service capacity and capability. Funding should therefore be available under the contract to support these activities. The evaluation report noted that all stakeholders, particularly voluntary, community and social enterprises, need to **carry out scenario testing** in terms of giving forethought to the impact of various outcome scenarios on their resilience.²⁷²
- Finally, interviewees argued that **bids should be assessed more rigorously**, particularly with respect to the provider's performance forecast and financial

²⁷¹ Gadenne et al. (2020).

²⁷² CBO (2017).

resilience in low-volume scenarios. KPIs and targets relating to engagement volume appear to be a constant pressure on service providers. Therefore, if participant volumes appear substantially lower or higher than expected, contracts should leave room for a **re-negotiation of terms**.

Impact on third-sector organisations and social enterprises

The service providers in the SIB argued that **the SIB's structure presented greater risks to them than traditional contracts**. When referrals and outcome revenues fell short of expectations, providers perceived that all of the financial risk had been transferred to them and were faced with a dilemma over whether to spend more money investing in the team/performance capability in the face of uncertain future outcome-only revenues. Furthermore, they argued that there is not enough margin for shortfalls in capacity, i.e. when a team member leaves, there can be gaps before new starters are found and get up to speed. The proportion of people was higher leaving compared with situations funded under block contracts, because the SIB payment structure created uncertainties. In addition, performance-based outcome payments could make planning and budgeting more difficult because of the unpredictability of future performance. Finally, the configuration of this SIB meant that the contracts were vulnerable to policy changes in community mental health services. The intermediary interviewed concluded that "if you totally relied on outcome payments, this is not feasible for small third-sector organisations that do not have the managerial and operational capacity."

1.9. JobPath (Ireland)

1.9.1. Background

In 2015, Ireland's Department of Employment Affairs and Social Protection (DEASP) introduced JobPath as a response to the increasing numbers of the **long-term unemployed** following a dramatic collapse in employment between 2009 and 2012²⁷³. The programme is still ongoing, but in this study, we focus on its implementation between 2015 and 2018, because the programme results from this period have already been evaluated.²⁷⁴

Through JobPath, employment services were contracted out. According to the DEASP, **this commissioning approach was chosen due to the cyclical nature of unemployment**: "providing high-quality activation and case management capacity using an in-house permanent staff cadre is challenging in a situation where there are cyclical, and sometimes rapid, changes in the number of unemployed jobseekers. That is why the Irish PES [Public Employment Service] has always had recourse to contracted services that are flexible across the economic cycle. This has ensured a responsive system of employment support for those in unemployment despite the recurring economic cycles of low unemployment followed by high unemployment, a risk that is associated with Ireland's position as a small open economy."²⁷⁵

To be selected for JobPath, all long-term unemployed jobseekers on the Live Register²⁷⁶, aged between 18 and 61 years old inclusive, were categorised into groups based on duration of unemployment (i.e. 1-2 years, 2-3 years, etc.). Selection for referral to JobPath

²⁷³ Department of Social Affairs and Social Protection (DEASP). (2019). Working paper: Evaluation of JobPath outcomes for Q1 2016 participants. Retrieved from: <https://assets.gov.ie/36499/ffdce98cddc34addb05cf41a70aaf4e7.pdf>, vii.

²⁷⁴ DEASP (2019).

²⁷⁵ DEASP (2019), p. 4.

²⁷⁶ The Live Register is used to provide a monthly series of the numbers of people (with some exceptions) registering for Jobseekers Benefit (JB) or Jobseekers Allowance (JA) or for various other statutory entitlements at local offices of the Department of Social Protection. Information is published in the form of a monthly release titled the Live Register. <https://www.cso.ie/en/interactivezone/statisticsexplained/labourmarket/whatistheliveregister/>

was by stratified random sampling using these duration-based categories.²⁷⁷ This meant that only a portion of all eligible long-term unemployed persons were referred to the programme. This was important, because those who are not referred serve as the control group to estimate the impacts of the programme (for details, please see Section 1.9.3). Participation in the programme was mandatory for selected participants. The period of engagement in the programme for each client was 52 weeks.²⁷⁸

With regard to programme **activities**, participants on JobPath received individual support from contracted providers to help them address barriers to employment and to assist them in finding jobs. During this time, jobseekers had access to a personal advisor who worked with them over two phases. During the first phase, which lasted up to 12 months, the personal advisor provided practical assistance in searching, preparing for, securing and sustaining employment. During the second phase, if the participant had secured employment, the service provider remained in contact with the participant for the first 26-52 weeks of employment.²⁷⁹

More specific services included:²⁸⁰

- Assessment of client skills, competencies and aptitudes.
- Development of a personal progression plan for each client and the review of this plan on regular basis.
- Assistance with job search.
- Development of the jobseeker's curriculum vitae.
- Development of job interview skills.
- Training, education and other interventions up to 26 weeks.
- Support in the transition to employment, including a period of 'in-employment' guidance/counselling.
- Access to computers, the internet and other facilities to aid clients in their search for employment, with support on how to use these tools.
- Support in developing key skills to assist clients to sustain employment, e.g. team working, organisation and time management skills.
- Support to deal with other issues that may make it harder for clients to find and sustain employment; for example, support with managing a health/disability-related condition or advice on managing finances.
- Other services or support to enhance the client's prospects of securing sustainable employment.

The prescription of common minimum service provision is set out in the JobPath **Service Guarantee**. This stipulates that providers must hold one-to-one meetings within 20 days of referral, and agree a personal progression plan with all clients. The plan must contain basic contact information and a detailed record of the client's employment experience, skills and barriers to employment, and a plan setting out the client's return to work trajectory. The provider must hold a further one-to-one meeting with the client every 20 working days until they gain employment, and a full review of the plan must take place periodically (at 13, 26, 39 and 52 weeks) if the client remains unemployed.²⁸¹

²⁷⁷ DEASP (2019), p. viii.

²⁷⁸ DEASP (2019), p. 8.

²⁷⁹ DEASP (2019), p. viii.

²⁸⁰ DEASP (2019), p. 17.

²⁸¹ Wiggan, J (2015), n.p.

With regard to **payment** arrangements, contractors were paid an initial referral fee (EUR 311 per participant on average). Further payments were made when jobseekers remained in employment for 13, 26, 39, or 52 weeks (see the table below). The contractors had to verify the duration of the employment with the co-operation of the jobseeker. This structure aimed to reward sustainable employment where the client remained in employment for at least 12 months.²⁸² JobPath providers could potentially receive up to EUR 3,718 per client on average, though in reality payments were lower because not all participants remained employed for 52 weeks.²⁸³

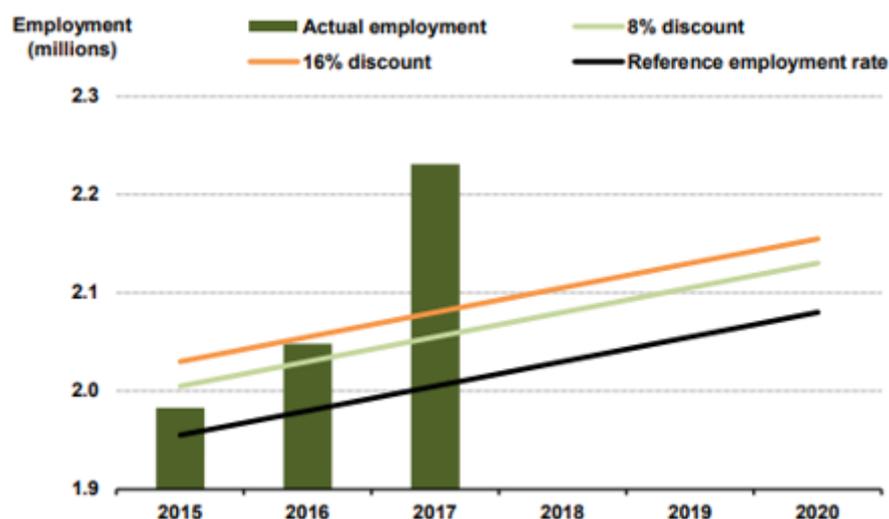
Table 36. Average payments per JobPath participant, based on weeks in employment

	Registration	13 weeks	26 weeks	39 weeks	52 weeks	Total
Share of total payment	8.40%	16.50%	19.80%	24%	31.30%	100%
Payment (EUR)	311	613	737	892	1,165	3,718

Source: DEASP 2019, 17-18.

Importantly, the contract prescribed **discounts** to registration and job sustainment fees in the event of more favourable economic conditions that would lead to greater than anticipated employment levels. The figure below shows the reference employment levels prescribed, and the levels of employment at which discounts of 8% and 16% could be applied by the DEASP.

Figure 3. Reference employment levels 2015-2020 and actual employment levels 2015-2017



Source: Office of the Comptroller and Auditor General 2018, 147.

Furthermore, if a prime provider failed to achieve **satisfactory ratings** in the annual customer survey or performance inspections, the Department could withhold up to 15% of the fees payable.²⁸⁴

Although providers were paid on the basis of outcomes achieved, the Department also specified **minimum performance metrics**. These included two thresholds – minimum

²⁸² DEASP (2019), p. 17-18.

²⁸³ Ibid.

²⁸⁴ Wiggan, J (2015), n.p.

performance and reference performance. Minimum performance was set at 30% above a benchmark figure of the total number of individuals that could be categorised as part of each of the customer groups that moved into employment during the 12 months between November 2012 and October 2013. Reference performance was set at 60% above this same benchmark for four of the groups, and 70% for two of the groups. Provision were also made for the application of penalties where performance targets were not met.²⁸⁵

Overall, the outcomes achieved exceeded expectations. During the first year, around 22% of long-term unemployed participants moved into full-time paid employment. This exceeded the 14.8% employment target.²⁸⁶

For a summary of stakeholders involved in JobPath, please see the table below.

Table 37. Summary of JobPath

Personally targeted social service	Labour market related
Target population	Long-term unemployed
SOC scheme type	Mixed PbR ²⁸⁷
Payment model	Frequency
Cost	2015: EUR 1.2 million 2016: EUR 28.6 million 2017: EUR 58.5 million ²⁸⁸
Commissioner(s)	Minister for Social Protection ²⁸⁹ , Department of Social Affairs and Social Protection (DEASP)
Social service provider(s)	Turas Nua delivers services in the southern counties; Seetec delivers services in the northern counties (both private for-profit)
Investor(s)	N/A
Intermediary	N/A
Evaluator(s)	Payment-related outcomes were reported by providers. Furthermore, an impact assessment was also carried out as part of a partnership between the Statistics and Business Intelligence Unit of the DEASP ²⁹⁰ and the Directorate for Employment, Labour and Social Affairs of the Organisation for Economic Co-operation and Development (OECD).

1.9.2. Equivalent TF scheme

Before the roll-out of JobPath, long-term unemployed individuals received education and training programmes, public employment programmes such as Community Employment, or self-employment incentives. No job search and assistance programme was designed for, or offered specifically to, long-term unemployed people.²⁹¹ Therefore, there is no exact equivalent TF scheme for JobPath. Nevertheless, there are two TF schemes that are similar to JobPath. One is **Intreo**, a single point of contact for all employment/income support for

²⁸⁵ Lowe, Simon (2015). JobPath: The proposed introduction of an employment programme in the Republic of Ireland. *The Public Sphere*, 124.

²⁸⁶ DEASP (2017).

²⁸⁷ In a **pure PbR**, the Commissioner does not provide up-front money, while in the **mixed PbR** type, a proportion of payment is provided upfront, which is not contingent on the achievement of a specified outcome. Please see more information on the conceptual framework developed in the Inception report.

²⁸⁸ DEASP (2019), p. 17.

²⁸⁹ Department of Employment Affairs and Social Protection (DEASP). (2015). JobPath - Request For Tenders. Retrieved from: <https://www.gov.ie/en/collection/8fbd78-jobpath-request-for-tenders/>

²⁹⁰ The Statistics and Business Intelligence Unit of the Department of Employment Affairs and Social Protection is part of the Irish Government Statistical Service, which is headed by the Central Statistics Office. It operates in association with, and contributes to, the wider Irish Government Economic and Evaluation Service (IGEES).

²⁹¹ DEASP (2019), p. 44.

social protection payment recipients and engagement with local employers. Intreo is responsible for the processing of all claims and payments, as well as any queries relating to DEASP payments. Intreo activation services are targeted at **short-term** Jobseekers (duration on Live Register of less than one year) as a priority for case officers to stem the flow into long term employment. It operates a 'work first' policy: jobs first for those that are ready to enter/re-enter the labour market; education and training for those who are not.²⁹² The unemployed engage with services provided by Intreo for a year prior to being referred to JobPath (i.e. prior to being classified as long-term unemployed).

Another programme is **Local Employment Services (LES)**. These services are contracted out to private providers by DEASP through traditional contracts. LES provides a range of services and facilities to assist unemployed persons to enter/re-enter employment. LES works with activation and non-activation clients: activation clients are those for whom the service is mandatory (if they fail to attend, their benefit rate may be cut); non-activation clients may still access the service on a voluntary basis with no penalty for failure to attend. Activation clients are referred to LES by Intreo, and are of working age and in receipt of specified social welfare payments. Non-activation clients are more distant from the labour market. These include persons with a disability, lone parents, early school leavers, non-Live Register returnees to employment, members of the travelling community, refugees, low income smallholders and ex-offenders.²⁹³

According to the Department of Public Expenditure and Reform, "it is unclear what criteria is used to determine the most appropriate provider for an individual's activation journey as there are a number of overlaps between providers of services to the Long-Term Unemployed."²⁹⁴ With the introduction of JobPath in 2015, referrals to LES were made on a replacement basis²⁹⁵ only, to deliver more intensive client engagement.²⁹⁶

In general, JobPath, Intreo and LES provide similar services (see the table below), are all overseen by the same Department, and are all running concurrently. However, Intreo focuses mostly on the short-term unemployed (the long-term unemployed can reach out to Intreo on a voluntary basis), which is both an advantage and a disadvantage for our study. It is a disadvantage because the target groups are different and hence their outcomes, for example in terms of employment rates, are not comparable, making it difficult to compare their effectiveness. It is an advantage because all persons referred to JobPath have previously engaged with Intreo and there is available research comparing their effectiveness. Meanwhile, although LES was designed to help the long-term unemployed, it is not always clear when the long-term unemployed would be referred to LES²⁹⁷ and the exact composition of LES clients varies. Therefore, any comparisons must take into account that the target groups in each of these programmes are not identical.

²⁹² Lavelle, Orlagh & Niamh Callaghan (2018). Public Employment Services – Mapping Activation. IGees & Department of Public Expenditure and Reform. Retrieved from: <https://igees.gov.ie/wp-content/uploads/2018/07/12.-Public-Employment-Services.pdf>, p. 2.

²⁹³ Indecon (2018). Indecon Review of Local Employment Services. Retrieved from:

<https://www.dsfa.ie/en/pdf/IndeconReviewofLocalEmployment%20Services-Final%20Report.pdf>, p. 10.

²⁹⁴ Lavelle and Callaghan (2018), p. 14.

²⁹⁵ Replacement basis means that, as of 2016, a maximum caseload of 120 clients per mediator must be maintained, with clients referred to LES on a replacement basis only (a new referral is possible only when a jobseeker/client enters employment or finishes their 12-month engagement period with the LES service).

²⁹⁶ Lavelle and Callaghan (2018), p. 2.

²⁹⁷ Lavelle and Callaghan (2018), p. 2.

Table 38. Comparison of Jobpath, Intreo and LES services

	JobPath	Intreo	LES
Provision of labour market information	✓	✓	✓
Active placement and recruitment service	✓	✓	✓
Career guidance/counselling and mediation	✓	x	✓
Employer liaison	✓	✓	✓
Referral to training, education, work experience and employment programmes	✓	✓	✓
Support after training, education, work experience and employment programmes	x	✓	✓
In-employment support	✓	x	✓
Drop-in service	✓	✓	✓
Non-activation clients	x	✓	✓
Interview preparation	✓	✓	✓
CV preparation	✓	✓	✓
Provides funding to remove barriers to employment	✓	✓	✓

Source: Adapted by PPMI from Lavelle and Callaghan 2018, 22 and updated on the basis of information from interviews.

For a comparison of stakeholders and costs for the three programmes, please see the table below.

Table 39. Comparability of jobPath, Intreo and LES

	Same as SOC	Comments
Personally targeted social service	Yes	Labour market-related
Target population	Yes/No	JobPath (SOC): Long-term unemployed Intreo (TF): Long-term and short-term unemployed (mostly the latter after the introduction of JobPath) LES (TF): Long-term unemployed and other groups further away from the labour market, for whom participation is not mandatory

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

	Same as SOC	Comments
Number of participants	Yes/No	<p>JobPath (SOC): 245,419 participants have been referred to JobPath between 2015 and 2018,²⁹⁸ equating to roughly 80,000 participants per year</p> <p>Intreo (TF): N/A</p> <p>LES (TF): roughly 90,000 participants per year in 2014 and 2015, and 47,000 in 2016²⁹⁹</p>
Location	Yes	Ireland
Cost	Yes/No	<p>JobPath (SOC): EUR 28.6 million in 2016 and EUR 58.5 million in 2017³⁰⁰</p> <p>Intreo (TF): EUR 55-58 million annually³⁰¹</p> <p>LES (TF): roughly EUR 19 million annually³⁰²</p>
Commissioner(s)	Yes	DEASP
Social service provider(s)	No	<p>JobPath (SOC): Turas Nua and Seetec (both private for-profit)</p> <p>Intreo (TF) is a service provider in itself (public agency)</p> <p>LES (TF): 22 contractors comprising 21 Local Development Companies and one community organisation.³⁰³ The vast majority of LES contracting organisations have 'company limited by guarantee' status, which is a typical status for non-profits and charities.³⁰⁴</p>

1.9.3. Outcomes measurement

In order to claim payments under JobPath, service providers were required to submit evidence of sustained employment on the part of programme participants. In addition, in 2019 an impact assessment was carried out by the Statistics and Business Intelligence Unit of DEASP in collaboration with the Directorate for Employment, Labour and Social Affairs of the Organisation for Economic Co-operation and Development (OECD) (an international organisation). Lastly, every year since the inception of the programme, customer satisfaction surveys have been carried out by the DEASP.

Measurement of outcomes relating to payment mechanisms

Providers were paid when they could provide sufficient evidence that programme participants had remained in employment for a pre-determined amount of time (13, 26, 39, or 52 weeks, with each time period triggering a separate payment). Providers were required

²⁹⁸ DEASP 2019, 19.

²⁹⁹ Indecon 2018, iv.

³⁰⁰ DEASP 2019, 17.

³⁰¹ Lavelle and Callaghan 2018, 18.

³⁰² Indecon 2018, 33.

³⁰³ Indecon 2018, i.

³⁰⁴ Indecon 2018, xi.

to furnish supporting evidence of employment, comprising verification by the participant and the participant's employer.

The expectation was that in the first year of the programme, 14.8% of the jobseekers who used JobPath would start work.³⁰⁵ This target is available in the Request to Tender³⁰⁶. The 14.8% employment expectation, together with detailed targets for each unemployment group, were based on historical data (the number of long-term unemployed people who stopped claiming unemployment benefits from November 2011 to October 2013 inclusive).³⁰⁷

Other evaluations

Impact assessment

The impact assessment used a quasi-experimental design, specifically a propensity score matching approach. The evaluation identified JobPath-eligible benefit claimants in the first quarter of 2016, and divided them into those who did not start JobPath and those who did. Please note that the long-term unemployed were randomly referred to the programme, resulting in these two groups. The data regarding both groups was drawn from the Jobseekers Longitudinal Dataset (JLD). The JLD is an administrative dataset that tracks social welfare claims, activation and training, and employment histories over time, covering people with jobseeker or one parent family claims since 2004. It also contains information on a claimant's gender, age, marital status, nationality, educational attainment, spousal earnings (to qualify for an adult dependent allowance), number of child dependents, family payment type (i.e. adult and child dependent allowances, adult only, etc.), and geographical location. For this evaluation, the JLD was supplemented with DEASP data on earnings from employment, social welfare payments data, and data on social welfare status.

Using the JLD data, the evaluator estimated the probability of being assigned to JobPath in the first quarter of 2016 using logistic regression. This procedure generated probability scores for each individual, and allowed the evaluator to estimate the inverse probability of treatment weights. Adding weights to each observation in the control group ensured that the treatment and control groups were adequately balanced and, consequently, that any comparison between them reflected only their differing status with respect to JobPath, and not underlying differences in their labour market characteristics.³⁰⁸

In addition to employment outcomes, the impact assessment focused on the following results:

- The amount of money earned from employment compared with the amount of money received from social welfare payments during the 2017 calendar year.
- The number of weeks of insurable employment during the 2017 calendar year.³⁰⁹
- These measures were compared between JobPath participants and the control group.
- The evaluation also looked at whether the impacts were significant and differed for different groups ('clusters') of the unemployed. These clusters are described below:

³⁰⁵ Department of Employment Affairs and Social Protection (DEASP). (2017). JobPath Performance Data Jobseekers Referred to JobPath July – September 2015. Retrieved from: <https://www.gov.ie/en/press-release/85249d-department-of-social-protection-publishes-first-performance-report-o/>

³⁰⁶ Department of Employment Affairs and Social Protection (DEASP). (2015), tables 7.1a and 7.1b.

³⁰⁷ DEASP 2015, 94.

³⁰⁸ DEASP 2019, ix.

³⁰⁹ DEASP 2019, 1.

- Younger Casual Claimants had the shortest claim durations, with comparatively good labour market attachment, even if they tended to have earnings only in the previous calendar year.
- Younger Professionals were largely young claimants, with a higher share of short claim durations; almost all had some history of employment.
- Intermittent Labour Market Attachment. People in this cluster had a poor employment history in the past year but evidence of intermittent employment/earnings over the past five years.
- Shorter Durations was the largest cluster. People in this cluster tended to be 30 to 40 years of age. The cluster had an above-average share of people with clerical and secretarial occupations.
- Older, With Strong Employment History was the smallest cluster. While long-term unemployed people in this cluster had a strong prior history of employment, they were largely male and included a greater share of people coming close to retirement age.
- Self-Employed. People in this cluster had often been self-employed prior to their claim. They had weak labour market attachment and their average claim duration was the second longest among all the clusters.
- Persistent Longer Durations. People in this cluster had generally been unemployed for more than two years and they rarely moved to another cluster. This cluster had the lowest share of people who were previously in managerial or professional occupations.³¹⁰

For the cluster analysis, probabilistic modelling was used to segment the available data into cohorts. A rich dataset was compiled and a clustering algorithm used to calculate the optimal number of clusters. Outcomes were compared for people in each cluster with an equivalent group who were not enrolled in JobPath.

The impact assessment measured the following outcomes (see Section 1.9.1):

- how many more people had found jobs with JobPath the year after the programme, compared with the number of people who had found employment without JobPath;
- how much more the JobPath participants earned, compared with what they would have earned without JobPath;
- how these effects differed for the different clusters analysed.

Customer Satisfaction Surveys

Two rounds of surveys – over the phone and online – were carried out with a random sample of JobPath participants every year from 2015. Both surveys shared the same questionnaire³¹¹. The **outcomes and indicators** measured through the surveys included:

- Overall satisfaction with Seetec/Turas Nua (service providers’) services.
- Satisfaction with Seetec/Turas Nua offices.
- Satisfaction with Seetec/Turas Nua staff.
- Detailed satisfaction with selected Seetec/Turas Nua services.

³¹⁰ DEASP 2019, viii-ix.

³¹¹ DEASP (2020). Surveys. Retrieved from: <https://www.gov.ie/en/publication/96cc45-surveys-2015/>

- Satisfaction with Seetec/Turas Nua processes.
- Comparison of customer satisfaction between JobPath and Intreo services.
- Employment status at the time of the survey.

Levels of satisfaction were measured via multiple Likert scale-type questions. For example, in relation to staff, respondents were asked to indicate on a scale from 'Completely agree' to 'Disagree completely' whether a provider's staff tried their best for the respondents. Similarly, respondents were asked to indicate their level of agreement with the statement that the services offered helped to improve their prospects of getting a job.³¹²

Barriers to and enablers of the measurement process

The **requirement for providers to collect proof of outcomes was perceived as inefficient**, and sometimes meant that the providers were not paid for the outcomes achieved, simply because they could not acquire the necessary proof from programme participants. Furthermore, using tax records as an alternative was not possible at the time due to data constraints.

The main strength of the impact assessment is that, given its quasi-experimental design, it can be used to draw causal inferences as to the effectiveness of JobPath. A few enabling factors contributed to the effectiveness of the impact assessment, for example, the **availability of the JLD**, which has already been developed in 2012, as well as the **random referral** of the long-term unemployed to JobPath. Random referral to the programme limited the likelihood that unobserved characteristics would be responsible for the better performance of JobPath participants when compared with the control group (in programmes with voluntary enrolment, those who enrolled might be more motivated to work than those in the control group). Attaching treatment weights to the control group ensured that the two groups were comparable, isolating the impact of JobPath.

However, the evaluation was carried out by the Statistics and Business Intelligence Unit of DEASP together with the OECD, so the evaluator was not entirely independent. On the other hand, this might have been beneficial with regard to access to the JLD.

³¹² DEASP (2020).

Table 40. Summary of the measurement process in JobPath

Measurement methodology	
Methods related to payment mechanisms	Outcomes were reported by providers, with proof of sustained employment
Other evaluations	Impact assessment using propensity score matching
Methods using experimental or quasi-experimental design	Quasi-experimental design
Control group	Yes
Causality of impact	Yes
Enabling factors and strengths of the evaluation process	Availability of data Random referral to the programme
Barriers to and weaknesses of the evaluation process	NA
Governance of measurement	
Independence of evaluator	No
Evaluator	Statistics and Business Intelligence Unit of the DEASP Directorate for Employment, Labour and Social Affairs of the Organisation for Economic Co-operation and Development (OECD)

1.9.4. Effectiveness

Achievement of outcomes

As mentioned in the background section, the **outcomes achieved exceeded expectations**. Out of the 65,868 people who started JobPath in 2016, 11,609 transitioned into employment and were still receiving support (JobPath phase 2). In 2017, 74,702 persons started the programme and 41,552 transitioned into the second phase³¹³. This translates into an employment rate among JobPath participants of 38% over the first two years of the programme.

Limited information is available to explain why JobPath was successful. The evaluator interviewed pointed out that the long-term unemployed people in the control group could access employment services on a voluntary basis, whereas JobPath was mandatory for the referred clients. It is possible, therefore, that relatively fewer participants in the control group accessed existing employment services such as Intreo and LES compared to JobPath.

Concerning the equivalent TF schemes, Intreo appeared to work for the short-term unemployed, but not the long-term unemployed. Compared with control areas in which the Intreo reform had not yet been rolled out, “Intreo reforms had a small but statistically significant positive impact on the probability of having exited the Live Register at three months, six months and nine months after the individual’s first interaction with an Intreo

³¹³ DEASP 2019, 20-22.

centre [...] . After 12 months, however, there was no statistically significant difference in the probability of having exited the Live Register.”³¹⁴

In LES, targets were set as follows: for the years 2013-2015, the target was that 50% of activation or non-activation clients should either find jobs or move onto a training programme. In 2016, this target was revised to 30% of clients entering employment placement, where placement is defined as employment of 30+ hrs per week or self-employment of 30+ hrs per week.³¹⁵ The programme did not achieve these outcomes, and the majority of service providers disagreed that outcome targets were appropriate (see the table below). They cited challenges in meeting the needs of different types of clients with multiple problems, and mentioned that targets should include all progressions made by clients, not just the 30-hour placements.³¹⁶

Table 41. LES Employment Placement Rate by Client type, 2013-2016

	2013	2014	2015	2016
Activation Client Placement Rate*	31.5%	35.3%	37.1%	28.8%
Non-Activation Client Placement Rate*	43.1%	47.2%	51.3%	27.9%
Target Rate	50%	50%	50%	30%

Source: Indecon 2018, 55.

Note: The placement/progression rates are based on the target outcomes as defined in the respective year. Over the period 2013-2015, the rate indicates the average share of clients who attended an interview that were placed into employment or self-employment or progressed into an education or training programmes. For 2016, the rate indicates the share of clients that are case-closed who were placed into employment or self-employment of 30+ hours per week. Because the target definitions changed from 2016, the 2016 target rate and placement rates are not comparable to the previous years.

Benefits and drawbacks of SOC compared with TF

Stakeholders noted that it is difficult to compare JobPath with other employment services in Ireland that are funded through traditional contracts. However, the results of the impact evaluation of JobPath give some insights into the added value and benefits of SOC.

The evaluation compared the performance of the JobPath treatment group with the control group (which consists of similar long-term unemployed persons on the Live Register who were not referred to JobPath)³¹⁷. The results, based on referrals in 2016 and 2017, show that **JobPath recipients were better off than those not participating in the programme** across different groups of participants.

According to the evaluation report:³¹⁸

- People who benefitted from JobPath in 2016 acquired 20% more jobs the following year than they would have done without JobPath, and 26% more jobs in 2018.
- People who did acquire jobs earned 16% more per week in 2017 and 17% more in 2018 if they had benefitted from JobPath in 2016.
- This means that, on average, people who benefitted from JobPath in 2016 had earnings from employment that were 35% higher than they would have earned without the programme in 2017, and 37% higher in 2018.

³¹⁴ Lavelle and Callaghan 2018, 32.

³¹⁵ Indecon 2018, 54.

³¹⁶ Indecon 2018, 128.

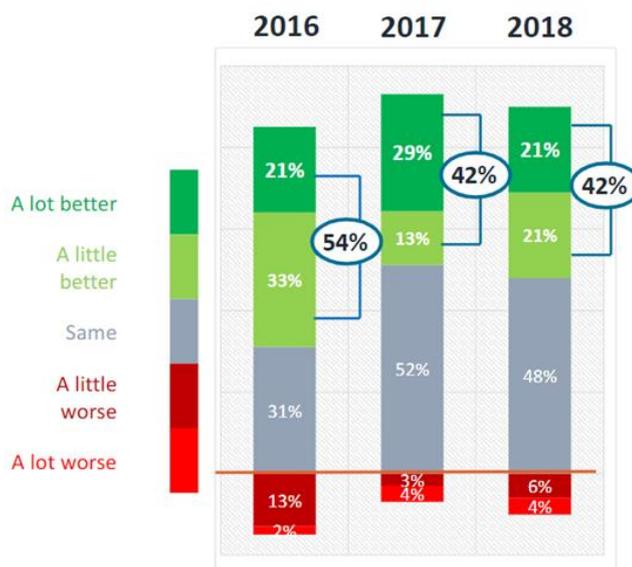
³¹⁷ The comparison involves a number of statistical adjustments, explained in Chapter 3.

³¹⁸ DEASP 2019, vii.

- Furthermore, the effect was positive for all cohorts who received the JobPath service, including those furthest from active participation in the labour market.

Regarding **user satisfaction**, since 2016, the Department has carried out customer satisfaction surveys of Intreo and JobPath customers twice a year. These provide an insight into jobseekers' levels of satisfaction with JobPath offices, staff, services and processes, as well as jobseekers' views of JobPath in comparison to Intreo. According to the phone survey conducted in 2018, **42% of JobPath participants thought the service was better than Intreo; 48% thought they were of similar quality, and 10% thought JobPath was worse** (see the figure below).³¹⁹

Figure 4. JobPath services compared with Intreo centre/Branch office



Source: DEASP (2019).

The JobPath survey results published in 2017 show that:

- Jobseekers express a high level of satisfaction with the service (76-81% satisfaction) and a low level of dissatisfaction (5-8%).
- JobPath staff made them feel valued (90%+) and clients had a good relationship with their JobPath advisor (90%+).
- The JobPath service has improved their chances of getting a job (68-77%).
- Of the small minority who provided suggestions for improvement, the initiatives suggested tended to concentrate on provisions for wider variety of jobs, more suitable courses and better services for older people.³²⁰

As in other schemes, providers noted that JobPath provided greater **flexibility in designing services**. Nevertheless, since all changes to service provision were decided after consultation with the DEASP, the contract still allowed a certain level of control.

According to the commissioner who was interviewed, the main benefit of JobPath compared with LES is its **focus on job sustainment**. As far as he was aware, service recipients in LES did not receive any support once they found employment. Nevertheless, he also noted

319 DEASP (2019). Satisfaction with JobPath service providers, Survey Phone results. Retrieved from: <https://www.gov.ie/pdf/?file=https://assets.gov.ie/75060/61bdf8e2-63d3-4c50-8011-b236921885c4.pdf#page=null>

320 DEASP 2017, 22.

that if he were to design another SOC, he would make follow-up services reactive rather than proactive because, while some participants appreciated and benefitted from the in-work support, others did not require or want it.

Another benefit pointed out by the commissioner of JobPath is its **job-first approach**, meaning that all of its focus is on helping people find jobs rather than helping them to develop skills or providing other services. However, the commissioner noted that the **job-first approach works better for people who are more job-ready**, while for other users it might be useful to focus on job preparedness (engaging in training, building confidence, etc.) and providing more holistic support (finding adequate housing, engaging in education, etc.).

Scalability and replicability

The DEASP is considering the possibility of replicating the PbR model for other employment services. However, the commissioner interviewed noted that the JobPath model would not necessarily be replicated in exactly the same way because other services might require a different balance between TF payment components and outcomes-based components.

1.9.5. Efficiency

Intervention costs

In JobPath, costs can vary depending on whether participants acquire jobs and how long they sustain them for (see section 1.9.1), but the average cost per participant in 2015-2018 was EUR 780.³²¹ **According to the JobPath evaluation, “this average cost [per participant] compares favourably to costs of other forms of activation such as LES...and Intreo, although exact cost comparisons can be difficult to quantify, particularly for Intreo.”**³²² The JobPath commissioner interviewed noted that the EUR 780 figure is a lower-bound estimate because some participants (who do not find jobs) re-enter the programme, yet their costs are not added together since they are counted as separate participants. Furthermore, the cost per participant will rise as participants spend more time on the programme and more of them find jobs. Nevertheless, the commissioner expected cost per participant to remain lower in JobPath than in LES, although the difference might become more marginal over time.

However, cost per participant tells us little regarding which programme is more cost-effective, and comparing cost per job acquired between the two programmes is problematic because they measure employment outcomes differently. In JobPath, for example, after the initial referral fee, providers only receive payments when the person has been employed for at least 13 weeks. The cost for such an employment outcome during the years 2015-2018 is equivalent to EUR 3,406.³²³ By contrast, LES reports participants who find jobs, regardless of how long they are sustained. According to an evaluation of LES, the cost per job averaged EUR 2,314 between 2014 and 2016.³²⁴ It is possible that this cost might increase noticeably if the measurement criteria were the same as JobPath's.

³²¹ DEASP 2019, 17.

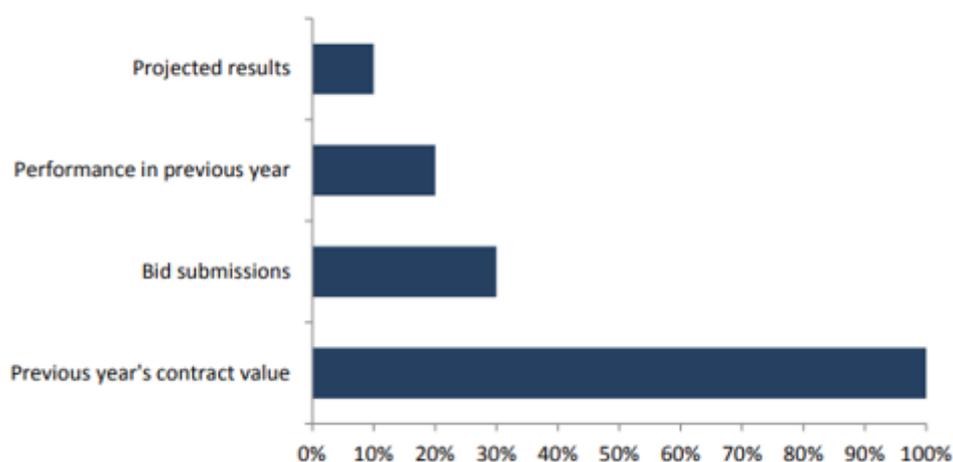
³²² DEASP 2019, 17.

³²³ The Office of the Comptroller and Auditor General (2018; 147, 142) shows that between July 2015 and March 2018, JobPath cost EUR 109 million and served roughly 160,000 participants, of which 20% found jobs (including both full-time and part-time). The quoted figure is based on this information.

³²⁴ More specifically, LES evaluator Indecon reports that in 2014 the average expenditure per job (counting both activation and non-activation clients) was EUR 1,981; in 2015, it was EUR 2,518; and in 2016, it was EUR 2,444 (Indecon 2018, 96).

The LES evaluation nevertheless argues that the programme would benefit from an outcomes focus. The evaluators showed that in the local DEASP divisions responsible for overseeing LES, there was a broad feeling that the decision whether or not to award a contract, and the value of the contract, were **not** significantly influenced by the ability of the service to meet contract targets.³²⁵ Research among contracted organisations indicates a similar view: all of the organisations interviewed believed that future contract values are based on the previous year's contract value, and only a few believe that it reflects an organisation's performance during the previous year (see the figure below). In fact, there have been instances in which DEASP continued to issue contracts even to those contractors who failed to meet existing contractual performance targets. The justification was that there were few other providers in the local area, and that sometimes contract targets might be overly ambitious.

Figure 5. Views of LES organisations' on the criteria for setting contract values



Source: Indecon 2018, 117.

Meanwhile, the JobPath commissioner interviewed believed that the JobPath programme still does not work for many people (those who are harder to help), at least in terms of acquiring jobs. Several times, the commissioner stressed that PbRs and TF do not represent a black-and-white picture. **There are many examples of mixed payment modes, and the commissioner suggested that contracts should be weighted more heavily towards the TF component for difficult-to-help groups, and further towards the PbR component for easier-to-help groups** (e.g. people who want to find jobs).

Operational costs

According to the JobPath commissioner interviewed, the **management costs for JobPath and TF programmes are different, and thus very difficult to compare**. The main costs involved in running JobPath (in addition to the cost of the intervention itself, which is paid out to the contractors) include salaries within DEASP, and the IT system used to coordinate with service providers. DEASP has the equivalent of 18 full-time staff working on JobPath, engaged in tasks such as the validation of outcomes and management of the contract. Five staff members work on outcome validation alone. Meanwhile, the equivalent of six full-time employees manage LES. However, many tasks that are performed centrally within the DEASC for JobPath are delegated to local-level government divisions for LES. The government also has to cover costs such as heating, rent, etc., for Intreo Centres, which

³²⁵ Indecon 2018, 117.

are not applicable to JobPath. Capital set-up costs for Intreo, for example, were estimated at EUR 40 million.³²⁶

1.9.6. Design of the scheme

Design features

Our research identified a number of design elements that were considered to make JobPath effective and/or efficient:

- The allocation of a sole provider per contract area was believed to be optimal for the design of JobPath in order to facilitate the creation of a new market, ensure economies of scale, and to render contracts more attractive to tenderers. It was also less demanding for the DEASC to manage only one provider in a given contract area. While competition might be appropriate in more densely populated areas, it may not be desirable in rural Ireland due to lower volumes of demand for employment services.³²⁷ However, the counter-argument to this set-up is that service users lose the freedom to choose providers, and providers are thus less accountable (which is why multiple providers were contracted per contract area in the UK's Work Programme, see Section 1.11). To address this counter-argument, the DEASP conducts semi-annual customer satisfaction surveys. If customer satisfaction drops, the DEASP can withhold up to 15% of the fees payable to a provider.³²⁸
- Another measure that ensured high-quality service was the Minimum Service Guarantee (MSG) (see Section 1.9.1). Since providers were contractually obliged as part of JobPath to provide a basic level of service to every client, it was not possible for providers to simply stop working with harder-to-help clients.³²⁹ However, one concern echoed by the service providers who were interviewed is that the MSG might undermine flexibility in returning the unemployed to work, and that overall service quality might suffer as a result.
- DEASP used a zero-based budget model for their commissioning process. This was in response to the PbR experiences in the UK (see Sections 1.10 and 1.11), during which tenderers submitted overly ambitious proposals for the price indicated, in order to win the bid. For JobPath, tenderers were required to determine the costs of delivery for each groups, as well as the performance levels that could be achieved at the stated prices. It was felt within the Department that providers would have a better idea of these costs. The main worry with the DEASP's approach was that the providers could get their calculations wrong, and that such an approach left potential for them to underbid for contracts – something that could have had disastrous implications for the programme's success, as the quality of service would have suffered and providers might ultimately have had to pull out. Nevertheless, these worries did not pan out, possibly because providers charged more than they felt was required to deliver the contract, due to the obvious degree of uncertainty.³³⁰
- The commissioner emphasised the importance of proper monitoring. For example, some JobPath participants at the beginning of the project were not motivated to look for work. Instead, they formed the impression that JobPath was a year-long

³²⁶ Murphy, M., Dulee-Kinsolving, A., Eustace, A. & Clarke, A (2016). *IDSS Country Study Ireland*, pp. 35-36. Retrieved from: http://mural.maynoothuniversity.ie/12063/1/MM_IDSS%20Country%20Study.pdf.

³²⁷ Lowe 2015, 120.

³²⁸ Wiggan 2015.

³²⁹ Lowe 2015, 122.

³³⁰ Lowe 2015, 123.

programme, so they were ready to engage with the programme for a whole year and only then look for work. This was observed both by service providers and the DEASC, who both worked to change this perception.

Although JobPath has so far delivered positive results, the commissioner interviewed noted a few elements of the programme that he would change if the contract were to be designed anew:

- The services provided after job acquisition could be made voluntary rather than mandatory (see Section 1.9.4).
- Irish employment services have developed an indicator to estimate **the probability of exit from employment**. The commissioner suggested that perhaps **payments (or the balance between TF and outcome-based components) could be determined on the basis of that indicator**. Administratively it would not be possible to tailor the payment structure on the basis of every individual; however, payment groups could be tailored on the basis of the groups of people targeted by the contract.

Impact on third-sector organisations and social enterprises

The position of NGOs appears to be more favourable in LES (TF) than in JobPath (PbR). This is because large contractors were invited to bid for JobPath contracts in order to withstand the programme's upfront costs, whereas LES are largely delivered by non-profits and charities (see Table 39). JobPath's Request for Tenders specified the following qualification criteria:³³¹

- the annual turnover of the prime contractor in each of the last three (3) audited financial years must be equal to or in excess of EUR 20 million (ex VAT) per annum;
- the tenderer must also have a clean audit opinion (not ongoing concern); and
- the tenderer must demonstrate the ability to finance the services.

With regard to technical and professional ability, Request for Tenders stated that:

- tenderers must have relevant experience;
- tenderers must have the necessary resources in terms of personnel.³³²

Nevertheless, smaller organisations do participate in JobPath as sub-contractors, as was envisaged in JobPath's Request for Tender:

"In line with Government policy the Department seeks to encourage participation by Small and Medium Enterprises (SMEs) in this procurement exercise. SMEs that believe that the scope of this procurement is beyond their technical or business capacity are encouraged to explore the possibilities of forming relationships with other SMEs or with larger enterprises. Through such relationships they can participate in and contribute to the successful implementation of any contracts, agreements or arrangements that result from this competition. Larger enterprises are also encouraged to consider the practical ways that SMEs can be involved in their proposals to maximise the social and economic benefits of the contracts that result from this public procurement competition."³³³

³³¹ DEASP (2015), 17-19.

³³² DEASP (2015), 20-21.

³³³ DEASP (2015), 5.

1.10. Provider-led Pathways to Work (the UK)

1.10.1. Background

Pathways to Work was an initiative aimed at improving labour market readiness and opportunities **for people claiming benefits for health reasons**. It was commissioned by the Department for Work and Pensions (DWP) and lasted between 2003 and 2011. Pathways was delivered by two types of service providers:

- Jobcentre Plus, and
- Contracted private and third-sector organisations.

Jobcentre Plus, a part of the DWP, is an employment agency with offices in various cities around the UK. In October 2003, the DWP piloted Pathways to Work in seven Jobcentre Plus areas, following which the programme was rolled out nationally. In rolling out Pathways, the DWP continued to deliver services through Jobcentre Plus across 39% of Great Britain, but in 2007 and 2008 it also contracted the delivery of services in the remaining 61% of the country to private and third-sector organisations through outcomes-based contracts.³³⁴ The contracted-out model is known as **Provider-led Pathways to Work** (PL Pathways) (as opposed to Jobcentre Plus-led Pathways to Work), and is one of the PbR schemes selected for this study. The DWP contracted the services out in order to foster innovation in service delivery and increase value for money.

Providers in PL Pathways were **contracted to deliver**:

- “Five, monthly mandatory Work Focused Interviews (WFIs) from the third to the eighth month of the [incapacity benefits] IB claim unless: the customer had been screened out or was exempt from these WFIs, or a WFI had been waived. (Jobcentre Plus advised if a waiver had been applied or the customer was screened out or exempt).
- Tailored, work-focused support (including better-off calculations and promotion of in-work benefits such as Return to Work Credit and tax credits, and other work-focused support.
- Referrals to Disability Employment Advisors and other specialist provisions (WORKSTEP, Access to Work, Residential Training College) as appropriate.
- A Condition Management Programme (CMP) that focused on at least the three main types of condition that give rise to the majority of IB claims, i.e., muscular-skeletal, cardiovascular, and mild to moderate mental health problems. The programme was required to conform to Department of Health Clinical Governance standards and data protection requirements and similar standards within Scotland and Wales.
- Prompt referral of the case to Jobcentre Plus for consideration of a benefit sanction where the customer has failed to attend or to participate. Jobcentre Plus would then establish if the customer had ‘good cause’ for non-participation and, if not, impose a sanction.”³³⁵

³³⁴ DWP. (2011). Jobcentre Plus Pathways to Work: Official Statistics. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/222922/jcppathways1011.pdf, 9.

³³⁵ Knight, G., Salis, S., Francavilla, F., Radu, D., Hevenstone, D., Mocca, E. & Tousley, B. (2013). Provider-led Pathways to Work: Net impacts on employment and benefits. Working Paper No 113. DWP. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/221223/WP113.pdf, 8.

In addition to these mandatory elements of the programme, providers could also provide additional services as they saw fit (the contract employed the so-called 'black-box' approach).

All new and repeat Incapacity Benefit (IB) claimants were required to participate in the mandatory elements of Pathways in order to continue receiving the benefit. Nevertheless, jobseekers who were not claiming health benefits could also join the programme voluntarily. Between 2007 and 2011, PL Pathways served 587,580 participants (444,600 on a mandatory basis, plus 142,980 who participated voluntarily).³³⁶

Service providers were **paid** in three ways:

- a service fee for taking people on to their caseloads (this service fee was worth 30% of the contract value and was paid in monthly instalments over the life of the contract);
- a job outcome payment when a client started work (these were one-off payments associated with successful outcomes and accounted for 50% of the contract value); and
- a sustained employment payment when a client maintained work for 26 weeks (a minimum of 16 hours a week for at least 13 of the previous 26 weeks. This component accounted for 20% of the contract value).³³⁷

Payments were made monthly, on the basis of targets or 'profiles' for job outcomes, sustained job outcomes, and the numbers of Work-Focused Interviews (WFIs) carried out. On average across service providers, contractors were expected to move one in three clients into employment (37%).³³⁸ The DWP agreed separate targets with contractors for voluntary Pathways participants, reflecting evidence that, on average, this group is closer to employment. Contractors were expected to move nearly one in two (47%) 'voluntary participants' into work.³³⁹

Despite a clear payment structure being laid out in the contract, actual payments were altered. This was mainly because the profiles specified in the contracts appeared overly ambitious, especially in the context of the 2008 recession, and providers ran into cash flow problems when they failed to meet them. As a result, the DWP brought forward the payment of service fees for nine out of 11 service providers, so that an additional GBP 24 million was paid in the first year of the three-year contract, reducing payments that were due later in the contract. Furthermore, the DWP removed restrictions on claiming outcome payments for voluntary participants.³⁴⁰

Both Provider-led and Jobcentre Plus-led Pathways to Work ended in 2011, when the initiative was replaced by the Work Programme (see Section 1.11). This replacement was partly motivated by **negative reviews** from the National Audit Office³⁴¹ and the Public Accounts Committee³⁴². According to Chris Grayling, the Minister in charge of the DWP at the time, Pathways to Work "was coming to an end anyway. It was due to be ended by the

³³⁶ DWP (2011), 14.

³³⁷ Knight et al. (2013), and Hudson, M., Phillips, J., Ray, K., Vegeris, S. & Davidson, R. (2010). The Influence of Outcome Based Contracting in Provider-led Pathways to Work, DWP Research Report 638, HMSO, Norwich.
<https://webarchive.nationalarchives.gov.uk/20130314010800/http://research.dwp.gov.uk/asd/asd5/rreports2009-2010/rrep638.pdf>, 8.

³³⁸ National Audit Office (NAO) (2010). Support to incapacity benefits claimants through Pathways to Work. Retrieved from:
<https://www.nao.org.uk/wp-content/uploads/2010/05/101121.pdf>, 26.

³³⁹ NAO (2010), 26.

³⁴⁰ NAO (2010)

³⁴¹ NAO (2010).

³⁴² House of Commons Committee of Public Accounts (2010). Support to incapacity benefits claimants through Pathways to Work. First Report of Session 2010–11. Retrieved from:
<https://publications.parliament.uk/pa/cm201011/cmselect/cmpubacc/404/404.pdf>.

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

previous Government. It [had] been given a pretty heavy thumbs-down by the National Audit Office (NAO) and by the Public Accounts Committee.”³⁴³ The NAO acknowledged that the Pathways programme had caused some new claimants to move off benefits sooner, but this was primarily because under Pathways, claimants received their medical assessment earlier. Some claimants failed the assessment, hence losing their benefit entitlement. There was little evidence that other parts of the programme were effective, or that the jobs acquired by participants were sustainable. Contractors also underperformed against the targets set, and job outcomes were achieved more often for those claimants who participated in the programme voluntarily (and hence were easier to help) than those for whom the programme was mandatory. Lastly, the Pathways programme was not financially sustainable for a third of contractors and two-thirds of subcontractors. The House of Commons Public Accounts Committee reiterated many of these points.

For a summary of stakeholders involved in PL Pathways, please see the table below.

Table 42. Summary of PL Pathways

Personally targeted social service	Disabilities Labour market-related
Target population	Recipients of health benefits, meaning people with physical or mental disabilities.
SOC scheme type	Mixed PbR
Payment model	Binary
Cost	Estimates vary. According to the NAO, the contracts for 2007-2011 had a total value of GBP 521 million ³⁴⁴ . According to the DWP, the total contract value for all Provider-led Pathways to Work contracts for 2007 – 2011 was GBP 469 million ³⁴⁵ .
Commissioner(s)	Department for Work and Pensions
Social service provider(s)	Eleven prime providers in total. Nine of these were private sector organisations; two were non-profit third-sector organisations. ³⁴⁶ Across prime contractors, there were 81 subcontractor organisations. Most sub-contractors were third-sector organisations. ³⁴⁷
Investor(s)	N/A
Intermediary	N/A
Evaluator(s)	Importantly, there were no evaluations of outcomes that were associated with payments. Service providers would report how many clients acquired and sustained jobs to the DWP. However, PL Pathways has been subject to multiple evaluations, impact assessments and progress checks. The Policy Studies Institute (PSI), a university research centre, carried out an impact study, while the National Centre for Social Research (NatCen) (a not-for-profit/charity) conducted survey fieldwork and produced a customer experience report. The qualitative research was carried out in collaboration between PSI, the Social Policy Research Unit (a university research centre) and NatCen. ³⁴⁸
Other parties involved	Jobcentre Plus staff, known as Third-Party Provision Managers (TPPMs), helped monitor the performance of providers.

³⁴³ House of Commons Work and Pensions Committee (2011). Work Programme: providers and contracting arrangements. Volume I: Report, together with formal minutes, oral and written evidence. Retrieved from: <https://publications.parliament.uk/pa/cm201012/cmselect/cmworpen/718/718i.pdf>, 53.

³⁴⁴ NAO (2010), 28.

³⁴⁵ Freedom of Information request submitted to DWP by PPMI (ref FOI2020/14648). Response received from the DWP Finance Group on 17 April 2020.

³⁴⁶ Knight et al. (2013).

³⁴⁷ NAO (2010), 29.

³⁴⁸ Knight et al. (2013), 2.

1.10.2. Equivalent TF scheme

Given that Jobcentre Plus delivered the same programme as PL Pathways but within the DWP, for the purposes of this study, we compare PL Pathways with Jobcentre Plus Pathways. Although providers were free to deliver any services to help participants find jobs in addition to the minimum programme components, cash flow issues among the providers meant that most delivered only the minimum services required.³⁴⁹ These were very similar to the services delivered by Jobcentre Plus, making the comparison valid, although the intensity of activities varied. For example, CMP was prescribed more frequently in Jobcentre Plus Pathways than in PL Pathways.³⁵⁰ Nevertheless, an impact assessment carried out as part of the Pathways to Work evaluation effectively compares the two programmes, providing useful insights for our study. For a summary of stakeholders involved in the two interventions, please see the table below.

Table 43. Comparability of PL Pathways (PbR) with Jobcentre plus Pathways (TF)

	Same as SOC	Comments
Personally targeted social service	Yes	Labour market-related Disabilities
Target population	Yes	Recipients of health-benefits, meaning people with physical or mental disabilities
Number of participants	Yes	SOC: In total, 587,580 participants took part in PL Pathways between 2007 and 2011, ³⁵¹ amounting to roughly 150,000 participants per year TF: 1,252,500 took part in the intervention from 2003 until 2011 ³⁵² , amounting to roughly 150,000 participants per year
Location	No	The PL and Jobcentre Plus models were rolled out across different parts of the UK
Cost	Yes/No	Comparable information is only available for 2008-09 SOC: the PL Pathways programme cost roughly GBP 117 million TF: GBP 130 million for Jobcentre Plus Pathways ³⁵³
Commissioner(s)	Yes	Department for Work and Pensions

³⁴⁹ Hudson et al. (2010), 45.

³⁵⁰ Nice, K., Davidson, J. & Sainsbury, R. (2009). Provider-led Pathways: Experiences and views of early implementation, DWP Research Report 595, HMSO, Norwich, 34.

³⁵¹ DWP (2011). Jobcentre Plus Pathways to Work: Official Statistics. Retrieved from:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/222922/jcppathways1011.pdf, 14.

³⁵² DWP (2011), 2.

³⁵³ NAO (2010), 12.

	Same as SOC	Comments
Social service provider(s)	No	PbR: mostly private for-profit organisations TF: Jobcentre Plus (public agency)

1.10.3. Outcomes measurement

The measurement of outcomes relating to payment was carried out by service providers. Aside from this, the scheme was also subject to a series of qualitative and quantitative academic evaluations, carried out by a consortium of universities and research centres (the Policy Studies Institute, the National Centre for Social Research, the PSI and NatCen).

Method used in the measurement of outcomes relating to payment mechanisms

Data concerning the achievement of the main outcomes upon which payment was conditioned were collected by the service providers. In order to prove sustained employment, the providers often had to call the employers for whom programme participants worked, and/or acquire their signature. This process was burdensome for the service providers and beneficiaries, and in some cases even discouraged employers from hiring programme participants.³⁵⁴

The DWP required service providers to retain documentary evidence of all jobs for which payments were claimed, undertaking sample checks on 10% of these.³⁵⁵ During these checks, DWP investigated the documentary evidence supplied by the providers.

Only at the end of the programme was a Pathways Evaluation Database created, which merged information from secondary sources including Jobcentre Plus's Labour Market System, the National Benefits Database, the Job Outcome Target database, the Pathways Screening Tool data, the New Deal for Disabled People (NDDP) Evaluation Database, and Return to Work Credit payments data.

The main indicators upon which payments were conditioned were:

- whether a client (benefit recipient) acquired a job;
- whether the client sustained employment for 26 weeks.

In relation to indicators and related payments, the programme set the main targets as follows: on average across all service providers, contractors were expected to move one in three clients into employment. Consequently, a 37% success rate was set as a target, but this target was not related to payments. Rather, it was used by contract managers to review service providers' performance.

Specific targets were set for the cohort of voluntary Pathways participants. For this group, the programme was expected to move into employment one out of every two voluntary participants. Thus, a success rate of 47% was deemed an appropriate target for these subjects.

³⁵⁴ Hudson et al. (2010), 44.

³⁵⁵ NAO (2010), 28.

Other evaluations

Several evaluations were also carried out that were not related to payments.

Qualitative studies – interviews

In total, four separate qualitative studies were carried out to evaluate PL Pathways. They were based on interviews with all stakeholders involved in the delivery of the intervention, as well as service recipients. Topics addressed included early implementation³⁵⁶, the influence of outcomes-based contracting on service delivery³⁵⁷, the experiences of service recipients³⁵⁸, and views on the Condition Management Programme.³⁵⁹

The qualitative research on the experiences of service recipients³⁶⁰ largely did not address the intervention's outcomes, but focused on participants' opinions regarding particular aspects of the programme, including how participants became aware of the programme, what they understood its purpose to be, how satisfied they were with the services provided, etc. Nevertheless, it did assess whether interviewees belong to one of the following groups:

- customers who were in full- or part-time paid work;
- customers who were thinking about paid work and/or had taken steps towards it;
- customers who were not thinking about paid work in the near future;
- customers who wanted paid work, but thought it to be an unlikely possibility.

The results from the interviews were analysed, taking into account which group each respondent belonged to.

Customer experiences – survey

Customer experiences were also assessed via telephone surveys (and an accompanying web-survey) of 3,095 new and repeat incapacity benefit customers in the Phase 1 areas of PL Pathways (December 2007), who claimed benefits between April and June 2008. Interviews were conducted during the period from June to mid-September 2009 (on average 14 months after their claim for benefits).³⁶¹

The survey focused on identifying customers' health and background characteristics, experiences of contacting Jobcentre Plus and individual providers' offices, the services provided to them and their assessment of these services, and respondents' work outcomes. The following outcomes were assessed:

- whether or not respondents were in paid employment;
- number of months (out of 13) spent in employment after a claim for Incapacity Benefit;
- number of work periods during the 13 months following a claim for Incapacity Benefit;

³⁵⁶ Nice et al. (2009).

³⁵⁷ Hudson et al. (2010).

³⁵⁸ Tennant, R., Kotecha, M. & Rahim, N. (2010). Provider-led Pathways: experiences and views of implementation in phase 2 districts, DWP Research Report 643, HMSO, Norwich.

³⁵⁹ Nice et al. (2009).

³⁶⁰ Tennant et al. (2010).

³⁶¹ Hayllar, O. and Wood, M. (2011) Provider-led Pathways to Work: the experiences of new and repeat customers in phase one areas, DWP Research Report 723, HMSO, Norwich.

- whether respondents worked as an employees, were self-employed, were looking for work or trying to start a business, or if they were not looking for work;
- number of hours worked each week;
- socio-economic classification of current work, based on the National Statistics socio-economic classification (from routine occupations to higher managerial and professional occupations); and
- the respondent's take-home pay.

Analysis using survey data also showed how some of these outcomes differed according to the respondent's health status and work history prior to claiming benefits. Furthermore, bivariate analysis was conducted in order to understand whether respondents' work outcomes depended on attending WFIs, assessment of services provided, demographic characteristics, qualifications, and problems with basic skills.

Impact assessment – differences in differences

The impact assessment of the programme³⁶² adopted a difference-in-differences methodology and compared the effectiveness of Jobcentre Plus-led model with the provider-led model. The outcomes measured included the **proportion of service recipients off benefits** and the **proportion of participants in employment**, using survey data as well as data from Her Majesty's Revenue and Customs (HMRC), the UK's tax authority. Jobcentre Plus Pathways pilot and expansion areas were used as comparison areas. Matching was undertaken with these comparison areas in order to select an appropriate comparison for each of the local Pathways areas. The matches identified were used for the survey and also to construct a comparison group based on administrative data. A number of pre-programme tests were performed to ensure the validity of the comparison.

Barriers to and enablers of the measurement process

The **high level of commitment and support given by the commissioning body** (the DWP) to the evaluation is considered an important enabling factor in the evaluation process. According to the evaluator, the DWP's support was essential in combining existing administrative data on employment and benefits to assess the impact of the programme, as this had not previously been done. Furthermore, the DWP took into account the evaluator's opinion when designing the intervention itself. Specifically, PL Pathways was rolled out in phases to enable a clear difference-in-differences comparison. Evaluation reports also mention overall support for the evaluators from service providers.

The main barriers to and weaknesses of the measurement process appear to be connected with the **unreliability of data** provided by the service providers and to the **lack of a monitoring and validating system** for the data provided. In fact, the same outcome reporting system had been subject to fraud in another employment programme in which one of the Pathways providers had participated.³⁶³ In relation to this, commissioners complained about the insufficiency of service providers' checks on data reliability.

For their part, providers complained that the outcomes measurement process hindered payments from the DWP, since the Department did not issue payments in cases where documents were considered incomplete. The providers also stressed the **administrative burden relating to data collection**, which took up time that could otherwise have been devoted to participants.

³⁶² Knight et al. (2013).

³⁶³ House of Commons Work and Pensions Committee (2010).

Table 44. Summary of the measurement process in PL Pathways

Measurement methodology	
Methods related to payment mechanisms	Self-reported data from service providers
Other evaluations	Qualitative studies – interviews and focus groups Survey on customer experiences Impact assessment – difference-in-difference
Methods using experimental or quasi-experimental design	Only in the impact assessment, which was based on a difference-in-differences approach (quasi-experimental)
Control group	Yes, in the impact assessment based on a difference-in-differences approach
Causality of impact	Causality of impact could not be attributed in the methodological approach that related outcomes to payments, but the impact assessment allowed causal inferences to be made in relation to the differences in outcomes achieved by PL Pathways and Jobcentre Plus Pathways.
Enabling factors and strengths of the evaluation process	<ul style="list-style-type: none"> • High level of commitment and support from commissioners Support for evaluations from service providers
Barriers to and weaknesses of the evaluation process	<ul style="list-style-type: none"> • Unreliability of data from the providers • Delays in data validation • Lack of independent control over data processing • Risk of fraudulent behaviour in the payment system • Administrative burden on the providers The range of outcomes was not wide enough to reflect the diversity of needs among the population with whom the service providers were assigned to work.
Governance of measurement	
Independence of evaluator	No, with regard to payment-related outcome verification; Yes, with regard to other evaluations
Evaluator	Service providers, for outcomes related to payment Policy Studies Institute (PSI) (university research centre) National Centre for Social Research (NatCen) (not-for-profit/charity) PSI, the Social Policy Research Unit and NatCen.

1.10.4. Effectiveness

Achievement of outcomes

Overall, service providers **failed by considerable margins to meet their contractual targets**. Actual job acquisition rates for claimants who were required to participate in Pathways ranged between 3 and 11% (June 2009), compared with the target of 37%.³⁶⁴

The main reasons for this were:

- a lower number of voluntary referrals than anticipated;
- the economic downturn, which resulted in greater competition for jobs;
- the competitive bidding process encouraged providers to set unrealistically high targets;
- administrative problems were also experienced by some providers; for example, difficulties gathering evidence of job outcomes, problems with administrative and payment flows.³⁶⁵

The contractors engaged in the PL Pathways scheme performed similarly to Jobcentre Plus.³⁶⁶ Based on the data for December 2007 to July 2009, 9% of participants in PL Pathways areas found jobs, compared with 11.1% of participants in Jobcentre Plus Pathways.³⁶⁷ Nevertheless, around 40% of all jobs achieved in PL Pathways areas were for claimants who have volunteered to participate, compared with around 9% in Jobcentre Plus Pathways areas. Volunteer participants are, arguably, easier to support because they are more likely to be motivated to take steps towards securing a job.³⁶⁸ Still, the Pathways evaluation found that employment effects were not statistically different between the two cohorts.³⁶⁹ There is also evidence that the rate of attendance was higher for PL Pathways, but that this was nevertheless associated with poorer work outcomes.³⁷⁰ The PL Pathways evaluation suggests that this was probably because providers **'parked'** some customer groups who were harder to help. For example, some customers would continue to be asked to attend meetings, but these meetings were cursory and did not lead to significant results.³⁷¹

Although differences in employment outcomes were not significantly different, PL Pathways providers outperformed Jobcentre Plus in terms of getting people off benefits. The impact of PL Pathways was to lower the proportion of benefit claimants by two percentage points (compared with the Jobcentre Plus model).³⁷² This was particularly true for people with mental health problems and other health conditions. The proportion of customers who used the health management services at PL Pathways was higher than those who used a similar Condition Management Programme in Jobcentre Plus.³⁷³ Moreover, the percentage of users (44%) who reported that the services delivered by Provider-led Pathways had helped them to think about work "a lot" was higher than the levels reported for Jobcentre Plus.³⁷⁴

³⁶⁴ NAO (2010), 26.

³⁶⁵ Hudson et al. (2010), 24.

³⁶⁶ House of Commons Committee of Public Accounts (2010), 3.

³⁶⁷ DWP (2011), 2.

³⁶⁸ NAO (2010), 9.

³⁶⁹ Knight et al. (2013), 4.

³⁷⁰ Knight et al. (2013), 52.

³⁷¹ Knight et al. (2013), 52-53.

³⁷² Knight et al. (2013), 4.

³⁷³ Knight et al. (2013), 5.

³⁷⁴ Hayllar and Wood (2011), 4.

Benefits and drawbacks of SOC compared with TF

The main **benefits** noted by the stakeholders of the Provider-led Pathways to Work PbR scheme relate to the fee structure, which was tied to job outcomes, because this was an efficient way to focus operations.³⁷⁵ Also positively assessed was the opportunity of subcontracting to other specialised providers.

With regard to user satisfaction, **the level of positive assessments among PL Pathways participants was higher than that revealed by the survey of customers in Jobcentre Plus areas** (44% of PL Pathways participants said the programme helped a lot in thinking about paid work in the future, compared with 30% in JobCentre Plus). However, when interpreting these findings it should be noted that the assessment is for all Work-Focused Interviews in Jobcentre Plus Pathways areas, and that a higher proportion of customers received services there than received services from providers.³⁷⁶

In terms of **drawbacks**, providing **long-term support** to those who were far away from the labour market was not a priority for the service providers, which focused more on helping people to acquire jobs:

“...I think given that the client group that we’re dealing with, a lot of them are those are going to need long-term support. I mean management do address that they will always ask who’s coming forward, those clients that are further away from work, what are you doing to move them forward, so it is addressed, but I think the very nature of it being outcome-based means that getting jobs on the board is always going to be a priority.”³⁷⁷

The quote above illustrates the issue of **creaming**. Advisers reported a huge management pressure to focus on job-ready clients, which led to less time being spent with clients who were further away from work.³⁷⁸ A strong sense of what needed to be done for business survival and job security meant that many advisers (and their senior managers) saw creaming as appropriate behaviour in a target-setting environment.³⁷⁹ Therefore, the scheme had problems in dealing with harder-to-help participants.

The focus on specific outcomes also meant that **other intermediary services** that did not lead directly to a customer’s employment were somewhat overlooked. Some prime providers recognised that there were other services that would benefit their customers, such as paying for training, financial management and debt advice, but they could not justify these costs because those services did not relate directly to job outcome targets.³⁸⁰

Another negative point concerns **outsourcing** more generally. According to the evaluator, PL Pathways emerged amid the broader context of outsourcing uncontroversial services such as food supply. The process of outsourcing social services was a political imperative at the time and, despite a lack of evidence that outcomes-based contracts resulted in better outcomes, unemployment services continued to be outsourced on a PbR basis due to political will. The evaluator argued that PL Pathways was not able to leverage existing relationships in the way that Jobcentre Plus could, for example, with the health sector. Furthermore, service outsourcing created disruptions in service provision, which confused service recipients as to where they should seek out services³⁸¹. It entailed a ‘handover step’,

³⁷⁵ Hudson et al. (2010), 40.

³⁷⁶ Becker, E., Hayllar, O. & Wood, M. (2010). Pathways to Work: programme engagement and work patterns. Research Report No 653. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/214423/rrep653.pdf, 38.

³⁷⁷ Hudson et al. (2010), 60.

³⁷⁸ Hudson et al. (2010), 56.

³⁷⁹ The issues of creaming and parking could be partly due to the fact that programme was mandatory for people who could not work. Based on advisers’ testimony, it seems that they were not comfortable ‘pushing’ these people to work even more, which resulted in spending less time with them.

³⁸⁰ Hudson et al. (2010), 43.

³⁸¹ Nice et al. (2009), 64; Hayllar, O. & Wood, M. (2011). Provider-led Pathways to Work: the experiences of new and repeat customers in phase one areas, DWP Research Summary, HMSO, Norwich. Retrieved from:

transferring users from Jobcentre Plus to private providers, which proved not to be very effective – and even problematic, in an environment of resource cutbacks within the agency.³⁸²

Third-sector organisations also criticised the outsourcing decision, arguing that it introduced high competitive pressures that affected the behaviour of Jobcentre Plus staff and **“squeezed out organisations that offer specialist and local knowledge that could be a lifeline to helping vulnerable people into employment, [...] prohibiting smaller organisations from bidding for contracts, and encouraging gaming within the system.”**³⁸³

Jobcentre Plus advisers expressed concern that providers’ **staff** were not adequately trained or experienced to work with the Pathways client group, particularly with regard to helping people with health problems, people who were difficult to engage in a mandatory programme, or people who were not job-ready. Some provider staff also lacked expertise in benefit and ‘better off’ calculations.

Lastly, service innovation remains one of the points debated. The PbR approach was expected to increase the innovativeness of services provided. Although a similar service was previously provided via Jobcentre Plus, PL Pathways providers had the option to provide whatever additional services they deemed necessary in order to get people into jobs (the ‘black box’ approach). Nevertheless, this approach did not provide the expected benefits to users. According to the commissioner interviewed, the private sector **was not able to achieve the high level of innovation** that had been hoped for. In practice, cash flow problems limited many providers to delivering only those services they were obliged to under the contract. When service providers offered additional services, these were often the same services in which the provider already had expertise. Nevertheless, some prime service providers subcontracted other organisations to provide specialist services, for example, for clients with particular health conditions (e.g. those who were visually or hearing impaired, or who had mental health conditions or learning disabilities) in which the prime service providers did not have sufficient in-house expertise.³⁸⁴

Respondents from various stakeholder groups pointed out that innovation is tied to extra resources and, in the economic climate during which PL Pathways took place, most prime providers were not in a position to spend on extra services for customers.³⁸⁵ However, one prime provider argued that passing on cases that required more specialised services to third parties was a service innovation, because third-party organisations had better knowledge as to how to help those cases³⁸⁶.

Nevertheless, the commissioner interviewed noted that even when outcomes-based payments increased under the subsequent Work Programme, providers still did not really innovate their services, but instead took on more participants who had a low chance of success. Thus, greater funding resulted in services for more people, but these services were not necessarily very innovative. Furthermore, for some groups with disabilities, such as people who had suffered a traumatic brain injury, there already existed at the time some established and effective rehabilitation programmes, but these tended to be expensive. Therefore, greater funding should not have been used on trying to further innovate services, but rather to provide services that had already been established as effective.

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/197262/723summ.pdf, 2.

³⁸² Interview quote.

³⁸³ The Centre for Social Justice (2020). Commissioning Excellence in Disability. Retrieved from: <https://www.centreforsocialjustice.org.uk/library/commissioning-excellence-in-disability>, 5.

³⁸⁴ Hudson et al. (2010), 19.

³⁸⁵ Hudson et al. (2010), 45.

³⁸⁶ Nice et al. (2009), 34.

Scalability and replicability

The Pathways to Work programme was initially implemented in seven districts by Jobcentre Plus, and was later scaled up across the UK using both the Jobcentre Plus model and the Provider-led model.

The main challenges associated with the national roll-out, according to the evaluator, were:

- building relationships among stakeholders;
- recruiting staff and appointing tasks.

1.10.5. Efficiency

Intervention costs

We could not compare the costs and benefits of the PL and Jobcentre Plus-led Pathways to Work schemes because the costs of the latter model were not made available to the research team. While a detailed cost-benefit analysis was conducted in the initial seven pilot areas,³⁸⁷ later assessment by the NAO concluded that the programme was not as effective when it was rolled out nationally,³⁸⁸ so it would be wrong to assume that the cost per job achieved remained the same after transitioning from pilot to national-scale programme. Nevertheless, the NAO conducted its own cost-benefit analysis. Though additional details were not provided, it concluded that Pathways as a whole, including both the Provider-Led and Jobcentre-plus-led models, had cost GBP 451 per participant ('programme start') or GBP 2,942 per job acquired for claimants starting Pathways up to the end of March 2009. The report states that "**jobs achieved through Provider-led Pathways are of similar cost to those delivered through Jobcentre Plus-led Pathways.**"³⁸⁹

Nevertheless, the commissioner noted that any comparison of costs between PL Pathways and Jobcentre Plus Pathways is subject to a number of **limitations**. In the contracted model, success is reflected by how much is spent on the programme. However, if the programme fails to attract many participants, or if it attracts many participants but fails to achieve payment-based outcomes, it will cost less. However, this is not because it was designed to cost less and not because it is more cost-effective. **In the TF model, the commissioner has more control:** the commissioner knows how many staff were recruited, how much was spent on training, etc. **In PbR models, the commissioner never knows how much is actually spent on the programme** (only how much the DWP allocates to the programme, which may potentially be substantially more than is actually spent). Furthermore, the costs of TF programmes are also not clear-cut. Local Jobcentre Plus offices had the opportunity to cross-subsidise from other programmes – for example, to recruit staff, if they thought that was appropriate. The staff recruited using Pathways funds might also end up working on other programmes. Lastly, costs for both programme ignore the price of 'free' services available to participants such as training or education, but these services are not free to the taxpayer. This is another example of how 'total programme costs' fail to capture the true costs of a programme.

Furthermore, the evaluator interviewed added that **many of the problems with creaming and parking are due to the supposed cost-effectiveness of outsourced services**. The evaluator further pointed out that Jobcentre Plus already had a monopoly in many areas in which they were delivered pre-defined services, so even in theory it is difficult to imagine

³⁸⁷ Adam, S., Bozio, A., Emmerson, C., Greenberg, D. & Knight, G. (2009). A cost-benefit analysis of Pathways to Work for new and repeat incapacity benefits claimants. Retrieved from: <https://www.ifs.org.uk/publications/4231>

³⁸⁸ NAO (2010).

³⁸⁹ NAO (2010), 9.

how new providers, with less experience in those areas, could find more cost-effective ways to deliver **services of the same or better quality**. The evaluator argued that new contractors did not have the skills to deliver (particularly in the short term), which was one of the reasons why the programme did not work well. The evaluator argued that the DWP continued to choose contracted PbR arrangements over in-house delivery for subsequent programmes due to theoretical rationalisations stating that contracted services must be more efficient, without any strong evidence that services could be delivered more cheaply while maintaining service quality.

Operational costs

Information on the operational costs of PL Pathways is limited, hence we cannot provide a side-by-side comparison of operational costs for PL Pathways and Jobcentre Plus Pathways. Nevertheless, in the discussion below we rely on the cost analysis conducted for the pilot areas of Jobcentre Plus Pathways only. We rely on interviews to infer whether the respective costs for PL Pathways are likely to have been lower or higher.

We estimate that the operational costs of Jobcentre Plus Pathways amounted to 12% of all programme costs (please see the table below for a detailed breakdown of what we classified as intervention and operational costs). Please note that information on the evaluation of the Pathways programme, as well as set up costs, were not available. Therefore, the 12% figure is a lower-bound estimate.

Table 45. Net costs (intervention and operational) for Jobcentre Plus Pathways per Incapacity Benefit enquiry (GBP) in seven pilot areas, April 2005 - March 2006

	Cost component	Cost (GBP)	% of total cost
Intervention costs	Total field staff salaries	68.99	88%
	Choices (NDDP; CMP; Accelerated Personal Capability Assessment)	68.69	
	Payments to participants (Return to Work Credit, Adviser Discretionary Fund, reimbursed expenses)	137.93	
Operational costs	Central administrative staff salaries	4.29	12%
	Field staff overhead	26.69	
	Central administration staff overhead	1.72	
	Field staff travel and support	1.45	
	Central administrative staff travel and support	0.56	
	Other field office costs (e.g. conferences, IT)	1.59	
	Other central administrative office costs	1.91	
Total		313.82	

Source: compiled by PPMI based on Adam et al. 2008, 44-47.

Notes: Adam et al. 2008 (p. 48) also transform the overall costs to market prices (accounting for VAT savings). We disregard these savings here because the available information would not be sufficient to break the costs into intervention and operational. Furthermore, costs are presented per Incapacity Benefit enquiry, even though not all people who enquired about the benefit ended up accessing it, and therefore not all were

enrolled into the Pathways programme. Programme costs per participant (presented under Intervention costs) are therefore higher.

While equivalent information for is not available PL Pathways, **interviews suggest that operational costs in the provider-led model were probably higher**. According to the commissioner, this is mostly because outsourced programmes require a greater level of oversight since private sector organisations have clear incentives to maximise profits and minimise costs, and because a greater auditing standard is applied before outcomes-based payments can be issued. To illustrate this, the DWP now uses administrative data to monitor Jobcentre Plus performance, assuming that everyone who stops claiming benefits finds a job, even though that is not always the case. By contrast, the DWP requires additional proof of job outcomes such as employment contracts before issuing payments to private providers. Providers have therefore developed their own systems to record when someone finds a job, and the DWP runs additional checks. In some cases, provider systems have been quite wasteful. For example, a private provider would contract a survey agency to track down programme participants, or the provider would provide a financial incentive for the participant to contact the provider again to provide proof of employment. Overall, the process of proving outcomes is more complex than monitoring Jobcentre Plus, and therefore entails greater costs.

1.10.6. Design of the scheme

Design features

Given that PL Pathways was one of the first large-scale PbR interventions in the UK, it offers many lessons learned about how to design these schemes more effectively.

- One lesson is the need to use administrative data to prove outcomes, rather than paper-based systems (see 1.10.3).
- All stakeholders emphasised the need for a longer contract duration. This would have made good economic sense on the commissioning side, requiring fewer staff resources to administer and manage. For providers, a longer contract would provide greater staff security, which would translate into greater continuity in services and organisational relationships. For example, the evaluator noted that short-term contracts do not provide enough time for prime providers to establish trustworthy relationships with their subcontractors, which can negatively impact the quality of the services and outcomes reported.³⁹⁰
- Evaluators noted the need for subcontractor performance monitoring and guidance to ensure high-quality delivery throughout the service supply chain. In fact, the DWP had little oversight over subcontractors, even though in some areas they delivered the bulk of the services. According to one contract manager, “where you’ve got a big subcontractor that delivers a big chunk of that contract then, you know, the standards between that subcontractor and [prime provider] should be very similar – I’m not sure that they [were] that similar. But I’ve got no evidence, because like I said we [didn’t] get into the nitty-gritty of subcontractor delivery.”³⁹¹ As a result, the NAO recommended that the DWP should conduct greater oversight over the supply chain and evaluate what impact its contracts have on subcontracted third-sector organisations, especially given that risk was disproportionately transferred to subcontractors.³⁹² Evaluators further noted that “if Pathways providers [were] to

³⁹⁰ Hudson et al. (2010), 42.

³⁹¹ Hudson et al. (2010), 31.

³⁹² NAO (2010).

continue subcontracting both work-focused interviews and service provision, it might be useful to supply providers and subcontractors with written guidance about acceptable standards regarding consistent practice, secure data transfers and accountability.”³⁹³

- On a related note, the NAO emphasised that the DWP should inspect the targets proposed by contractors with more rigour, to make sure they are not overly ambitious even if the primary risk involved in not meeting those targets lies with the contractor. In PL Pathways, when contractors failed to meet targets, continuity of service provision was at risk, because providers experienced a financial loss. Furthermore, the NAO recommended that if contractors fail to meet the targets, their contracts should be terminated rather than renegotiated, especially with respect to those contractors who perform the worst.³⁹⁴
- Providers noted the need to foresee in contracts – including those with subcontractors – what would happen if the volumes of referrals were much lower (or higher) than anticipated during programme design. When referrals turned out to be lower than expected, the Pathways programme became no longer financially viable for a number of providers (see the following sub-section).
- Another lesson learned was the need to consider whether PbR schemes are appropriate for mandatory client engagement. In their bids for PL Pathways, providers based their performance expectations on their experiences of working on programmes in which participants engaged voluntarily. Therefore, performance expectations were unreasonably high when applied to participants who were required to participate in Pathways, but who were not willing or able to start working, which is why most providers subsequently failed to meet those expectations. Although the subsequent PbR programme – the Work Programme (Section 1.11) – was also mandatory for some participants, only voluntary clients were invited to participate in the current Work and Health Programme. Part of the justification for engaging only voluntary clients is that providers should not be penalised for failing to achieve job outcomes for people who are not in a position to work. For more information, please see Section 1.11. On the other hand, engaging only voluntary clients might make it more difficult to predict participant volumes, which in turn might place more strain on the providers to make PbR schemes financially viable.
- Lastly, stakeholders gave mixed feedback regarding the ‘black box’ approach. On the one hand, the black box approach provided flexibility for providers to test out new service approaches. Nevertheless, providers argued that PL Pathways was still too strict in terms of the service requirements it prescribed (see Section 1.10.1), leaving little room for innovation. The recession, coupled with limited funding (due to failure to meet outcomes), made it virtually impossible for providers to offer any additional services beyond those minimally prescribed. Commissioners, on the other hand, felt that the black box approach did not allow for sufficient insights into what services work best. As a result, the commissioner interviewed stated that “the ideal is probably neither the black box, nor the grey box, but the glass box.” Furthermore, the black box failed to generate as much innovation as the commissioner expected. Asked whether more funding would have encouraged innovation even with the black box, the commissioner was sceptical. He noted that when outcomes-based payments increased under the subsequent Work Programme, providers did not innovate the service, but could instead afford to take on more participants who had low chances of success. So, greater funding resulted in services for more people, but these services were not necessarily very innovative. Furthermore, for some groups with disabilities, such as people who have suffered a traumatic brain injury,

³⁹³ Nice et al. (2009), 89.

³⁹⁴ NAO (2010), 13.

there were quite established and effective rehabilitation programmes, but these tended to be expensive. Therefore, the greater funding should not have been used on trying to further innovate the service, but to provide services that had already been established as effective. Overall, PL Pathways demonstrates a tension between service providers who push against prescriptive services for the sake of personalised support, and commissioners who want to ensure participants receive adequate services and programmes provide evidence of what services work best.

Impact on third-sector organisations and social enterprises

As mentioned in Table 42, nine out of 11 prime providers in PL Pathways were large for-profit organisations which then subcontracted all or part of the services to smaller NGOs. In general, the PL Pathways programme **was not financially viable** for a number of both for-profit and non-profit service providers. The economic downturn and resulting decline in job vacancies was felt to have exacerbated the financial risks associated with achieving customer employment targets. In that economic climate, the parameters of the contract were not considered feasible.³⁹⁵ Respondents on both the commissioning and the contracting side affirmed that providers had experienced a financial loss.³⁹⁶ The Shaw Trust, for example, the only major voluntary sector provider in Pathways to Work, lost almost GBP 3 million in the last financial year of its involvement in the DWP programme.³⁹⁷ Prime providers, and some partner agencies reported that their Pathways operations were being subsidised with revenue from other service contracts and funding streams.³⁹⁸ More than half of contractors (service providers) said they would not participate in Pathways again under the same terms.³⁹⁹

The financial risks were disproportionately felt by subcontracted NGOs. In PL Pathways, 70% of payments were conditioned on outcomes (see Section 1.10.1), so the risk was too great for small NGOs to participate as prime providers, and they had less power to negotiate more favourable agreements with prime providers. For example, prime providers chose how many and what type of clients to refer to subcontractors. As a result, some subcontractors argued that prime providers cherry picked clients and only referred those participants who were harder-to-help.⁴⁰⁰ Ultimately, two-thirds of subcontractors surveyed by the NAO expected to make a financial loss on Pathways, compared with one-third of prime providers. This illustrates the disproportionate risk carried by the third-sector NGOs who participated in Pathways.⁴⁰¹

1.11. The Work Programme (UK)

1.11.1. Background

The Work Programme (WP) followed two other outcomes-based contracts providing **employment services** in the UK. These were the Flexible New Deal (FND) and the Pathways to Work programme (see previous chapter). According to the evaluation of the WP, “in the summer of 2010, the Coalition Government announced that **in order to seek improvements in programme performance**, both the FND and Pathways to Work would

³⁹⁵ Hudson et al. (2010), 3.

³⁹⁶ Hudson et al. (2010), 40-41.

³⁹⁷ Benefits and Work (2009). Shaw Trust Loses Millions on Pathways. Retrieved from: <https://www.benefitsandwork.co.uk/news/1100-shaw-trust-loses-millions-on-pathways>

³⁹⁸ Hudson et al. (2010), 40-41.

³⁹⁹ NAO (2010), 10.

⁴⁰⁰ Hudson et al. (2010), 64.

⁴⁰¹ NAO (2010), 11.

be replaced by the Work Programme.”⁴⁰² According to the evaluator interviewed, the political climate also had a strong influence on why the PbR models continued to be chosen to deliver work-related services in the UK, despite a lack of clear evidence that they resulted in better outcomes for service recipients than traditionally financed services. Outcomes-based contracts emulated practices coming from Australia, and were considered to suggest effective incentives. Delivery of services by the public sector was seen as not being cost-efficient in comparison to the private sector – which, according to prevailing narrative, could innovate in a way that the public sector could not. The key new elements introduced into the WP compared with Pathways included higher performance expectations, the presence of multiple providers within a single contract area, a different outcomes-based payment model, longer contract duration, and less prescription to service providers in terms of the services that must be provided. The target population was also different: whereas Pathways to Work targeted those claiming benefits for health reasons, the WP substantially expanded this scope to include most unemployed adults (more specifically, those claiming Jobseeker’s Allowance [JSA] and Employment and Support Allowance [ESA]).⁴⁰³

These changes were a direct result of lessons learned from the Pathways to Work programme and FND. In Pathways, for example, contractors themselves estimated how many programme participants would find jobs and eventual targets were agreed with the DWP, based on the performance targets indicated in contractors’ bids. Nevertheless, according to the commissioner: “With so many different bidders for one contract, everyone is trying to outbid each other on numbers...so I don’t think it’s at all controversial to say that the vast majority of Pathways targets were never achievable”.⁴⁰⁴ By contrast, the WP set performance targets based on an analysis of historical data (for more details, see Section 1.11.3).

Furthermore, the decision was made to have multiple providers within a single contract area, in order to foster competition. More programme participants were referred to the providers with the best performance. Unlike in Pathways, payments (described in more detail below) were no longer issued when participants found jobs, but only after they had sustained them for a number of weeks, in order to incentivise sustainable employment and prevent deadweight⁴⁰⁵. Contracts were designed to last longer, to provide more stability for contractors and motivate them to invest in service innovation. To further encourage innovation, few activities were prescribed. This contrasted with both Pathways and FND.⁴⁰⁶ Finally, the target group was expanded substantially in order to simplify the process for benefit recipients: instead of navigating more than 20 different welfare-to-work programmes prior to 2011, participants could be referred to a single programme—the WP.⁴⁰⁷

In total, nine different groups of people could be referred to the programme for a maximum of two years⁴⁰⁸. Jobcentre Plus (the main employment agency in the UK, and part of the

⁴⁰² Lane, P., Foster, R. Gardiner, L., Lanceley, L. & Purvis, A. (2013b). Work Programme Evaluation: Procurement, supply chains and implementation of the commissioning model. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/197710/rrep832.pdf, 8 (emphasis added).

⁴⁰³ Lane et al. (2013b).

⁴⁰⁴ Hudson et al. (2010), 24.

⁴⁰⁵ Deadweight is the extent to which job outcomes would have been achieved without programme intervention.

⁴⁰⁶ Lane et al. (2013b), 8-9.

⁴⁰⁷ DWP (2012). The Work Programme: the First Year. Retrieved from:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/49882/work-programme-first-year.pdf, 2.

⁴⁰⁸ These groups included: 1. JSA claimants aged 18-24; 2. JSA claimants aged 25 and over; 3. JSA Early Entrants; 4. JSA ex-Incapacity Benefit (IB) claimants; 5. ESA Volunteers; 6. New ESA Customers (a. excluding 12 month prognosis claimants, and b 12 month prognosis claimants only); 7. ESA ex-IB claimants; 8. Incapacity Benefit/Income Support Volunteers; and 9. JSA Prison Leavers. DWP (2018). Work Programme National Statistics background information note. Retrieved from: <https://www.gov.uk/government/publications/work-programme-official-statistics-background-information-note/work-programme-national-statistics-background-information-note>

DWP) staff would refer unemployed individuals to WP service providers after they had been claiming JSA or ESA for a minimum amount of time. This minimum period varied, depending on the group the claimant belonged to, up to a maximum of 12 months. For example, those aged 18-24 were referred to the programme after nine months of claiming JSA, whereas prison leavers were immediately referred to the WP if they claimed JSA within three months of leaving prison. Referrals could be either mandatory or voluntary for different groups of people (see the table below).⁴⁰⁹

Table 46. Work Programme participant groups

Group	Time of Referral	Basis for referral
JSA claimants aged 18-24	From 9 months	Mandatory
JSA claimants aged 25 and over	From 12 months	Mandatory
JSA claimants who have recently moved from Incapacity Benefit	From 3 months	Mandatory
JSA claimants facing significant disadvantage (e.g. young people with significant barriers, ex-armed services personnel)	From 3 months	Mandatory or voluntary depending on circumstance
ESA claimants who are unlikely to be fit for work in the short term	When Work Capability Assessment is concluded	Voluntary
ESA claimants who are expected to be fit for work within 3 to 12 months [ESA Flow]	When Work Capability Assessment is concluded	Mandatory
ESA claimants who have recently moved from Incapacity Benefit	- Any time if expected to be fit for work within 3 to 12 months	Mandatory
	- Any time if not expected to be fit for work in short term	Voluntary
Incapacity Benefit and Income Support (In England only)	Any time	Voluntary
Prison leavers claiming JSA	Immediately, once benefit claim made within 3 months from leaving prison	Mandatory

Source: Dar 2016, 4.

With regard to **activities**, providers had the freedom to introduce and implement their own ideas and schemes to help unemployed participants find work.⁴¹⁰

According to the programme evaluation, “the **procurement** of the Work Programme...was a two-stage process where potential providers first bid to join DWP’s Employment-Related Support Services Framework and then took part in ‘mini-competitions’ for Work Programme delivery within 18 contract package areas (CPAs)... The mini-competitions attracted 177 bids, with between nine and 17 bids in each CPA. Thirty of the 35 framework providers bid, 18 of which were successful. Many of these providers were successful in more than one CPA and there are two or three ‘prime contractors’ or ‘primes’ in each of the CPAs”⁴¹¹.

⁴⁰⁹ Dar, Aliyah (2016). Work Programme: background and statistics. Briefing Paper Number 6340, House of Commons Library, 4.

⁴¹⁰ Dar (2016), 3.

⁴¹¹ Lane et al. (2013b), 14.

Large organisations were chosen to serve as prime providers, so that they would have sufficient upfront capital to deliver the services until programme participants found jobs – on the basis of which, providers would receive outcomes-based payments. Prime providers also subcontracted various organisations to deliver the services. Roughly half of organisations involved in service delivery were small (fewer than 50 employees); one-fifth were medium sized (50 to 250 employees); and the remaining (slightly less than one-third) were large organisations (250+ employees). Smaller organisations were more likely to provide specialist services (mental health, disabilities, other health issues, ex-offender services, etc.) whereas larger ones delivered end-to-end services. Almost half of all providers were voluntary, community, or social enterprise organisations; one-fifth were public sector (which also tended to be larger); and almost one-third were private⁴¹².

For a summary of the various parties involved in the programme, please see the table below.

Table 47. Summary of the Work Programme

Personally targeted social service	Labour market-related Disabilities Social exclusion
Target population	Various groups of unemployed persons (see Table 47)
SOC scheme type	Mixed PbR
Payment model	Binary, Frequency and Hybrid
Cost	According to the commissioner interviewed, the cost totalled roughly GBP 3 billion between 2011 and 2019.
Commissioner(s)	Department for Work and Pensions (DWP)
Social service provider(s)	18 private, public and voluntary-sector organisations. For a full list of providers and the areas they cover, please see the footnote ⁴¹³ . Prime providers subcontracted various organisations to deliver the services.
Investor(s)	N/A
Intermediary	N/A
Evaluator(s)	Payment-related outcomes were not evaluated externally (service providers claimed payments for each employed individual and relied on a paper-based trail of documents to prove this). The DWP verified the claimed outcomes using off-benefit checks (see Section 1.11.3). Other external evaluations were conducted by the Institute for Employment Studies (not-for-profit research centre) and the National Institute of Economic and Social Research (not-for-profit research centre).
Other parties involved	Jobcentre Plus European Social Fund provided funding for volunteer Income Support (IS) and Incapacity Benefit (IB) recipients to participate in the Work Programme until March 2015. ⁴¹⁴

With regard to **payment arrangements**, there were four payment elements made to WP prime contractors:

⁴¹² Foster, S., Metcalf, H., Purvis, A., Lanceley, L., Foster, R., Lane, P., Tufekci, L., Rolfe, H., Newton, B., Bertram, C. & Garlick, M. (2014b). Work Programme evaluation: Operation of the commissioning model, finance and programme delivery. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/425301/rr893-report.pdf, 71-72.

⁴¹³ DWP. (2013). Work Programme contract package areas and providers. Guidance. Retrieved from: <https://www.gov.uk/government/publications/work-programme-contract-package-area-and-prime-providers/work-programme-contract-package-area-and-providers>

⁴¹⁴ DWP 2018b.

- “An attachment payment for each individual participant. This [was] paid when an individual referral to the Work Programme provider [resulted] in a successful ‘attachment’, usually triggered in the first meeting with an adviser. The attachment fee [diminished] over the duration of the contract and was reduced to nil at the start of the fourth year of the contract. [Starting in 2014], the programme [was] solely funded by outcome payments.
- A job outcome payment for each individual successfully placed in a job. [Normally, payments would be made after an individual was employed for six months, or three months for harder-to-help groups]. Job outcome payments [were] only paid once for a participant over a two-year period. No payment [was] made for an initial ‘job entry’.
- A sustainment payment for each individual successfully retained in employment. This [was] paid every four weeks for keeping a participant in employment after a job outcome payment [had] been made. The maximum number of sustainment payments [differed] between payment groups, with up to 26 sustainment payments possible for those facing the most complex barriers to work (‘harder-to-help’ groups).
- An incentive payment: This flat rate fee [was supposed to] be paid only for jobs sustained by Jobseeker’s Allowance (JSA) participants above a given performance level, defined by DWP as 30% above the non-intervention rate (NIR), where the NIR is the number of participants who would have found employment without assistance from the Work Programme.”⁴¹⁵

For different participant groups, the DWP specified different levels of payment. These are outlined in the table below.

Table 48. Outcomes payments for specific groups of Work Programme participants (GBP)

Payment group	Max year 1 attachment fee	Max year 1 job outcome fee	Sustainment fees*	Total
JSA aged 18-24	400	1,200	2,210	3,810
JSA aged 25+	400	1,200	2,795	4,395
JSA Early Access	400	1,200	5,000	6,600
JSA Ex-IB	400	1,200	5,000	6,600
ESA Volunteers	400	1,200	2,300	3,700
New ESA claimants	600	1,200	4,700	6,500
ESA Ex-IB	600	3,500	9,620	13,720
IB/IS	400	1,000	2,300	3,700
JSA Prison leavers	300	1,200	4,000	5,500

*Paid monthly for varying periods, with the longest payment periods targeted at mandated clients receiving ESA.

Source: Finn, D. (2013). 6-7.

⁴¹⁵ Foster et al. 2014b, 125.

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

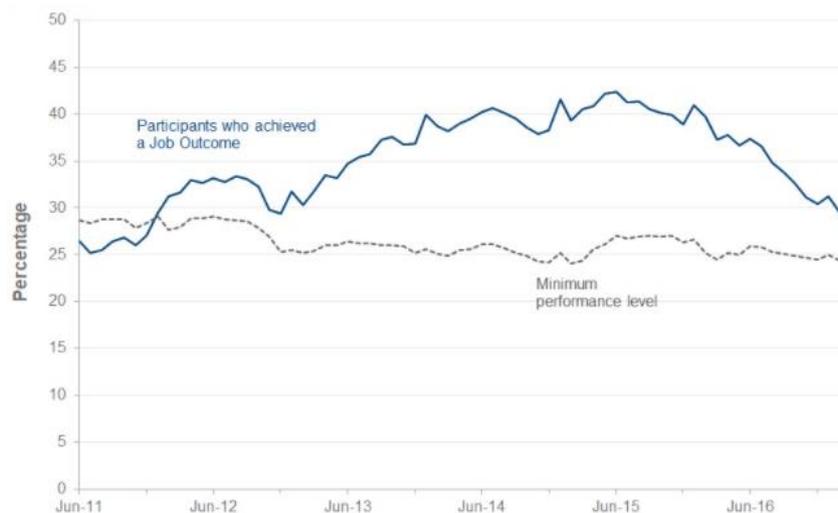
Note: JSA = Jobseekers Allowance; ESA = Employment Support Allowance; IB = Incapacity Benefit

In response to rising unemployment and revised forecasts from the Office for Budget Responsibility, six months after the Work Programme went live the DWP released new estimates in which referral volumes increased from 2.5 to 3.3 million. These referral volumes – far higher than anticipated – presented a much greater challenge for providers of the Work Programme.

In late 2014, the DWP renegotiated its contracts with prime contractors in light of the fact that most providers were failing to meet performance expectations. The DWP admitted that expectations – particularly those for ESA claimants – had been set too high, and subsequently lowered them.⁴¹⁶ As a result of these changes, the Department brought in a new performance measure based on the actual number of referrals, against which providers' performance was assessed.⁴¹⁷ For target and actual performance, please see Figure 6.

Between 2011 and 2017 (when referrals to the programme stopped), 1.9 million people were referred to the programme.⁴¹⁸ Of these, roughly 610,000 spent at least three/six months in work (the duration measured depended on the group participants belonged in)⁴¹⁹. While service providers struggled to meet performance expectations at the beginning of the programme, performance levels were exceeded after targets were adjusted (see the figure below).

Figure 6. Percentage of referral intakes that achieved a job outcome, 2011-2017



Source: DWP 2020.

Although the results delivered by the WP were good overall, however, the programme was not very effective for those groups that were harder-to-help. For example, please see 'Other ESA/IB' claimants in the figure below.

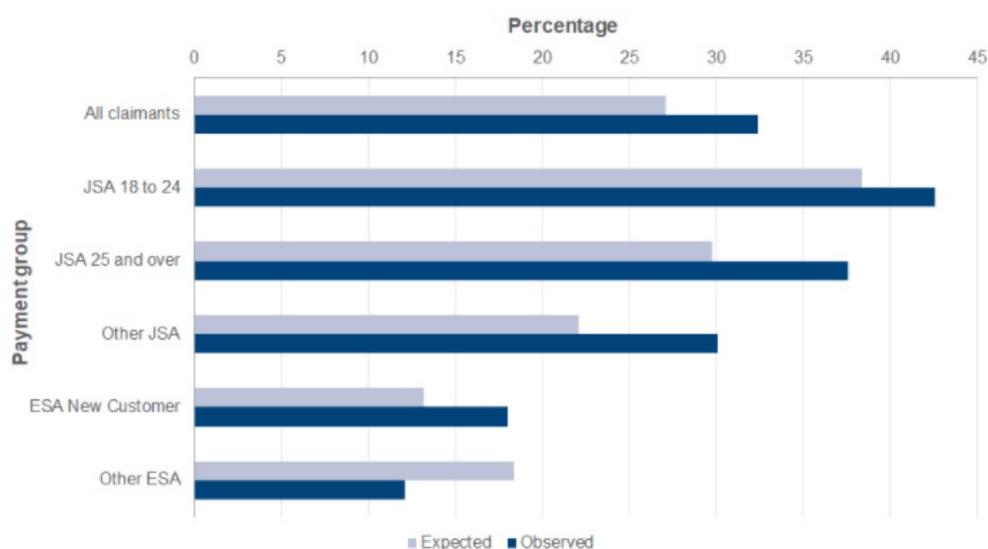
⁴¹⁶ NAO (2014). The Work Programme. Retrieved from: <https://www.nao.org.uk/wp-content/uploads/2014/07/The-work-programme.pdf>

⁴¹⁷ DWP 2018b.

⁴¹⁸ DWP (2017a). Annual Report and Accounts 2016-2017. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/624319/dwp-annual-report-and-accounts-2016-2017.pdf, 25.

⁴¹⁹ DWP. (2018a). Work Programme National Statistics. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/692566/work-programme-statistics-to-december-2017.pdf, 1.

Figure 7. Work Programme results by payment group, 2011-2020



Source: DWP 2020.

1.11.2. Equivalent TF scheme

The WP does not have a valid TF counterfactual, because the PbR model was rolled out across the entire UK. Nevertheless, according to interviewees, some comparisons can be made with the New Deal (ND) programme. The ND was a national UK employment programme that was in effect between 1998 and 2009. Although ND programmes originally targeted the young unemployed (18- to 24-year-olds), they eventually provided labour market activation measures for various subgroups who had a high propensity for being out of employment. These groups included those aged 25+ who had been unemployed for longer than 18 months, lone parents, partners of the unemployed, disabled people, those aged 50+ who were claiming various benefits for six months or more, and musicians claiming JSA for longer than six months. However, to ensure comparability with the WP, wherever possible we limit our comparison to two groups:

- New Deal for Young People (NDYP), which targeted youth (aged 18–24) unemployed for six months or longer.
- New Deal 25Plus (ND25pl), which targeted youth (aged 25+) unemployed for 18 months or more.

The ND programmes provided training, subsidised employment and voluntary work for the unemployed. These programmes also introduced the power to levy sanctions on benefit withdrawal for those who refused ‘genuine’ opportunities for employment (as in WP). ND programmes in the UK were operated by Jobcentre Plus, a network of government offices that administered cash benefits and offered employment services, as well as private and third-sector organisations.⁴²⁰

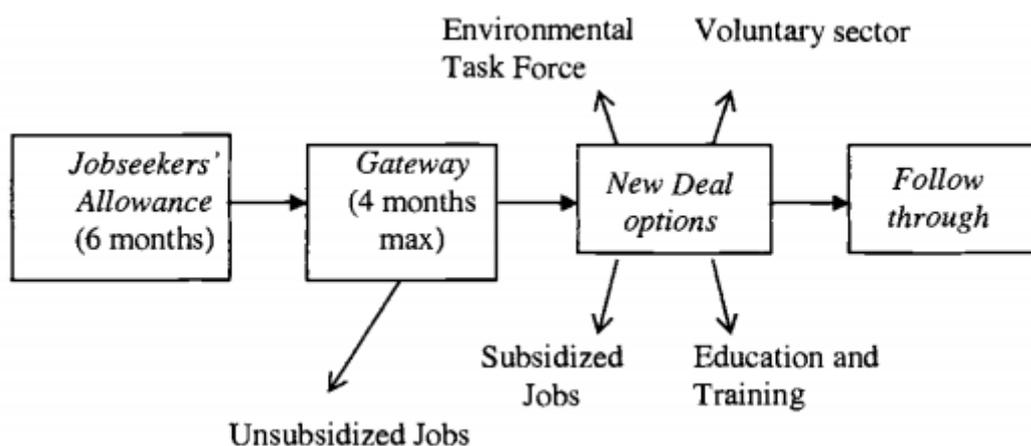
As the figure below shows⁴²¹, ND usually began with a period known as the Gateway. During the Gateway, participants received up to four months of intensive, personalised help and support, initially designed to help them find an unsubsidised job. If the participant did not acquire a job straight away, they would be directed towards one of four ‘ND Options’. The

⁴²⁰ Zaidi, A. (2009). Welfare-to-Work Programmes in the UK and Lessons for other Countries. *Policy brief. European Centre for social welfare policy and research, Vienna October.*

⁴²¹ Van Reenen, J. (2004). Active labor market policies and the British New Deal for the young unemployed in context. In *Seeking a premier economy: The economic effects of British economic reforms, 1980-2000.* University of Chicago Press, p. 15.

Options available were: 1) subsidised work; 2) full-time education and training; 3) work in the voluntary sector; or 4) work with the Environment Task Force. These Options typically lasted for six months, after which participants entered a period known as ‘Follow through’, which provided similar support to that available under the Gateway. NDYP and ND25pl were mandatory programmes, meaning that there was no option to not participate and continue claiming JSA.

Figure 8. The New Deal customer journey



Source: Van Reenen 2004, 15.

Both the ND and WP targeted unemployment and were large in scope. According to the commissioner, the main differences between the WP and the New Deals were that the WP was outcomes-focused, and involved more prime providers, personally tailored provision, with the ‘black box’ approach being more typical. There were no large differences between the two programmes in terms of target populations and activities delivered, although it is notable that people were referred to each programme after different periods of unemployment. Those aged 18-24 were referred to the WP after nine months in unemployment, compared with six months in NDYP; those aged 25+ entered the WP after 12 months of unemployment, compared with 18 months in the ND. These differences may or may not have impacted employment outcomes. Furthermore, the New Deal was in effect during the 2008 recession. Such different macroeconomic conditions also probably had an impact on how effective the two programmes were in helping people find work. For other aspects of the comparison between NDYP/ND25pl and the WP, please see the table below.

Table 49. Comparability of NDYP/ND25pl (TF) with the Work Programme (PbR)

	Same as SOC	Comments
Personally targeted social service	Yes	Labour market-related Disabilities
Target population	Yes/No	Overall, the WP targeted some groups of unemployed persons that were not targeted by the ND and vice versa. However, wherever possible, we limit the comparison to JSA claimants aged between 18 and 24, and those aged 25+. These two groups participated in both the ND and WP, although they were referred to these programmes after different durations in unemployment.
Number of participants	No	SOC: Between 2011 and 2017, 1,144,580 18 to 24-year-old JSA claimants and JSA claimants aged 25

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

	Same as SOC	Comments
		and over participated in the WP, amounting to roughly 160,000 participants per year. ⁴²² TF: From 1998 to 2010, the number of starts on the NDYP and ND25pl together was 3,825,930, which equates to roughly 300,000 participants per year. ⁴²³
Location	Yes	Both the WP and nds were rolled out nationally.
Cost	Unknown	Detailed breakdowns of costs by participant group are not available for the WP, so it is not possible to state whether or not expenditure was similar to that for the ND. However, expenditure on all WP groups amounted to roughly GBP 375 million per year, ⁴²⁴ compared with GBP 557 million ⁴²⁵ for all groups of ND.
Commissioner(s)	Yes	Department for Work and Pensions
Social service provider(s)	Yes/No	Whereas all WP services were contracted on a pbr basis, the nds were either delivered in-house (through employment services), contracted out through traditional financing, or delivered through both in-house delivery and contracted delivery, depending on the local area.

1.11.3. Outcomes measurement

Outcomes relating to payment mechanisms in the Work Programme were reported by the service providers and verified by the DWP using off-benefit checks, while other external evaluations were conducted by Employment Studies (a not-for-profit research centre) and the National Institute of Economic and Social Research (a not-for-profit research centre).⁴²⁶

Measurement of outcomes relating to payment mechanisms

Providers reported customer attachments (initial contacts with referred customers) and claimed Job Outcomes and Sustainment Payments through an electronic Provider Referral and Payment system (PRaP). The DWP did not prescribe the way in which the providers should track participants or record information about their employment. There was, however, specific information that the providers were required to input to PRaP when they submitted a claim. This included the contact details of the participant and their employer,

⁴²² Elaborated by PPMI, based on DWP (2018) Work Programme statistics, Table 1_2, available from: <https://www.gov.uk/government/statistics/work-programme-statistical-summary-data-to-december-2017>.

⁴²³ DWP (2011). *The numbers of starts on New Deal for Young People and New Deal 25plus, and the number of times they have started each programme*. DWP, pp. 2-3. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/223116/nd_repeater_s_2011.pdf

⁴²⁴ Derived by PPMI from the cost of the programme stated by the commissioner as being roughly GBP 3 billion between 2011 and 2019.

⁴²⁵ Between 1997 and 2006, GBP 3.5 billion was spent in total on the six New Deal programmes (Field, F., & White, P. (2007). *Welfare Isn't Working: The New Deal for Young People*, p. 11. Reform, London). We divide this number by nine (the number of years) and adjust for inflation between 2006 and 2019 to arrive at the GBP 557 million figure (assuming that inflation averaged 2.9% per year). Please note that without inflationary adjustment, the programme expenditure per year would amount to GBP 382 million.

⁴²⁶ DWP 2018b.

the participant's job title, employment start and end dates, number of hours worked, etc. This information had to be provided for every job the participant held during the claim period. Failure to provide all or part of this information could result in payment being recovered.⁴²⁷

Before issuing payments, the DWP checked the claimed outcomes against administrative benefits data to ensure that the same people were no longer claiming benefits. Furthermore, after the payments were issued, the DWP conducted post-payment validations. In essence, these involved drawing a random sample of job outcome claims and checking against HMRC data to see whether those participants were indeed employed (hence, a different dataset was used from that used in off-benefit checks). The rate of invalidated claims in the sample was extrapolated to all outcomes claimed, which would trigger back payments from the providers. Providers were given the opportunity to contest the findings and present evidence of employment before paying the DWP back.

There were two main outcomes upon which payments were conditioned: the beneficiary finding work and remaining in employment (the duration required for payments associated with sustained employment varied between different groups of the participants).

To prove these outcomes, providers had to enter the following information into PRaP regarding every programme participant:

- whether the period is in employment or self-employment (or a combination of both);
- the employer's contact details (including address, contact name, full business telephone number and email address);
- the participant's contact details;
- job title;
- job start date;
- job end date;
- the number of hours worked each week;
- the work/shift pattern;
- an employee identifier such as works or payroll number (optional).
- In order to claim sustained job outcomes, the same information above had to be provided. However, as proof of sustained employment the DWP required additional information, which included written records showing that the provider:
 - has established that the participant is in sustained employment;
 - has agreed on an on-going support mechanism with the participant and that this support mechanism remains available to the participant, regardless of whether they choose to access it;
 - has agreed with the participant that they are established in their job and that they do not need to further contact the provider unless circumstances change;
 - has agreed with the participant that they will notify the provider of any relevant change in their circumstances, for example if they require more support, start a different job, and/or leave their current job.

⁴²⁷ ESF. 2016. Work Programme Provider Guidance: Chapter 9 - Work Programme evidencing/validating payments.

Retrieved from:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/564883/work-programme-provider-guidance-chapter-9.pdf, 6.

In addition to the main indicators – the number of people who find jobs and sustain them – the DWP also calculated the expectations for what providers should deliver. These expectations were only applied to three of the nine groups of people who were eligible for the programme. These expectations were contractual minimum performance levels (MPLs). In other words, if the providers failed to achieve them, they could be subject to a formal performance improvement process.⁴²⁸ Specific targets (called End of Cohort Profiles) are specified below⁴²⁹:

- JSA claimants aged 18-24: **38.4%** (i.e. it was expected that contractors would get at least 38.4% of all JSA claimants aged 18-24 and referred to the contractor, into work during the 30 months for which participants are allowed to participate in the programme).
- JSA claimants aged 25 and over: **29.8%**.
- New ESA customers:
 - Excluding the 12-month prognosis claimants: **13.2%**;
 - 12-month prognosis claimants only: **6.6%**.

In addition to these rolling 12-month MPLs, the DWP also specified quarterly MPLs for the same groups. MLPs were set only for those groups because these were the three largest participant groups for whom viable estimates could be made from analysing the historical off-benefit flows and job entry rates.⁴³⁰

All expectations regarding performance were based on the analysis of historical data. The levels were set 10% above those that would be expected to occur in the absence of the Work Programme.⁴³¹

Other evaluations

Qualitative evaluation⁴³²

A qualitative analysis was carried out regarding programme delivery. The study drew on qualitative interviews with staff from 56 different Work Programme provider organisations, and with Jobcentre Plus staff across 12 districts; qualitative interviews with over 90 Work Programme participants; and observations of meetings between frontline staff in the Work Programme provider organisations and participants in the programme.

The evaluation focused on early programme implementation, so there was little focus on outcomes. Nevertheless, participants' views on various aspects of the programme were assessed. Furthermore, the evaluation specifically assessed how the programme had changed the aspirations and motivation of participants towards finding work, and the effort they had put into finding work prior to the Work Programme and during it.

⁴²⁸ DWP 2018b.

⁴²⁹ European Social Fund. 2014. Work Programme Provider Guidance: Chapter 12 – Performance Management. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/723312/work-programme-chapter-12.pdf, 4.

⁴³⁰ Foster et al. 2014b, 97.

⁴³¹ DWP 2018b.

⁴³² Newton, B, Meager, N., Bertram, C., Corden, A., George, A., Lalani, M., Metcalf, H., Rolfe, H., Sainsbury, R. & Weston, K. (2012). Work Programme Evaluation: Findings from the first phase of qualitative research on programme delivery. Research Report 821. Department for Work and Pensions.

Evaluation of Procurement, supply chains and the implementation of the commissioning model⁴³³

The Evaluation of Procurement, supply chains and the implementation of the commissioning model also relied on the qualitative analysis of interviews. Interviews were conducted with provider organisations outside of the Work Programme's supply chains, with the Work Programme providers, and with DWP and Jobcentre Plus staff. An online survey was also conducted addressing the subcontracted Work Programme providers.

Evaluation of Day One Access to the Work Programme for prison leavers and Employment and Reoffending Pilot⁴³⁴

The Evaluation of Day One Access to the Work Programme for prison leavers and Employment and Reoffending Pilot encompassed qualitative work with the providers, employers and individuals, as well as quantitative research including a survey of individuals and a management information assessment.

Both the interviews and the survey with prison leavers focused on their experiences of receiving services and their evaluation of those services. In addition, the survey assessed work outcomes and impact by asking about the following:

- whether respondents were in paid employment at the time of the survey;
- whether they were self-employed;
- whether they were working on a full-time or a part-time basis;
- how long they had been in employment;
- the type of contract held;
- how many hours the respondent worked;
- the level to which employment matched the respondent's experience, skills and interests;
- to what extent the Work Programme had helped in finding work;
- to what extent their existing employment offered the possibility of promotion or taking on more responsibility;
- to what extent their existing employment offered the possibility of increased pay;
- how the Work Programme had helped respondents after finding work;
- if respondents were engaged in any training activities.

Effectiveness of prime providers⁴³⁵

The evaluation of the effectiveness of the prime providers analysed factors associated with differences in the relative effectiveness of the prime providers. The analysis used a two-stage regression approach to relate provider impacts to a range of service indicators. An important limitation of this analysis is that it cannot provide causal estimates. That is, the

⁴³³ Lane, P., Foster, R., Gardiner, L., Lanceley, L. & Purvis, A. (2013b). Work Programme Evaluation: Procurement, supply chains and implementation of the commissioning model.

⁴³⁴ George, A., Metcalf, H., Hunter, G., Bertram, C., Newton, B., Skrine, O. & Turnbull, P. (2014). Evaluation of day one mandate of prison leavers to the Work Programme. Research Report No 897. DWP.

⁴³⁵ Dorsett, R. & Lucchino, P. (2016). The Work Programme: factors associated with differences in the relative effectiveness of prime providers. DWP ad hoc research report no. 26.

results should not be understood in such a way as to indicate that the differences in measured service type are responsible for the differences in prime providers' relative effectiveness.

Evaluation of participant experience⁴³⁶

The evaluation of the participant's experience entailed analyses from two waves of a large-scale longitudinal surveys of participants and a multi-wave (partly cross-sectional, partly longitudinal) programme of in-depth qualitative fieldwork with the participants.

The evaluation includes a separate chapter dedicated to programme outcomes, which focuses on the following:

- entry to employment: the proportions and characteristics of the Work Programme participants obtaining paid work;
- sustained employment of six months or longer, including the characteristics of those most likely to obtain this outcome;
- characteristics of those who did not find employment.

The evaluators also asked those who completed the programme whether the intervention had made any difference to them, even if they did not obtain jobs. The most often cited responses included gains in confidence or self-esteem, the ability to get short-listed for a job interview, learning more effective job-search techniques, gaining qualifications in English, maths and IT skills, or learning something new from the training activities they had engaged in.

Barriers to and enablers of the measurement process

Some providers felt that the job **targets for Employment and Support Allowance (ESA) customers** were unrealistic, and did not reflect their view as to how far away from the labour market many of these participants were. This was argued by providers across the spectrum, including disability specialists and prime providers.

From the providers' perspective, these issues were related to **how the DWP measured performance**. The DWP's MPL targets were problematic, as they were highly affected by referral patterns. The DWP's focus on performance in just three payment groups encouraged providers to prioritise support to these groups over others. Providers also criticised the outcomes chosen. A recurring criticism made by providers working with participants who were judged to be furthest away from the labour market, was that the Work Programme model did not give sufficient weight to '**soft outcomes**' (for instance, progress in areas such as confidence and motivation building). Some providers felt that they had to put in a lot of effort and resources in order to make significant improvements in the employability levels of these participants, but these efforts were not adequately rewarded due to the Work Programme's evaluation design.

Table 50. Summary of the measurement process in the Work Programme

Measurement methodology	
Methods relating to payment mechanisms	Self-reported data (verified by DWP using existing databases)

⁴³⁶ Meager, N, Newton, B., Sainsbury, R., Corden, A. & Irvine, A. (2014). Work Programme evaluation: the participant experience report.

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

Other evaluations	Qualitative evaluation Evaluation of Procurement, supply chains and implementation of the commissioning model Evaluation of Day One Access to the Work Programme for prison leavers and Employment and Reoffending Pilot Effectiveness of prime providers Evaluation of participant experience
Methods using experimental or quasi-experimental design	No
Control group	No
Causality of impact	The methodology adopted could not assess causality of impact
Enabling factors and strengths of the evaluation process	The availability of benefit and HMRC data allowed the outcomes claimed by service providers to be verified
Barriers to and weaknesses of the evaluation process	Choice of outcomes: according to providers, insufficient weight was given to 'soft outcomes' Outcomes were subject to referral volumes, which were highly variable
Governance of measurement	
Independence of evaluator	No
Evaluator	Department for Work and Pensions Institute for Employment Studies (not-for-profit research centre) National Institute of Economic and Social Research (not-for-profit research centre)

1.11.4. Effectiveness

Achievement of outcomes

Based on official statistics, after targets were downgraded (see Section 1.11.1), **providers exceeded minimum performance levels for most of the programme's duration in terms of helping people to find jobs.**⁴³⁷ Almost two million people were referred to the Work Programme during its lifetime. Of these, 630,000 participants achieved job outcomes, compared with an expected number of 530,000. This amounted to 32% of all claimants referred to the programme.⁴³⁸

However, outcomes varied substantially by group, with those **harder-to-help achieving worse results.**⁴³⁹ Personal characteristics made a difference to the likelihood of participants

⁴³⁷ DWP (2017b). Work Programme National Statistics. Data up to December 2017. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/692566/work-programme-statistics-to-december-2017.pdf

⁴³⁸ DWP (2020). Work Programme statistical summary: data to June 2020. Retrieved from: <https://www.gov.uk/government/publications/work-programme-statistical-summary-data-to-june-2020/work-programme-statistical-summary-data-to-june-2020>

⁴³⁹ DWP 2017b, 4.

finding and sustaining work while on the programme. In particular, multivariate statistical analysis (i.e. controlling for other factors) showed that **after the programme had been running for two years, participants' total duration of employment while on the programme was higher if they were female, young, did not have a disability or health condition, had recent work experience prior to joining the programme, and lived in an area with a less deprived local labour market.**⁴⁴⁰

On this point, the WP has been criticised by various scholars: for instance, the programme did not perform as well for participants with long-term health conditions as for those without. The actual proportion of **people with long-term health conditions** moving into work and sustaining it for a defined period of time (three or six months) remained low. Indeed, the Work Programme fell short of expectations regarding the employment outcomes achieved for those claiming Employment and Support Allowance (ESA). Just one in seven ESA claimants achieved a job outcome via the Work Programme, compared with one in four who were claiming Jobseeker's Allowance (JSA). For those individuals who shifted to ESA from Incapacity Benefit, outcomes were worse: only one in 20 obtained a job.⁴⁴¹

Performance also varied by provider: 31.9% of eligible participants referred to the highest-performing contract on this measure (Ingeus UK, East of England) achieved job outcomes, compared with 23.4% for the lowest performing contract (Rehab jobfit, Wales).⁴⁴² Furthermore, part-time and temporary jobs were much more common among Work Programme participants than among the UK workforce as a whole, but the proportion of self-employed was similar to the national average.

The government set the traditionally funded NDYP programme the target of getting 250,000 under-25s off benefit and into work in 1997. This goal was supposed to be achieved by 2002. In fact, the target was achieved in September 2000.⁴⁴³ The NAO report on the NDYP showed that, of those who completed the programme, 40% had gone into unsubsidised employment.⁴⁴⁴ **The results are similar to those of the Work Programme:** 42% of JSA claimants aged 18-24 achieved a job outcome, according to data up to March 2018.⁴⁴⁵

As for ND25pl, only 25% of participants moved from benefits into sustainable and unsubsidised jobs. Almost half (46%) of those who left the programme ended up back on welfare.⁴⁴⁶ For these participants, the Work Programme appears to have performed better: 37% of JSA claimants aged 25 and over achieved a job outcome, according to data up to March 2018.⁴⁴⁷ Nevertheless, it is important to remember that participants within this age cohort were referred to the New Deal after 18 months in unemployment, compared with 12 months under the Work Programme. Longer spells of unemployment may have reduced the participants' chances of finding the job under ND25pl, which would explain why the programme appears less effective for these participants.

⁴⁴⁰ Meager, N., Newton, B., Sainsbury, R., Corden, A. & Irvine, A. (2014). Work Programme evaluation: the participant experience report. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/425081/rr892.pdf, 25.

⁴⁴¹ Dudley, C., McEnhill, L. & Steadman, K. (2016). Is welfare to work, working well? Improving employment rates for people with disabilities and long-term conditions. The Work Foundation, Part of Lancaster University. Retrieved from: http://englishbulletin.adapt.it/wp-content/uploads/2016/06/405_Work-Programmes.pdf, 5-6.

⁴⁴² Dar 2016, 7

⁴⁴³ NAO 2002, 3.

⁴⁴⁴ National Audit Office. (2002). The New Deal for Young People, p. 10. Retrieved from: <https://www.nao.org.uk/wp-content/uploads/2002/02/0102639.pdf>

⁴⁴⁵ DWP (2018d). Work Programme statistics, Table 2_8. Retrieved from: <https://www.gov.uk/government/statistics/work-programme-statistical-summary-data-to-december-2017>.

⁴⁴⁶ PricewaterhouseCoopers International Limited (2004). New Deal 25+ Evaluation Report No. 9. PWC. Retrieved from: <https://www.communities-ni.gov.uk/sites/default/files/publications/del/new-deal-25plus-evaluation-report-no9.pdf>, 158.

⁴⁴⁷ DWP (2018d). Work Programme statistics, Table 2_8. Retrieved from: <https://www.gov.uk/government/statistics/work-programme-statistical-summary-data-to-december-2017>.

Benefits and drawbacks of SOC compared with TF

In terms of **user satisfaction**, beneficiaries of the Work Programme, the NDYP and the ND25pl who found jobs reported **very similar levels of job satisfaction**. Of those Work Programme participants who found work, 85% said their most recent job placement matched their interests and needs fairly well or very well;⁴⁴⁸ 88% of employed NDYP participants were fairly, very or completely satisfied with the job they found;⁴⁴⁹ and 85% of employed ND25pl participants were either very or quite satisfied with the work they found.⁴⁵⁰

However, **based on participant surveys, the Work Programme appears to have been somewhat more effective in helping beneficiaries find work**. Roughly two-thirds (64%) said the programme was helpful or very helpful in terms of support in finding work. This compares with 31% of NDYP participants who said that the programme was helpful in terms of getting a job, and 58% who thought NDYP was helpful in terms of looking for work.⁴⁵¹ Similarly, 57% of ND25pl participants said the programme had improved their chances of getting a job.⁴⁵²

Importantly, surveys of Work Programme participants also asked why they took up employment that did not match their skills or needs. Of these participants, 18% did so because of pressure from the Work Programme provider. Nevertheless, 23% said they felt pressure from the traditionally funded Jobcentre Plus.⁴⁵³ Therefore, the pressure to take up unsuitable employment appears to relate more to the mandatory nature of the programme and benefit sanctions, rather than the funding mechanism.

The following are some of the main benefits of the SOC scheme in comparison to traditionally financed models:

- **The focus on outcomes drove the organisation to accomplish the goal and to better consider economic sustainability.** This had several consequences: first, it gave the DWP greater 'power' over prime providers, making them easier to manage and the programme more cost effective. Providers also felt that performance targets helped them to develop a clearer understanding of what they should prioritise. One respondent remarked that outcomes-based contracts encourage the 'right' behaviour, in terms of developing the right solutions for employers and investing in the right processes at an organisational level. In the respondent's view, outcomes-based models motivate providers to think hard about the service offer that is put in place.
- **Increased transparency.** During the programme, the DWP also committed to publishing all performance data to allow providers to better assess and understand their own performance in different areas, as well as to be transparent regarding costs.
- In some cases, **the black box approach enabled innovation in terms of process.** One example provided by the interviewee was providers hiring people from the health sector, training them, and using their background to help those with health conditions at the time of the WP. Staff with a health background were able to engage with the participants and their doctors in a much more meaningful way about the types of employment that may or may not be appropriate for an individual, given their particular health condition (e.g. chronic back pain). Other innovative services included holistic support for programme participants. The programme evaluation

⁴⁴⁸ Meager et al. 2014, Table 7.5.

⁴⁴⁹ Bonjour et al. 2001, 67.

⁴⁵⁰ PricewaterhouseCoopers International Limited 2004, 110.

⁴⁵¹ Bonjour et al. 2001, 67.

⁴⁵² PricewaterhouseCoopers International Limited 2004, 61.

⁴⁵³ Meager et al. 2014, Table 7.6.

revealed some evidence that **competition between various providers** may have encouraged the provision of better services for disabled participants, or for those with a health condition.

- **Collaboration between providers with regard to employer engagement was common.** Examples of collaboration include: best practice sharing forums, the sharing of efficient processes (for example, in relation to tracking individuals and evidencing outcomes), cross-provider job shadowing and buddying, supply chain-wide training, joint employer screening days and vacancy sharing.⁴⁵⁴
- The **inclusion of new Employment and Support Allowance (ESA) customers** as one of the three payment groups appeared to have a positive impact on the development of services for participants who were disabled or had a health condition. In both 2013 and 2014, a number of prime providers described the development of these specific services. Some also noted that their relative improvement in performance with regard to outcomes for this payment group had been an important factor in their overall performance assessment by the DWP.

Conversely, the main **drawbacks of the WP** were:

- Although providers usually had to employ additional measures **to assist harder-to-help people** in getting closer to the labour market, the PbR contract did **not reward the distance travelled** toward the labour market in cases where no job was achieved. In such cases, the providers felt that they were not always sufficiently well paid by the PbR model, which resulted in financial risks for them as well as the **parking** of some harder-to-help participants. Several providers mentioned a need to prioritise those participants who were the closest to the labour market.⁴⁵⁵ **Differential pricing** was introduced, under which providers were rewarded more for working with harder-to-help clients. In theory, this aimed to prevent creaming and parking, but the research suggests that it had little impact in driving the providers' behaviour. For example, the development of services for the ESA participant group appeared to be driven by factors other than the payment model. These included more intensive performance management by the DWP. Providers reported that the key reason why differential payment groups did not drive their customer segmentation and delivery was heterogeneity, both within and between payment groups, in terms of the levels of support required. In practice, they also felt it was not possible to manage delivery and performance, given the complexity of the payment model. In addition to this, providers suggested that in some cases the costs of support for those with the greatest needs exceeded the payments available. Both inside and outside the supply chain, providers suggested a need to review the Work Programme's financial model.⁴⁵⁶
- Some DWP staff felt that initially, many prime providers did not fully understand PbR contracts, and that it took them time to adopt the right mindset and learn how to deliver and track their performance accordingly.⁴⁵⁷
- **Quick and chaotic introduction of the programme** (for elaboration, please see Section 1.11.5).
- **Lack of common understanding of volumes** between providers and authorities. According to one respondent, it was difficult to get a realistic estimate of volumes of participants that would be referred to the programme. It took around a year to obtain this, which resulted in greater risk for providers in the early stages.

⁴⁵⁴ Foster et al. 2014a, 108.

⁴⁵⁵ Foster et al. 2014a, 147.

⁴⁵⁶ Foster et al. 2014b, 5-6.

⁴⁵⁷ Foster et al. 2014a, 99.

- **The role of endogenous factors such as the economic situation.** If the economy is booming, this resulted in falling revenues, as clients in such an economic setting tend to be those who are harder to help. The interviewee noted that organisations' activities did not correctly match as fluctuations in unemployment rates. Even when there is economic growth, not all organisations (which then begin to make excessive profits) increase their help for those who are hardest to help.
- **PbR favours organisations with 'deep pockets' and a strong balance sheets.**

Scalability and replicability

Referrals to the Work Programme ended in March 2017, and no attempts have been made to replicate the programme. Based on the interviews and desk research outlined below, the programme was replaced by the new Work and Health Programme, in part because it had been largely ineffective for hard-to-help groups, even though it exceeded expectations for other participants. The subsequent Work and Health Programme (WHP) did, however, inherit some elements of the WP. For example, it too is a PbR scheme, rolled out across the UK, and focusing on employment outcomes. For more information on the WHP and specifically which elements have been changed in comparison with those of the Work Programme, please see Section 1.11.6.⁴⁵⁸

1.11.5. Efficiency

Intervention costs

No cost comparison can be made between the WP and the NDYP/ND25pl programmes because there limited information is available on the share of the total cost of the WP that can be attributed specifically to working with JSA claimants.⁴⁵⁹ Nevertheless, the commissioner interviewed noted that DWP is in the process of conducting an impact assessment of the Work Programme, which should provide greater clarity in the future. With regard to the New Deal, estimates of cost per job achieved vary widely, depending on the assumptions made as to the programme's impact as well as its savings. Estimate therefore range between GBP 4,000 and GBP 16,000 per job achieved.⁴⁶⁰

The evaluator interviewed warned that the efficiency which comes with PbR schemes may not necessarily stem from them delivering more effective services, but rather from fostering competition between service providers. To remain competitive, **some providers might cut costs by offering lower salaries to the advisers who work with service recipients, which might in turn affect service quality.** The commissioner interviewed argued that there is little information to determine whether outcomes-based contracts are more cost-effective than TF models. Importantly, he noted that **the current Work and Health Programme (which followed the WP) will be evaluated using a randomised controlled trial in which a small percentage of eligible participants will be referred to Jobcentre Plus instead of private providers.** The forthcoming Work and Health Programme evaluation will also provide useful insights for our study, because it will control for various factors that make SOC and TF programmes difficult to compare.

⁴⁵⁸ Department for Work and Pensions & Department of Health. (2017). *Improving lives: the future of work, health and disability*. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/663399/improving-lives-the-future-of-work-health-and-disability.PDF

⁴⁵⁹ Taking all of WP participants into account, the cost per participant would be equivalent to GBP 1,579, derived from the fact that 1.9 million participants were referred to the programme which cost GBP 3 billion during its duration.

⁴⁶⁰ Bogdanor, A. (2004). *Not Working: Why Workfare Should Replace the New Deal*. Policy Exchange. Retrieved from: <https://www.policyexchange.org.uk/wp-content/uploads/2016/09/not-working-apr-04.pdf>, 29-32.

Operational costs

According to the WP commissioner, **overall operational WP costs** amount to “a few tens of millions”, while programme expenditure totalled roughly GBP 3 billion between 2011 and 2019. This share appears substantially lower than the administrative burden of the New Deal, which over the course of the period 1997-2008 averaged 38% of total expenditure.⁴⁶¹ Nevertheless, until the detailed impact assessment of the WP is published, we cannot compare what exactly is included within these amounts to establish whether or not they are truly comparable.

The **set-up costs** of the WP appear fairly minimal. Procurement took place between July 2010 and June 2011. With just **six months** between the Invitation to Tender and the programme going live, the Work Programme procurement process was substantially quicker than procurement for the previous programmes.⁴⁶² However, **no information is available regarding the set-up of the ND programmes to compare costs.**

Furthermore, while the rapid set-up of the WP can be seen as efficient, it also created **challenges** for the providers. Most notably, providers faced difficulties in securing staff and premises in areas where they had not previously delivered. In addition, some providers encountered legal issues surrounding the Transfer of Undertakings (Protection of Employment) Regulations (TUPE). TUPE protects employees' terms and conditions of employment when a business is transferred from one owner to another. When the business changes hands, employees of the previous owner automatically become employees of the new employer on the same terms and conditions. Whether and how TUPE applies is difficult for providers to discern, and this often requires legal assistance. As a result, TUPE considerations made planning for adequate staffing very difficult, affecting the quality of the services delivered.⁴⁶³

Higher than expected referral volumes also made the start of the programme challenging: “In response to rising unemployment and revised Office for Budget Responsibility forecasts, six months after the Work Programme went live DWP released new estimates in which referral volumes increased from 2.5 to 3.3 million. These far higher than anticipated referral volumes presented a much bigger challenge for Work Programme providers. In the short term, income from attachment fees increased but providers reported that this increase in volumes at a time when they were in the early phases of operation created pressure on physical resources, leading to increased use of group sessions and online support. In the longer term, the difficulties reported by many providers in sourcing sufficient up-front funding to boost delivery, combined with their view that slow economic growth was making job outcomes and sustainment payments harder to achieve, may be a continuing influence on the shape of the programme in steady-state.”⁴⁶⁴ According to one provider: “we didn't have enough staff in place which meant we then were flooded with referrals which meant that all

⁴⁶¹ The 38% is the authors' own calculation, based on the information provided in DWP (2005). DWP Departmental Report 2004. Retrieved from:
<https://webarchive.nationalarchives.gov.uk/20050303081443/http://www.dwp.gov.uk/publications/dwp/2004/dr04/expenditure/table7.asp>. Planned expenditure is available here:
<https://webarchive.nationalarchives.gov.uk/20070305141234/http://www.dwp.gov.uk/publications/dwp/2006/dr06/annexa/table10.asp>

⁴⁶² Lane, P., Foster, R., Gardiner, L., Lanceley, L. & Purvis, A. (2013a). Research Summary: Work Programme Evaluation: Procurement, supply chains and implementation of the commissioning model. Retrieved from:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/197712/832summ.pdf, 2.

⁴⁶³ Lane, P., Foster, R., Gardiner, L., Lanceley, L. & Purvis, A. (2013a). Research Summary: Work Programme Evaluation: Procurement, supply chains and implementation of the commissioning model. Retrieved from:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/197712/832summ.pdf, 2.

⁴⁶⁴ Lane et al. 2013a, 2-4.

we were doing was attaching people. [As a result] delivering the quality journey that the organisation wants to do and obviously intends to do was severely impacted.⁴⁶⁵

In terms of overseeing SOC and TF schemes, **the WP commissioner thought that oversight costs were greater for PbR contracts than for TF schemes**, but noted that this was also because there is a greater need to monitor contractors' performance, which may potentially lead to better results (which the commissioner claimed was the case in the WP). Indeed, the outcome validation procedures employed in the WP were quite rigorous, and are described in Section 1.11.3. The end-to-end post-payment validation process took approximately eight and a half months to complete. Routine sampling, checks and production of validation rates took just over one month. These were performed monthly, one for each of the previous six months of job outcomes payments.⁴⁶⁶

The WP service provider interviewed added that their organisation spent **well in excess of a million pounds a year on evidence tracking**. A portion of this budget was dedicated to building an IT structure that communicated outcomes to the commissioning authority and conveyed all of the evidence required. The provider noted that evidence tracking diverted substantial amounts of time away from working with programme participants. In fact, **if evidence of outcomes achieved was too difficult to acquire** from certain types of people, **the provider in a PbR contract avoided working with those people**. The provider supported the opinion of the commissioner that TF schemes financed by the DWP usually involved a lower reporting burden. However, **the provider also added that there were vast differences between TF schemes in terms of reporting requirements**. Grants funded by the European Commission, for example, demanded more time to comply with reporting obligations than the Work Programme.

1.11.6. Design of the scheme

Design features

As mentioned in Section 1.11.1, the WP was in many ways designed as a response to the issues encountered during the Pathways to Work programme and FND. Some of the changes introduced were welcomed, whereas others were later reconsidered during the subsequent Work and Health Programme. Below, we outline which design features of the WP were endorsed by stakeholders, and which were regarded as problematic.

The stakeholders agreed that a longer contract provided greater stability to programme providers. The WP was implemented via long-term contracts of up to five years (with payments being issued up to seven years in respect of clients who entered during the final year of the contract). The greater market stability offered by this contractual framework was intended to facilitate the development of provider capacity and expertise, and to encourage investment that would support innovation in service delivery.⁴⁶⁷ The service provider interviewed appreciated the longer duration of the WP, arguing that his organisation made a loss in the first two years of the programme, and therefore time was needed in order to learn how to deliver services effectively, and for participants to find jobs that would result in outcomes-based payments. The subsequent Work and Health programme functions in a similar way.

Nevertheless, WP also offered notable lessons learned about effective programme design:

- The WP was envisaged as a single programme for different groups of unemployed individuals, making it easier for the end user to navigate employment services. In

⁴⁶⁵ Lane et al. 2013b, 16.

⁴⁶⁶ DWP 2018b.

⁴⁶⁷ Foster et al. 2014a, 126.

practice, however, the presence of nine different participant groups, each with unique needs, made it very difficult for service providers to manage the programme effectively. Providers not only had to work with diverse groups of participants, but they also lacked information on the volumes of clients with different needs that would be referred to the programme. Without knowing these volumes in advance, it was difficult to appoint the right number of experts across contract areas to work with customers from specific groups: “So, if I've got a contract across a big geographical area, let's say, Scotland, again, I might have two people who are experts at dealing with people who've left prison. But they can never provide the coverage, they can never provide the support across the geography.” As a result, the subsequent Work and Health Programme was scaled down to focus primarily on three groups: people with disabilities, the long-term unemployed, and people with various needs who might access the programme voluntarily.

- To foster competition and drive good performance, the WP allowed multiple providers to deliver services in the same contract areas, shifting referral volumes toward those providers with better performance ('market share shift'). While reasonable in theory, this idea was not very effective in practice, and was abandoned in the subsequent Work and Health Programme, in which there is now only one provider per contract area.⁴⁶⁸ The downside of having more than one provider was that it led to confusion, as employers had become frustrated with the different approaches of contractors seeking vacancies.⁴⁶⁹ Furthermore, the first market share shift occurred in summer 2013 but DWP staff interviewed in 2014 did not report it having any noticeable impact on performance. Many providers also reported that shifting volumes of referrals lacked impact due to the relatively small percentage shift (contractually – up to 5% of referrals could be shifted⁴⁷⁰) against the backdrop of falling referrals.⁴⁷¹
- The DWP's experience with the WP suggests that mandating might be more appropriate for easier-to-help groups than for more vulnerable unemployed groups. While the debate on sanctions in social services is beyond the scope of this paper,⁴⁷² it is relevant here for two reasons. First, mandatory clients generally have worse outcome success rates than voluntary clients, which in turn affects payments to providers. Second, if PbR schemes are designed to help voluntary clients only, their volumes will be more difficult to predict, and it will be more difficult to ensure sufficient scale for programme cost-effectiveness. Under the WP, certain groups of ESA claimants – many of whom claim benefits for health reasons – were required to participate in the programme in order to receive the benefit. This was subsequently changed under the Work and Health Programme, which mandates only long-term unemployed JSA claimants to participate, whereas clients with disabilities, mental health issues or other complex needs can engage voluntarily. The change was implemented after a 2016 consultation with DWP programme participants, organisations who represent them, employers, and the wide variety of

⁴⁶⁸ Powell, A. (2020). Work and Health Programme. House of Commons Library Briefing Paper Number 7845.

⁴⁶⁹ Lowe, S. (2015). JobPath: The proposed introduction of an employment programme in the Republic of Ireland. *The Public Sphere*, 120.

⁴⁷⁰ Dar 2016, 7.

⁴⁷¹ Foster et al. 2014b, 4.

⁴⁷² For more information, please see Dwyer, P.J. (2018). *Punitive and ineffective : benefit sanctions within social security*. Journal of Social Security Law. pp. 142-157; Griggs, J. & Evans, M. (2010). Sanctions within Chapter heading conditional benefit systems: A review of evidence. Joseph Rowntree Foundation. Retrieved from: <https://pdfs.semanticscholar.org/be8f/80220314e3890635eafe28a2de1ab7a23d13.pdf>; Oakley, Matthew (2014). Independent review of the operation of Jobseeker's Allowance sanctions validated by the Jobseekers Act 2013. Retrieved from: <http://www.haringeyadvice.org.uk/uploads/jsa-sanctions-independent-review-print.pdf> and others.

professionals who deliver services – including local authorities and health care professionals.⁴⁷³

- The commissioner interviewed also noted some lessons learned about the programme's performance targets and evaluation methods. The commissioner argued that it is important to have flexibility to change performance expectations to better reflect what is happening in the labour market, and to ensure that the programme is financially viable, but not excessively profitable, for providers. The service provider endorsed this opinion. To be able to adjust performance expectations, control groups are needed. Control groups are necessary in order to estimate value for money, and to know what outcomes would have been achieved without the programme. Allowing such flexibility within the contract would nevertheless make the design process even longer and the contract potentially more expensive (as it entails greater risk for the providers). Lastly, the commissioner noted that the Work and Health programme uses real-time HMRC earnings information, which works better in terms of transparency, independence, and in saving money and resources, than the paper-based system that was in place during the WP (real-time HMRC information was not used in the WP programme because it was not yet available).
- Lastly, interviewees disagreed over the usefulness of the black box approach. While the service provider argued that it fosters innovation (see Section 1.11.4), the evaluator noted that it places all the trust in the provider to deliver high-quality services and makes it difficult to understand why programmes work or do not work. Importantly, during the WP, service providers and commissioners differed over their understanding of how much flexibility providers were allowed in their delivery models. Performance managers generally viewed the black box as having only applied during contracting, while many providers believed they had the freedom to flex delivery during live running to meet participant needs.⁴⁷⁴ Therefore, if the black box approach is employed in similar PbR schemes, it is important to clarify what the minimum delivery standards are.

Impact on third-sector organisations and social enterprises

Like Pathways to Work (see Section 1.10), **the Work Programme highlighted the power imbalance between large prime providers and smaller subcontractors, many of which were NGOs.** According to one of the programme's evaluations:

- “One of the Work Programme's aims was to increase the flexibility of service provision by being open to small, local providers and allowing the provision of innovative, targeted schemes. It was hoped that this would enable organisations to deliver more specialised, person-centred services that would achieve better outcomes for 'hard to help' groups... unfortunately this ambition [was not wholly] realised. Although there [were] some specialists operating at [prime provider level] (for example, the Shaw Trust), specialist provision [was] mainly concentrated at 'Tier 2' [subcontractors]. Tier 2 subcontractors [were] dependent on referrals. There was no guarantee that they would receive sufficient numbers of referrals, and **in practice numbers were often insufficient to guarantee a viable income stream.** Referrals were also often highly complex cases, with less likelihood of payment. The risk for these providers [was] significant and [could have acted] as a barrier to involvement, even though they may [have been] highly effective in helping those with complex

⁴⁷³ DWP (2017). *Improving lives: the future of work, health and disability*. Retrieved from:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/663399/improving-lives-the-future-of-work-health-and-disability.PDF

⁴⁷⁴ Lane et al. 2013a, 2-4.

health-related barriers. **Many smaller and specialist providers had to leave the programme in its earlier years**, impacting on the diversity of providers and the quality of support available for people with health conditions...

- Providers noted that participation in the Work Programme initially resulted in financial losses, but became (slightly) profitable by the fourth year. Once attachment fees ended and the economic climate improved (hence there were fewer unemployed clients to be referred to the programme), the Work Programme became less viable for subcontractors.⁴⁷⁵

By 2014, prime providers largely agreed that the Work Programme was a financially viable and commercially attractive proposition, whereas just one in ten subcontractors reported that the Work Programme was commercially attractive.⁴⁷⁶

Smaller NGOs cannot act as prime providers in large PbR programmes such as the WP because their heavily outcomes-based payment structure entails too much risk.

If all or most of the funding is outcomes-based – as it was in the WP – only large organisations can afford to sustain their activities even if they fail to acquire outcomes-based payments. According to the commissioner, in earlier DWP programmes only a small share of the total funding was outcomes-based, and the purpose of this was to reward providers for the very best performance. In such cases, the organisation would not go bankrupt even if it failed to achieve the agreed outcomes. In programmes such as the WP, where outcomes-based funding is the main source of funding, some organisations (particularly smaller ones) would be at risk of financial ruin if they failed to deliver.

1.12. The Youth Contract (UK)

1.12.1. Background

The Youth Contract (YC), launched in April 2012, was a package of schemes aimed at helping young people into sustained employment.⁴⁷⁷ The YC encompassed a number of activities and target groups (including participants aged up to 24 years old). In this study, however, we pay the closest attention to only one aspect of the YC – **support for 16- and 17-year-olds not in education, employment or training (NEETs)**, as we have identified this as the only part of the YC that used a PbR model. The commissioning body for the YC 16-17 strand was the Education Funding Agency (EFA) within the Department for Education (DfE). Other strands of the YC included the Apprenticeship Grant for Employers of 16-24 year olds (AGE 16-24)⁴⁷⁸, work experience⁴⁷⁹, sector-based work academies⁴⁸⁰, extra support at Jobcentre Plus⁴⁸¹, funding for localised Youth Contracts⁴⁸² and wage incentives^{483, 484}. Other strands were also commissioned by other government departments,

⁴⁷⁵ Foster et al. 2014a, 126-127, emphasis added.

⁴⁷⁶ Foster et al. 2014a, 140.

⁴⁷⁷ Mirza-Davies, J. (2015). Youth Contract article. Retrieved from: <https://commonslibrary.parliament.uk/research-briefings/sn06387/>

⁴⁷⁸ Payments of GBP 1,500 are available to employers with less than 50 employees that take on young apprentices.

⁴⁷⁹ Placements are available through Jobcentre Plus for 16 to 24-year-olds who have been claiming JSA for at least 13 weeks.

⁴⁸⁰ Some 18 to 24-year-old JSA claimants will be offered a mixture of training, work experience, and a job interview at a local firm through Jobcentre Plus.

⁴⁸¹ 18 to 24-year-old JSA claimants will be offered weekly rather fortnightly signing-on meetings at Jobcentre Plus.

⁴⁸² Localised Youth Contracts are available in Leeds City Region, Liverpool and Newcastle. The cities designed their own local schemes using national funding.

⁴⁸³ Payments of up to GBP 2,275 were available to employers who took on young people (aged 18-24) claiming Jobseeker's Allowance (JSA) for more than six months.

⁴⁸⁴ Mirza-Davies, J. (2015).

namely the Department for Business, Innovation and Skills and the Department for Work and Pensions (DWP).

In the YC 16-17 strand, payments of up to GBP 2,200 were made to providers who took on 16- and 17-year-old NEETs who had low or no qualifications, or came from other disadvantaged backgrounds. The eligibility criteria were initially set to those youth who had no General Certificates of Secondary Educations (GCSEs) graded A*-C. However, in January 2013 the eligibility criteria were expanded to include youth with one GCSE qualification, young offenders released from custody, and (from August 2013) those serving community sentences with one or more GCSEs; and young people in care/leaving care with one or more GCSEs. Eligibility was extended due to lower than anticipated costs in the contracts procured (the prime providers bid at low cost) and because contractors struggled to recruit participants on the basis of the initial criteria.⁴⁸⁵

Providers offered targeted support to disengaged young people to get them into education, an apprenticeship or a job with training that was sustained for five of six months. Prime providers were free to design **activities** and tailor support, e.g. personal support, mentoring, work experience, work related skills training, confidence building exercises, outreach activities, etc. (the 'black box' approach was employed).

The YC strand for 16 to 17-year-olds operated two contrasting and distinct models:

- **A national model**, under which the Department commissioned a small number of national prime providers to develop supply chains of subcontractors in different localities, with delivery governed by a 'black box' approach, PbR and strict eligibility criteria.
- **Three core city areas**, with responsibility for determining the composition and delivery of the YC devolved to individual local authorities (LAs), which also had much greater flexibility in managing the application of PbR, their programme entry criteria, and in determining the measurement of sustained outcomes. The three core city areas comprised different numbers of LAs:
 - the first core city area comprised three LAs: Leeds, Bradford and Wakefield;
 - the second core city area comprised two LAs: Newcastle and Gateshead; and
 - the third comprised one LA: Liverpool.⁴⁸⁶

Similarly, the YC as a whole (including participants aged up to 24) was also delivered using the national and core city models. Each of the LAs collaborating as part of a core city area established its own delivery model. In some, but not all of these, prime provider-subcontractor delivery with PbR was used. Reasons for rejecting PbR in some areas included the implications of PbR for young people, as well as for organisations – where the financial risks of PbR would have been too high for local authorities.

Youth Contract **payments** were triggered when young people re-engaged with education, employment or training and then sustained these pursuits for five of six months:⁴⁸⁷

- An initial payment of up to 20% of the unit cost for one participant in year one of the programme, and up to 10% in year two. To receive the initial payment, providers had to identify an eligible NEET and work with the young person to complete an effective and clear action plan for their re-engagement.

⁴⁸⁵ Mason et al. 2015, 54.

⁴⁸⁶ Newton et al. (2014). pg. 129

⁴⁸⁷ Mirza-Davies, J. (2015).

- A further payment, based on the outcome of re-engaging the young person in education, training or work with training (the re-engagement point), was worth up to 30% of the unit cost.
- A final payment, based on the young person still being in education, training or work with training six months after re-engagement (the sustainability point) was the largest, amounting to 50% of the unit cost in year one or 60% of the unit cost in year two. Sustainability was defined as participation in full-time education, training or work with training (including apprenticeships) for five months out of six.⁴⁸⁸

Payments to prime providers and their subcontractors (when they were able to secure the full payment for sustained re-engagement) were issued on time; however, they were considered to be too low by the service providers themselves, with regard to the cost of identifying and supporting young people furthest away from re-engagement. This was further exacerbated by lower than expected volumes of young people entering the programme, which led many respondents to question both their rationale for delivering the YC, and the financial returns.⁴⁸⁹

The PbR approach was chosen because there was consensus among policymakers that the key goals for the national YC were to test PbR in a re-engagement programme and to focus support on the hardest-to-reach young people.⁴⁹⁰ Implementing the YC through PbR, with an emphasis on sustained outcomes and a black box delivery approach, granted freedom to providers to determine the nature of the intervention and also to allow innovative and effective practices to emerge. This new approach to delivering youth unemployment programmes was partly a reaction to the failure of the preceding, traditionally-financed Future Jobs Fund programme, and the relative success of the Work Programme (see Section 1.11).⁴⁹¹

According to the evaluator interviewed, the decision to use PbR emulated practices from Australia at the time, when public services were seen as inefficient in comparison to the private sector. According to prevailing narrative, the private sector could innovate in a way that the public sector could not.⁴⁹²

Under the national model, the EFA reported setting a **contractual expectation** that a sustained re-engagement rate of 50% would be possible, but most prime providers interviewed during the evaluation reported that the expectation was to sustain re-engagement for 70% of young people that entered the scheme (which in the provider's opinion was very high).⁴⁹³ Between 2012 and 2016, 47% of the young people referred to the programme achieved a sustained outcome.

The YC ended in March 2016. It was discontinued due to lack of political will (in 2015, the government changed from a Conservative-Liberal Democrat coalition to a Conservative government); the scheme's hefty price tag; the fact that participation levels in labour force and education at the time were high, making the rationale no longer relevant; together with the well documented failure of the wage incentive scheme within the 18-24 strand⁴⁹⁴. A summary of stakeholders involved in the YC is provided in the table below.

⁴⁸⁸ Newton et al. (2014) pg. 21

⁴⁸⁹ Newton et al. (2014) pg. 20

⁴⁹⁰ Mason et al. 2015, 54.

⁴⁹¹ Newton et al. (2014). The Youth Contract for 16-17 year olds not in education, employment or training evaluation, Research report, pg. 2 Retrieved from: <https://www.employment-studies.co.uk/resource/youth-contract-16-17-year-olds-not-education-employment-or-training-evaluation>

⁴⁹² Interview with the evaluator (2020).

⁴⁹⁴ Left Foot Forward, Here's why the government's Youth Contract failed, 2014, '<https://leftfootforward.org/2014/08/heres-why-the-governments-youth-contract-failed/>

Table 51. Summary of Youth Contract Strand 16-17

Personally targeted social service	Labour market-related
Target population	Youth not in employment, education or training (NEET) aged 16-17
SOC scheme type	Mixed PbR
Payment model	Binary
Cost (incl. currency)	Total YC cost GBP 1 billion, of which GBP 126 million was allocated to the 16-17 strand. ⁴⁹⁵ Due to the programme's early discontinuation, the true cost is likely to be closer to GBP 55 million (please see Table 52)
Commissioner(s)	Department for Education
Social service provider(s)	<p><u>National service providers:</u></p> <p>Groundwork (federation of charities)</p> <p>The Consultancy Home Counties Limited (for-profit enterprise)</p> <p>Pertemps People Development Group Limited (currently APM) (for-profit enterprise)</p> <p>Skills Training UK (for-profit enterprise)</p> <p>Prospect Training Services (Gloucester) Limited (for-profit enterprise)</p> <p>Prospects Services (for-profit enterprise)</p> <p><u>Core city providers:</u></p> <p>Six local authorities (LAs)</p>
Investor(s)	N/A
Intermediary	N/A
Evaluator(s)	<p>Institute of Employment Studies (IES) – university/research centre</p> <p>University of Warwick - university/research centre</p> <p>PRI Leeds Metropolitan - university/research centre</p>

1.12.2. Equivalent TF scheme

The **Activity Agreement (AA) Pilots** were chosen as a comparison to the YC 16-17 strand. Under the AA pilots, young people (and, for a two-year period in some areas, their parents) were offered a weekly allowance in return for agreeing to a plan to integrate them back into learning. Overall, the activities of AA can be divided into the following groups:

- Job-related: work-experience placements, work-related skills, work taster courses.
- Personal development: activities, courses or training not leading to a qualification, e.g., confidence-building, healthy living, sport/outward bounds, and specific skills development, e.g., drama.
- College-based: activities relating to going to college or doing formal qualification, basic skills.⁴⁹⁶

The scheme was due to end in 2011, but was discontinued in December 2010 due to budget cuts. In a letter to the pilot areas, the DfE stated that the decision was taken as part of a plan to make GBP 6 billion savings across government.⁴⁹⁷

⁴⁹⁵ Newton et al. (2014) pg. 22.

⁴⁹⁶ Tanner et al. (2009) pg. 29

⁴⁹⁷ Children and Young People Now, Activity Agreement And Entry To Learning Pilots To End In December, 2010, <https://www.cypnow.co.uk/news/article/activity-agreement-and-entry-to-learning-pilots-to-end-in-december>

The TF scheme is similar to the YC because it tackles a comparable target population, its activities were similar, and it shared a delivery approach under which local authorities, government departments and agencies were involved. Notable differences include financial incentives, such as weekly allowances, being given to AA participants. Such incentives not present in the YC. For a summary of how the two programmes compare, please see the table below.

Table 52. Comparability of Youth Contract 16-17 strand (PbR) with AA Pilots (TF)

	Same as SOC	Comments
Personally targeted social service	Yes	Labour market-related
Target population	Yes/No	16 to 17-year-old NEETs were the target population for both, with the YC prioritising 'harder-to-reach' cases
Number of participants	No	SOC: 37,969 between September 2012 and March 2016 ⁴⁹⁸ , amounting to roughly 9,500 participants per year. TF: Roughly 25,000 between April 2006 and March 2010 ⁴⁹⁹ , amounting to roughly 6,250 participants per year.
Location	Yes/No	Although both the YC (16-17 strand) and AA Pilots were implemented in England, some regions overlapped, while others did not. ⁵⁰⁰
Cost	Yes	SOC: Based on attachments, re-engagements and sustained re-engagements (see Figure 9), we estimate the total cost of the YC 16-17 strand to be roughly GBP 54 million, or GBP 13.5 million per year ⁵⁰¹ . TF: We estimate the cost to be roughly GBP 53 million, ⁵⁰² or GBP 13.5 million per year.
Commissioner(s)	Yes	Department for Education
Social service provider(s)	No	According to interviewees, smaller scale NGOs and private enterprises bid to deliver AA Pilots compared to YC

1.12.3. Outcomes measurement

The Youth Contract involved a simple system for self-reporting data concerning the measurement of outcomes upon which payment was conditioned.

Three main independent programme evaluations were conducted for outcomes not related to payments: Evaluation of the 16/17 Strand (led by Institute of Employment Studies

⁴⁹⁸ DfE 2016.

⁴⁹⁹ DfE 2010, 4.

⁵⁰⁰ For a list of AA Pilots regions, please see Maguire et al. (2011), p. 45 and DfE (2016) for Youth Contract.

⁵⁰¹ This is based on the fact that the maximum payment per participant across the different models is GBP 2,200, 20% of which is paid as an attachment fee, 30% is paid for re-engagement, and 50% for sustained re-engagement. For assumptions associated with these calculations, please see Newton et al. 2014, 123. Please note that the actual cost of the YC 16-17 strand is not available.

⁵⁰² The DfE (2010; 4) has quoted the cost to be GBP 2,122 per participant. The same source indicates that there were roughly 25,000 AA participants, bringing the total cost to roughly GBP 53 million.

[(IES)]⁵⁰³; Evaluation of the 18/24 Strand (led by Kantar [formerly TNS BMRB]); and Evaluation of AGE (Apprenticeships Programme - 16/24), led by BMG Research.

Measurement of outcomes relating to payment mechanisms

Outcomes relating to payment were measured using a simple self-reporting system by local providers, or alternatively by local subcontractors.

The outcome measured related to the key domain of the programme – the engagement of NEETS in education, training or employment. A variety of indicators were used. These can be distinguished as indicating either initial re-engagement outcomes, or sustained engagement outcomes. The initial re-engagement outcomes were:

- Participation in full-time education or training leading to an accredited qualification funded by the EFA.
- Participation in part-time education, including re-engagement provision, funded by the EFA. Young people were required to participate in at least seven hours of directed learning per week.
- Participation in an apprenticeship.
- Participation in full-time employment (20 hours or more each week) with part-time training equivalent to at least 280 guided learning hours per year (around one day per week).
- At the point of sustained engagement, acceptable positive outcomes were:
 - Sustained participation for at least five out of six months in full-time education or training leading to an accredited qualification funded by the EFA.
 - Sustained participation for at least five out of six months in an apprenticeship.
 - Participation for at least five out of six months in full-time employment with part-time training equivalent to at least 280 guided learning hours per year (around one day per week).

Other evaluations

Impact evaluation

The evaluation of the 16/17 strand also included an *impact assessment* conducted using the propensity score matching approach. Propensity scores were estimated using Probit models explaining the individual participation of particular groups (by age and by gender) in the Youth Contract, compared with non-participation on the basis of the following observable characteristics:

- ethnicity;
- regional or local areas;
- educational achievement in GCSEs and at Key Stage 3;
- exclusions and absence during the year of Key Stage 4;

⁵⁰³ Newton et al. (2014), p. 26

- time since leaving secondary education;
- duration of initial NEET spell;
- young person's level of need prior to joining the YC;
- pre-Youth Contract employment and education experiences.

A monetisation of outcomes evaluated in the impact assessment allowed a *cost-benefit analysis* (value for money analysis).⁵⁰⁴

Moreover, the two strand assessments also included a *process evaluation* carried out through interviews with national stakeholders and with prime contractors, a survey for local authorities, and through extensive, longitudinal, and multi-perspective case studies.

Barriers to and enablers of the measurement process

Overall, the **cooperative behaviour of commissioners** and the transparency of data from the service providers can be considered enablers of the measurement process. Nonetheless, it is possible to identify several barriers to the measurement process.

Problems concerning the **communication and disclosure of data** were reported, although these were not directly associated with the evaluation itself. **Data quality** was another weakness. For this evaluation, young people taking part in the YC were identified using the National Client Caseload Information System (NCCIS). In the NCCIS data, young people were tracked more or less frequently depending on their participation status, if participation in learning or training is presumed to continue until the end of the academic year. Therefore, it automatically assumes that the period of learning or training is longer than it may actually be, thus 'overselling' the achieved impact if the participant leaves learning or training activities prior to the end of the academic year.

Table 53. Summary of the measurement process in Youth Contract

Measurement methodology	
Methods related to payment mechanisms	Data self-reported by providers
Other evaluations	Impact evaluation of 16/17 strand
Methods using experimental or quasi-experimental design	Quasi experimental – propensity score matching
Control group	Only for the impact assessment components of Evaluations Strand 16/17 – not related to payment
Causality of impact	Causality of impact can be detected only in the impact assessment components of Evaluations Strand 16/17 – not related to payment
Enabling factors and strengths of the evaluation process	- Cooperative behaviour of commissioner Transparency of data from the service providers
Barriers to and weaknesses of the evaluation process	- Communication and disclosure of data Lack of unified evaluation design

⁵⁰⁴ Nafilyan et al. (2014), p. 29

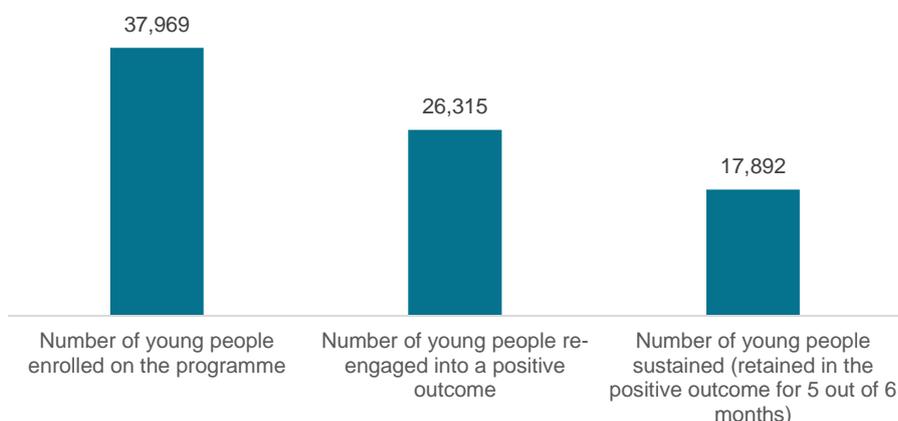
Governance of measurement	
Independence of evaluator	Yes, for outcomes not related to payment.
Evaluator	Local service providers (or subcontractors) Institute of Employment Studies (IES)-led consortium – university/research centres

1.12.4. Effectiveness

Achievement of outcomes

As presented in the figure below, out of 37,969 young people that enrolled in the programme, 26,315 were re-engaged with a positive outcome, and 17,892 sustained this outcome for five out of six months.⁵⁰⁵ This corresponds to 47% of the young people referred to the programme achieving a sustained outcome, compared with the target of 50%.

Figure 9. Youth Contract outcomes (16-17 strand), 2012-2016



Source: elaborated by PPMI, based on data from Department for Education (2016).

Note: the figure includes both the original and extended cohorts.

There is overall consensus that the YC's target was achieved only to some extent, since a large proportion of the planned budget remained unused after the scheme was discontinued.

Comparison with the outcomes achieved by the AA programme is problematic. It must be taken into account that the **client group** for the YC included clients who were more difficult to reach, and that the outcome under AA had to be sustained for three months, compared to five months under the YC. That said, the rate of re-engagement was remarkably similar for the two programmes: 49% for AA, and 47% for YC.

Opinions as to how the two schemes compare in terms of effectiveness are diverse. The commissioner noted that the YC performed fairly well, considering its tight eligibility criteria and its more stringent outcome measurement. One evaluator interviewed, who worked on both the YC and AA, felt that both interventions **performed rather poorly** in reaching young people, and only partly achieved the objectives that were set for them.

⁵⁰⁵ Department for Education (2016). Youth Contract delivery data. Retrieved from: <https://www.gov.uk/government/publications/youth-contract-delivery-data>

Benefits and drawbacks of SOC compared with TF

The PbR model in the Youth Contracts seemed to have **enhanced collaboration between prime providers and LAs**. For instance, some of the prime providers engaged with LAs in the contract package areas in which they proposed to bid, and then designed the supply chain taking into consideration the role that LAs could play and the support they could offer. As one provider said: “We went to local authorities with a blank piece of paper and asked them what they wanted us to do and I think we won a lot of brownie points”⁵⁰⁶. This early engagement with LAs may have underpinned effective relationships for delivery.

Various stakeholders identified the **risk of creaming** as the most critical flaw of the scheme. According to the commissioner interviewed, the creaming of participants in the YC was a significant issue. He stated that due to the focus of the YC on hard-to-reach and hard-to-help beneficiaries, some providers chose young people with less disadvantaged backgrounds to ensure that re-engagement and sustainability payments were achieved. One respondent, who was also involved in the evaluation of the TF scheme, claimed that creaming of participants was the main issue for the YC,⁵⁰⁷ and that it was evident that larger companies were effectively engaging with easier-to-reach target groups, at least for the first payment.

With regard to the ability of SOC to promote service innovation, the service provided under the YC was not particularly innovative. Many national stakeholders highlighted the use of the **black box approach** linked to PbR, and emphasised that such design features should encourage individualisation and service innovation. In practice, however, there was a lot of scepticism surrounding this delivery model after the scheme had finished. The commissioner mentioned that the service providers **did not innovate as much as had been anticipated** before the scheme commenced.⁵⁰⁸

Also, the opinions of the stakeholders interviewed were mixed with regard to the ability of this SOC scheme to meet the needs of the beneficiaries better than the TF equivalent. The commissioner remained optimistic about the PbR scheme reaching the most disadvantaged. In the commissioner’s opinion, the YC was designed to target more vulnerable target groups. However, according to providers, the **qualification-based eligibility criteria** for entering the national programme proved to be restrictive. Crucially, the criteria seemed to make it harder to reach some of the groups who were hardest to reach and help, but who fell outside these tight boundaries.⁵⁰⁹

Scalability and replicability

According to the commissioner, no attempts have been made to replicate or scale the YC. As he described it, the current political climate and the negative baggage of the programme would not allow any replication or scaling.

1.12.5. Efficiency

Intervention costs

Given that the total costs of neither YC nor AA Pilots were disclosed, the cost estimates available in Table 55 are based on a number of assumptions. First, regarding the YC, we estimate the total cost of the programme based on the payments associated with

⁵⁰⁶ Newton et al. (2014) pg. 51

⁵⁰⁷ Interview with the evaluator (2020).

⁵⁰⁸ Interview with the commissioner (2020).

⁵⁰⁹ Newton et al. (2014) pg 129

programme outcomes and statistical information available about the number of participants as well as how many of them were re-engaged. In general, we apply the same methodology as the programme’s evaluators (for details, please see the footnotes in the table below).⁵¹⁰ It is important to note, however, that these numbers are likely to be overestimates. Second, for AA Pilots, we derive the total cost of the programme, as well as the cost per sustained re-engagement outcome, from the available information on the programme’s cost per participant, programme success rate, and the number of participants.

Importantly, we have chosen to calculate programme costs per sustained re-engagement outcome, rather than per re-engagement outcome because the results for AA Pilots were reported three months after participants completed the programme. This reporting criterion was not identical – yet nevertheless similar – to that for the YC, in which participants had to remain in education or employment for at least five out of six months in order for the ‘sustained re-engagement’ outcome to be met.

The comparison is also subject to certain limitations. Although both programmes targeted 16-17 year-old NEET youth, we have no information as to the extent to which participants in the two groups were actually similar. For example, the two groups may have differed significantly in terms of how long they had been out of school, their familial situation, etc. These differences, coupled with other factors such as a different economic climate (the AA pilots took place during the 2008 recession), might at least in part explain the differences we observe in Table 55. We therefore cannot claim that one scheme cost more per participant strictly because of its funding arrangement.

Bearing these limitations in mind, the YC 16-17 strand appears to have been more efficient than AA Pilots. Both the cost per participant and the cost per sustained re-engagement outcome in the YC programme are roughly two-thirds of the equivalent costs for the AA Pilots. According to the commissioner interviewed, PbR drives the contracted service providers to strive for quality over quantity, thus ensuring efficiency through innovation and additional effort.⁵¹¹

Table 54. Cost comparison between the YC strand 16-17 and AA Pilots (GBP)

	YC strand 16-17 (PbR) 2012-2016	AA pilots (TF) 2006-2010
Participants	37,969 ⁵¹²	25,000 ⁵¹³
Sustained re-engagements	17,892 ⁵¹⁴	12,250 ⁵¹⁵
Total cost	53,755,460 ⁵¹⁶	53,050,000 ⁵¹⁷

⁵¹⁰ See Newton et al. 2014, 122-123.

⁵¹¹ Interview with the commissioner (2020).

⁵¹² Authors’ own elaboration, based on DfE (2016). Youth Contract: delivery data: September 2012 to end of March 2016. Retrieved from: <https://www.gov.uk/government/publications/youth-contract-delivery-data>

⁵¹³ DfE 2010, 4.

⁵¹⁴ Authors’ own elaboration, based on DfE 2016.

⁵¹⁵ DfE (2010; 5) notes that 49% of AA Pilot participants were engaged in education and employment-related activities three months after participation in the programme. We used this information to estimate the total number of re-engaged participants.

⁵¹⁶ This is based on the fact that the maximum payment per participant across the different models is GBP 2,200, 20% of which is paid as an attachment fee, 30% is paid for re-engagement, and 50% for sustained re-engagement. For the assumptions associated with these calculations, please see Newton et al. 2014, 123. Please note that the actual cost of YC 16-17 strand is not available.

⁵¹⁷ DfE (2010; 4) has quoted the cost to be GBP 2,122 per participant. The same source indicates that there were roughly 25,000 AA participants, bringing the total cost to roughly GBP 53 million.

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE
PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

	YC strand 16-17 (PbR) 2012-2016	AA pilots (TF) 2006-2010
Cost per participant	1,416	2,122 ⁵¹⁸
Cost per sustained re-engagement outcome	3,004	4,331

Source: elaborated by PPMI, based on the available information.

Note: the costs have not been adjusted for inflation. Such adjustment would further increase the costs of the AA Pilots relative to those of the Youth Contract. Furthermore, sustained re-engagement outcomes were measured differently between the two programmes. Namely, YC participants had to be enrolled in school, work or apprenticeship for five out of six months whereas information for AA Pilots indicates how many youth were in employment or education three months after participating in the programme.

Operational costs

Although information was limited as to the overall operational costs of either programme, **the commissioner interviewed stated that the YC programme was less expensive to set up and manage than similar TF schemes such as AA.**

There are two reasons for this: first, the set-up costs for the YC were minimal in comparison to its predecessors, AA Pilots and the Work Programme, because the YC borrowed concepts from these programmes.⁵¹⁹ For example, YC contractors were required to adopt the Merlin Standard developed under the WP to ensure the quality of services across their supply chains.⁵²⁰ Some providers nevertheless perceived this quick start-up negatively, arguing that they had insufficient time to fill-out pre-qualification questionnaires. This led to the disqualification of some bidders at the initial stage, who were later allowed to re-enter the bidding process.⁵²¹

Second, contractors in the YC comprised a few large service providers, in comparison to a multitude of smaller organisations delivering services under the AA – some of which were less experienced than their YC counterparts – making the implementation of the latter programme more resource intensive for the DfE.⁵²² Furthermore, under programmes such as the WP, substantial commissioner resources were dedicated to verifying the outcomes achieved. By contrast, the EFA adopted a lighter approach. For example, instead of verifying all outcomes (as had been the case under the WP), the EFA ran spot checks and audits of the outcomes obtained.⁵²³ Moreover, in many ways YC prime providers acted as contract managers in the way that a commissioner would in traditional models. For example, they set minimum delivery standards for subcontractors, including minimum expectations for the number of key workers, key principles of delivery, and in some cases daily monitoring of outputs and outcomes. All prime providers had at least one manager dedicated to these tasks.⁵²⁴ Prime contractors deducted a management fee from subcontractor payments for performing this function.⁵²⁵ Therefore, the YC was operationally less expensive than similar

⁵¹⁸ DfE 2010, 4.

⁵¹⁹ Interview with the Commissioner (2020).

⁵²⁰ The Merlin Standard was designed by the Department for Work and Pensions (DWP) to help evolve successful, high performing supply chains, and champion positive behaviours and relationships in service provision, and to ensure fairness within supply chains.

⁵²¹ Newton et al. 2014.

⁵²² Interview with the commissioner.

⁵²³ YPLA, DfE and DWP (2012). Youth Contract – Support for 16-17 Year Olds who are Not in Education, Employment or Training. Retrieved from: <https://data.gov.uk/data/contracts-finder-archive/download/432688/6ebdbbc6-9ecd-4b2e-bff7-a3fb6bc8de2e>

⁵²⁴ Newton et al. (2014) pg. 65

⁵²⁵ Mason et al. (2015). pg. 8

TF or PbR programmes because prime contractors took over some of the functions typically performed by the commissioning body.

While implementation was not resource-intensive for the commissioner of the YC, it was for service providers. One core city contact, for example, mentioned that “the amount of information required for that type of programme is quite intense, and we [mirrored] **the same process that we do on the projects that we’re managing under ESF [European Social Fund]**. So we collect information and data for as much, everything that you can think of, as much data as we can.”⁵²⁶ Most prime providers introduced new systems to manage data and comply with differing data requirements for providers, young people and for themselves in respect of capturing information about performance and outputs. Some subcontractors managed multiple IT systems in order to monitor and track the young people they engaged, keep track of payment claims, and provide the information required by prime providers and LAs. For these subcontractors, the administration of the YC was seen as particularly resource-intensive, marginalising the financial benefits of the contract (for more information on the financial viability of the YC, please see Section 1.12.6).⁵²⁷

1.12.6. Design of the scheme

Design features

One of the key design features of the YC was its tight **qualification-based eligibility criteria** (see Section 1.12.1). This aimed to ensure that funding was targeted at those young people who were hardest to reach or help. However, **acquiring proof of participants’ qualifications proved a challenge**. Most providers anticipated that LAs would act as a source of this information, but LAs were not ready to share it with providers, citing legal reasons (although national guidance stated this should not be an impediment, provided data sharing agreements were in place).⁵²⁸ Therefore, some providers reported not being able to claim payments for some outcome, even though the outcomes had been achieved:

“The sustainability rules, which looked alright at the beginning, are very harsh – we have to have [the participants] in for five out of six months... And if they drop out – and we’ve had lots of issues where the provider has closed – sometimes it takes five or six weeks to get them into something else and it’s over the month gap they have to start again. And by that time they’re often 18! Our providers have a number of examples of where young people have re-engaged and the providers know but they haven’t been able to claim it properly – because they’ve moved area, etc. We’re not being able to claim for a good 20% of young people that we should be able to claim for because of the rules.”⁵²⁹

Furthermore, despite broadening its eligibility criteria, the programme still ran into **recruitment issues**. According to one provider, “no provider is ever going to get above 80, 90% of the volumes that we’ve contracted for.”⁵³⁰

Designers of PbR schemes should consider whether the eligibility criteria are: 1) possible to prove; and 2) how they will affect the volumes of participants. Disagreements between providers and commissioners over the eligibility criteria also show that **sufficient time should be given for programme start-up** to resolve these and other issues. There is also a need for **flexibility in contracts**. This enables commissioners and providers to work together to recognise issues in the design and to make adjustments to metrics and payments accordingly. The Youth Contract evaluation nevertheless suggests that scope for

⁵²⁶ Newton et al. (2014) pg. 64

⁵²⁷ Newton et al. (2014) pg. 65

⁵²⁸ Newton et al. (2014). pg. 76

⁵²⁹ Newtown et al. 2014, 72.

⁵³⁰ Newton et al. (2014) pg 47.

this could be limited by the risk of legal action from unsuccessful providers, should the terms change significantly.⁵³¹

Another important lesson about programme design concerns the **timing of the programme**. A DWP representative remarked that at the time the YC was in operation, the labour market was doing well, meaning that the programme was as necessary, because there were fewer unemployed youth than there had been during the time period post-2008.⁵³² This, **coupled with a lack of political will** to continue the programme, resulted in its early termination. The YC experience demonstrates the need for a **feasibility study** prior to launch, which did not take place in the case of the YC.⁵³³ Such a feasibility study could provide a **realistic estimate** of how many people are likely to enter the labour market at a given time. According to the evaluator, this information is crucial to determining sufficient payments to contractors.

Nevertheless, one of the strongest features of the YC was **flexibility for LAs to design the scheme in the best way they saw fit**. Therefore, the YC encompassed a number of different payment arrangements, which could better reflect the needs of the target group present in that area. One LA, for example, offered different total budgets per participant to different subcontractors, reflecting the more or less disadvantaged populations each subcontractor would attract. The LA also designed a six-stage, rather than a three-stage, payment model (see Section 1.12.1). Contractors were paid for attachment (12%); initial assessment (12%); development of an action plan (12%); a mid-term review (12%); initial re-engagement (20%); and six-month sustained re-engagement (32%).⁵³⁴

Impact on third-sector organisations and social enterprises

By the time the YC evaluation was conducted, **more service providers were in favour of PbR than against it**. Nevertheless, both for-profit primes and smaller subcontractors argued that **the contract was too focused on sustained outcomes**, the expected levels of which were impossible to attain, making **the contract not financially viable for some organisations** – a situation that was made worse still by low volumes of participants. Many would have preferred for payments to reflect the distance travelled by participants, even if they had not achieved a sustained outcome.⁵³⁵ It was common for prime providers to state, on the basis of their delivery experiences, that if they could go back in time, they would not bid for the YC: “If this had to be tendered out again, I’m not sure how many people who have had experience of it would go for it again because you cannot earn.”⁵³⁶ The provider interviewed nevertheless contrasted the national model with the core city models, which provided more upfront funding: “The other ‘pilots’ that weren’t tendered out – their payment model is entirely different and much more up-front and their programme is much more ‘successful’.”⁵³⁷

Little upfront funding also meant that **many LAs did not bid for the contracts**. Some did not have the resources to develop their bids, whereas others were uncertain of staffing levels in the future due to funding cuts, and could not afford the risk that the YC entailed. This caused additional issues: when primes went to work with LAs, the latter were not

⁵³¹ Mason et al. (2015), pg. 9

⁵³² Interview with the commissioner (2020).

⁵³³ Interview with the evaluator (2020).

⁵³⁴ Newton et al. 2014, 54.

⁵³⁵ Newton et al. (2014) pg 43

⁵³⁶ Newton et al. (2014) pg. 70

⁵³⁷ Newton et al. (2014) pg. 70

always cooperative (see Section 1.12.3 on obtaining proof of participant qualifications). Nevertheless, this improved over time.⁵³⁸

1.13. Drug and Alcohol Recovery Pilots (UK)

1.13.1. Background

In April 2012, the Department of Health piloted a PbR scheme regarding **drug and alcohol recovery** in eight local administrative areas in the UK. The programme ended in 2013, having had involved 6,582 drug clients and 3,081 alcohol patients. The programme provided services and treatments against drug and alcohol misuse, aimed at re-integrating clients into community life.

The set-up process for the scheme involved a six-month co-deciding phase, during which the Department of Health, service providers, as well as independent consultants (the Policy Innovation Research Unit) defined various outcomes and targets⁵³⁹. The **rationale** for the programme was based on the 2010 Drug Strategy Agenda, which stated that it is not sufficient to help addicts achieve abstinence from drugs or alcohol. In line with the Agenda, a full recovery and sustained life without substance abuse was set as the programme's main goal. The Pilots included multi-dimensional outcomes, moving away from measuring only health-related indicators to encompass other spheres of life such as employment and reoffending rates. Outcome targets were set at both local and national levels.

The implementation of the scheme was guided by the principle of localism: local authorities commissioned locally-based service providers and managed the evaluation through the Local Area Single Assessment and Referral System⁵⁴⁰. Local authorities decided to implement the Pilots for a variety of reasons:

- to challenge historical performance and attract new providers;
- to build on a developing recovery system that was already in place;
- to create a more efficient and effective recovery system;
- to take a broad approach to recovery;
- to support clients in their recovery ambitions; and

to achieve a measurable increase in the number of people exiting the services in a successful and planned way.⁵⁴¹

Among the eight piloting sites, a diversity of **payment** arrangements existed. At three sites, all payments were outcome-based, although one of these sites eventually allocated 30% of the contract value as attachment fees. At the other sites, the portion of the contract value that was based on outcomes varied between 10% and 30%. Only a share of payments was outcome-based due to the necessity to increase funding stability for the providers, and to ensure sufficient cash flow for service delivery. More specifically, the funding mechanism was upfront, based on an attachment fee provided with a results-based claim-back clause in the majority of contracts. Such an approach allowed payments to be made earlier in the

⁵³⁸ Mason et al. (2015) pg. 54

⁵³⁹ Department of Health. (n.d.). Drug and Alcohol Recovery Pilots. Lessons learnt from co-design and commissioning with payment by results. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/118036/pbr-lessons-learnt.pdf, 7-9.

⁵⁴⁰ Donmall et al. (2017). Evaluation of the Drugs and Alcohol Recovery Payment by Results Pilot Programme Final Report. The University of Manchester, 10-12

⁵⁴¹ Department of Health n.d., 6.

scheme, in order to ease the cash-flow issues faced by providers. These payments could then be recovered at a later date if performance did not meet expectations.⁵⁴²

Moreover, with the exception of a key outcome named ‘outcome for reliable change’, all the pilot areas were free to set the price and to weight differently for each outcome in the payment system. The ‘outcome for reliable change’ clause meant that the initial outcomes relating to significant improvement in the client’s drug and alcohol usage could not weigh more than 20% of the money made available to providers for the ‘free from drug(s) of dependence’ outcome domain.⁵⁴³ This clause was set in coherence with the ‘full recovery’ objectives of the programme, which did not end with drug and alcohol abstinence. In determining how much to pay for each outcome, the pilot areas sought to ensure that recovery outcomes remained key, while at the same time ensuring sufficient incentives in the system for providers to engage with clients who were least likely to achieve outcomes.

For instance, at Site A in the first year, 90% of payment to providers was upfront, 5% was based on outputs (e.g. Hep C test, Hep B vaccine, Treatment Outcome Profile [TOP] completion, waiting times, etc.), and 5% was based on national outcomes. In the second year, 80% of payments were based on final outcomes, and 20% on interim/process outcomes (for more details on interim and final outcomes and outcomes that were set nationally, please see Section 1.13.3). At Site G, the individual tariffs set for each client were based on an initial screening and risk assessment in six domains (substance misuse; risk to self; risk to others; risk to children; risk from others; and offending). At this site, 30% was paid upfront as an attachment fee; 39% was payable on interim performance measures; while 31% on final outcomes.⁵⁴⁴

Lastly, payments and tariffs were also associated and clustered according to levels of patients’ complexity.⁵⁴⁵

Details about the achievement of the outcomes at each local site are not publicly available. However, the impact evaluation provides some insights into the impact of the project. It shows that the programme had **mixed effects**. Patients undergoing programme services had lower rates of misuse, treatment initiation and completion than other services, but higher rates of in-treatment abstinence and non-injecting.⁵⁴⁶

Table 55. Summary of Drug and Alcohol Recovery Pilots

Personally targeted social service	Social exclusion
Target population	People addicted to drugs and alcohol: 6,582 drug clients; 3,081 alcohol clients ⁵⁴⁷
SOC scheme type	Mixed PbR
Payment model	Binary
Cost	Total set-up cost: GBP 1,363,856 ⁵⁴⁸ (set-up costs for individual piloting sites differed).

⁵⁴² Donmall et al. 2017, 63; Mason et al. (2015). The impact of paying treatment providers for outcomes: difference- in-differences analysis of the ‘payment by results for drugs recovery’ pilot. *Addiction*, 110(7), p.1120.

⁵⁴³ Donmall et al. 2017, 46, 47.

⁵⁴⁴ Donmall et al. 2017, 132-140

⁵⁴⁵ Donmall et al. 2017, 54.

⁵⁴⁶ Jones, A., Pierce, M., Sutton, M., Mason, T., & Millar, T. (2018). Does paying service providers by results improve recovery outcomes for drug misusers in treatment in England? *Addiction*, 113(2), 279

⁵⁴⁷ Mason et al. 2015, 1122.

⁵⁴⁸ Donmall et al. 2017, 63.

	Total costs varied widely among the Drug Addiction Teams (DATs) (see Section 1.13.5)
Commissioner(s)	Two levels of commissioning can be distinguished: <ul style="list-style-type: none"> - Department of Health.⁵⁴⁹ - Eight local commissioners: Bracknell Forest, Enfield, Lincolnshire, Oxfordshire, Kent (west Kent only), Stockport, Wakefield and Wigan.
Social service provider(s)	Subcontracted local providers. Information about the identity of providers is not available for all areas. In Oxfordshire, the private provider was the charity Aquarius. ⁵⁵⁰ In Lincolnshire, the providers were a public agency of the NHS taking care of drug addiction services, and a third-sector organisation named Addaction. ⁵⁵¹
Investor(s)	N/A
Intermediary	N/A
Evaluator(s)	Evaluation of outcomes <u>associated with payments</u> was mostly based on self-reported answers on individual measures, and validated by Local Areas Single Assessment and Referral System. ⁵⁵² An <u>independent impact evaluation</u> of the programme was led by the University of Manchester. Rand Europe; Birkbeck, University of London; and User Voice contributed to this evaluation. ⁵⁵³

1.13.2. Equivalent TF scheme

The search for a TF scheme that was equivalent to the Drug and Alcohol Pilots appears challenging. One interviewee noted that the programme’s focus on the outcome of ‘full recovery’, its ‘localist’, and the inclusion of both alcohol and drug recovery aims, were new features for a recovery programme, preventing easy comparison with previous programmes in the field.⁵⁵⁴

In the impact evaluation mentioned above, the eight pilot sites were compared with 141 non-pilot commissioning areas. In the latter areas, however, other PbR schemes that were different from the Drug and Alcohol Recovery Pilot were implemented by local authorities. Thus, the 141 non-pilot areas were not subject to a single comparable TF programme, and hence the results of the impact assessment cannot be used to compare SOC and TF schemes.

For this reason, the provider interviewed suggested comparing the intervention with the same service provided before the launch of the scheme.⁵⁵⁵ The comparison has been carried out with specific reference to the Lincolnshire LA, using data retrieved from the interviews.

Before the pilot, services in Lincolnshire were provided under a 100% block contract system. Given that under this system, the same two providers deliver services to the same target population of the area, comparison is therefore possible.⁵⁵⁶

⁵⁴⁹ Policy Innovation Research Unit (PIRU). (2011). Payment by Results (PbR) Drug and Alcohol Recovery Pilot Programme: a note of advice to the Department of Health (DH) on the proposed evaluation, 3.

⁵⁵⁰ Department of Health n.d., 10-11.

⁵⁵¹ Interview with the commissioner.

⁵⁵² Department of Health n.d., 9; PIRU 2011, 4.

⁵⁵³ Donmall et al. 2017, 1.

⁵⁵⁴ Interview with the evaluator, conducted on April 21st 2020.

⁵⁵⁵ Interview with the provider.

⁵⁵⁶ Interview with the commissioner.

Table 56. Comparability of the Drug and Alcohol recovery pilots with the previous TF intervention in Lincolnshire

	Same as SOC	Comments
Personally targeted social service	Yes	Social exclusion
Target population	Yes	Alcohol and drug clients
Number of participants	Yes/No	SOC: 2,013 in 2012; 2,050 in 2013 (in Lincolnshire) TF: (only data for drug clients were available) 1,916 drug clients in 2010; 1,983 drug clients in 2011 ⁵⁵⁷
Location	Yes	Both in Lincolnshire.
Cost	Unknown	Detailed breakdowns of costs were not available
Commissioner(s)	Yes	Department of Health and local authority
Social service provider(s)	Yes	Lincolnshire Partnership NHS Trust and Addaction (non-profit organisation)

1.13.3. Outcomes measurement

The pilot scheme entailed a system of data reporting and monitoring at local level regarding the outcomes upon which payment was conditioned. This process was mainly monitored by a specific local system or Local Area Single Assessment and Referral Service (LASARS).

Aside from this, the scheme was also subject to a qualitative evaluation, an impact evaluation, and an economic evaluation, carried out by the University of Manchester; Rand Europe; Birkbeck, University of London; and User Voice.⁵⁵⁸

Measurement of outcomes relating to payment mechanisms

Data regarding the achievement of outcomes upon which payments were conditioned were self-reported by patients during Treatment Outcome Profile (TOP) reviews. These were periodic meetings with patients to assess their treatment profile.

These data were monitored, assessed and linked to the tariffs scheme by a Local Area Single Assessment and Referral Service (LASARS). LASARS were developed for the purposes of the PbR scheme, and were specifically charged with confirming outcomes to trigger payments.

LASARS represented a key service and part of the PbR scheme. The service was present in every local authority and entailed:

- Auditing functions to ensure that service users and outcomes are correctly assessed.
- Monitoring the payment systems, ensuring they are correctly formulated.

⁵⁵⁷ These data are not publicly available and were provided by the commissioner interviewed.

⁵⁵⁸ Unless stated otherwise, the information in this section is taken from Donmall, M., Sutton, M. et al. (2017). Evaluation of the Drugs and Alcohol Recovery Payment by Results Pilot Programme Final Report. The University of Manchester.

- Monitoring the progress of the programmes in achieving the outcomes.

LASARS evaluations were led partly by commissioners, partly by providers and local authorities.⁵⁵⁹ Approaches to LASARS operations varied across the eight sites and included: fully independent services; services operated by dedicated staff but within provider settings; services operated by dedicated staff but managed by the local Drug and Alcohol Team. This model was altered in the local pilots, such that some took very different approaches (e.g. some did not have an independent LASARS).

The outcomes upon which payment was conditioned fall under the following domains:

- Drug and alcohol addiction recovery;
- Offending;
- Health and wellbeing.

A fourth domain concerning employment was initially included, but was later eliminated from the payment scheme. For each domain, specific indicators were identified. These are listed in the table below.

Table 57. Indicators for each outcome domain

Outcome domain: Drug and Alcohol addiction recovery	
Interim/initial	<ul style="list-style-type: none"> • Improvement in drug and/or alcohol use (i.e. consumption reduced by a statistically significant level) for all presenting substances at any two Treatment Outcomes Profile (TOP) reviews within the last 12 months. No more than 20% of the total value assigned to this outcome domain can be placed on this initial outcome (see above about the possibility for pilot areas to set the price and weighting of each outcome). Where a client exits treatment before 12 months, the appropriate exit TOP will be used instead of a review TOP. • Abstinence from all presenting substances at any two TOP reviews in the last 12 months. (Where a client exits treatment before 12 months, the appropriate exit TOP will be used instead of a review TOP). <p>Planned exit from the treatment component of the recovery journey, free from drug(s) of dependence including alcohol, and including abstinence from heroin and crack cocaine.</p>
Final	Successfully discharged from treatment (free of drug(s) of dependence) and does not re-present in either of the treatment systems or in the criminal justice system in the following 12 months.

Outcome domain: Offending	
Interim	No proven offending during a six month period from the initiation of the recovery intervention by a provider.
Final	No proven offending during a 12 month period from the initiation of the recovery intervention by a provider
Final (cohort)	Reduction in average offending of cohort compared with the baseline, calculated and paid quarterly.

Note: Pilot areas were able to choose whether to use a cohort or individual measure. The majority adopted a group measure to minimise the risk of paying for outcomes that might have occurred anyway. Cohorts were made up of those individuals in and recently discharged from treatment.

⁵⁵⁹ Donmall, M., Sutton, M. et al. (2017). 118-119).

Outcome domain: Health and wellbeing	
Interim	<ul style="list-style-type: none"> Injecting: Of those injecting at the start of treatment, those who reported 0 days of injecting at any two review TOPs within the last 12 months. No Fixed Abode (NFA) / Housing problem: Of those NFA or with a housing problem at the start of treatment, those who no longer had any housing problems at any two review TOP where these were during the last 12 months, or at their exit TOP ('no' to both housing questions). Hep B Vac: Of those eligible, those that had appropriately completed a course of Hepatitis B vaccinations within the previous 12 months. <p>Health and wellbeing: Client achieves a normative quality of life score at any two TOP reviews, where these were during the last 12 months.</p>

Besides nationally decided outcomes, local pilot sites had the opportunity to insert **local indicators** for outcomes. Locally decided outcome indicators could differ slightly from those decided nationally in the Drug Recovery Strategy, in order to maintain local specificities, avoid cash problems in providers, avoid uncertainty relating to the service, and also to maintain a link with performance. The percentage of local indicators in the pilot sites varies between 0% (sites B,F,G) and 50% (site A) of the total indicators used at the pilot sites. It was not possible to retrieve local indicators using desk research.

Other evaluations

Impact evaluation - qualitative analysis

A qualitative analysis was conducted as part of the impact evaluation of the scheme. As part of this, 201 interviews were conducted with commissioners, providers (senior managers), practitioners, carers and services users. The data collected were analysed using thematic analysis.

Impact evaluation – difference-in-differences

The impact evaluation used a quasi-experimental design, namely a difference-in-differences approach. Data were collected from the National Drug Treatment Monitoring System (NDTMS) dataset.

All data for the impact evaluation were taken from the NDTMS dataset. NDTMS provides detailed data on people receiving structured treatment for drug and alcohol misuse in England, and is used to report on alcohol treatment activity, drug treatment activity, and young people in specialist drug and alcohol services. The cohort used in the impact analysis consisted of adult clients in contact with structured treatment services over four years, from 1 April 2010 to 31 March 2014.

The impact evaluation considered the following indicators for each outcome domain.

For drug clients:

Outcome domain 1: Drug addiction recovery	
Indicators	<ul style="list-style-type: none"> Abstinence rates Cessation of injecting Injecting at review

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

Outcome domain 1: Drug addiction recovery	
	Presence of housing problems
	Treatment completion within 6 months
	Treatment completion within 12 months
	Re-presentation for treatment
	Rate of re-presentation
	Unplanned discharge within 6 months
	Unplanned discharge within 12 months
	Retention in treatment within 6 months
	Retention in treatment within 12 months

Outcome domain 2: Offending	
Indicators	Number of recorded crimes within 6/12 months

Outcome domain 3: Health and wellbeing	
Indicators	Mortality rate (drug-related poisoning) Mortality rate (non-drug related poisoning)

For alcohol clients:

Outcome domain 1: Alcohol addiction recovery	
Indicators	Successful completion within 6 months
	Successful completion within 12 months
	Rate of completion
	Rate of completion and no re-presentation within 6 months
	Rate of re-presentation
	Rate of re-presentation after completion
	Unplanned discharge within 6 months
	Unplanned discharge within 12 months
	Retention in treatment within 6 months
	Retention in treatment within 12 months

Outcome domain 2: Offending	
Indicators	Offending (rate of recorded crimes per person within 6 and 12 months)

Outcome domain 3: Health and wellbeing	
Indicators	Mortality rate (alcohol-disease related) Mortality (not related to alcohol-disease) Presence of housing problems

Economic evaluation – difference-in-differences

An *economic evaluation* was implemented that also relied on a quasi-experimental approach using difference-in-differences. This economic evaluation relied on a monetisation of the following indicators:

- The volume of individuals treated for substance misuse via structured treatment, and the associated treatment costs for these individuals.
- The volume of recorded crimes committed by individuals in structured treatment for substance misuse, and their associated costs.
- The volume of drug-related A&E attendances and hospital admissions, and the costs associated with these attendances and admissions.

To examine the impact of PbR on the volume and cost of treatment, data were collected from NDTMS for the financial years 2010-11, 2011-12 and 2013-14. For the analysis of the volume and costs of recorded crimes, evaluators exploited data from Police National Computer database linked to NDTMS for the financial years 2010-11, 2011-12, and 2012-13. Anonymised patient-level data from the Hospital Episode Statistics were used for both hospital admissions and A&E attendances during the financial years 2009-10 to 2013-14.

Barriers to and enablers of the overall evaluation process

The **initial co-deciding phase** of the programme is considered a strength, since it enabled service providers to express their viewpoints concerning outcome indicators, and to push for the presence of interim outcomes and in-treatment indicators. Regarding the impact evaluation, it is important to mention the high level of internal validity provided by the difference-in-differences methodology, thanks to its quasi-experimental design.

Setting caps and floors relative to every outcome indicator and in relation to **patient complexity, rather than precise figures, has been an effective design feature** implemented at all eight pilot sites. This contractual arrangement is especially useful whenever there is uncertainty about what level of outcome is possible within a fixed budget.

560

The main barriers to and weaknesses of the measurement process that emerged from the desk research concern the **costs associated with measurement**, which were greater than expected for the service providers and local authorities. The service providers also highlighted challenges concerning **data availability and a lack of skills among staff** in retrieving and using these data. A recurring issue for local authorities was the development and management of data systems. For the impact evaluation, difficulties were encountered in **aggregating heterogeneous data from different sites**. Finally, the evaluators noted that the presence of LASARS, which were supposed to monitor and validate the

measurement process, often worsened the actual accessibility of treatment, weakening entire treatment effects.

Table 58. Summary of the measurement process in the Drug and Alcohol Recovery Pilots

Measurement methodology	
Methods related to payment mechanisms	Data self-reported by patients during Treatment Outcome Profile (TOP) reviews, monitored and validated by LASARS
Other evaluations	Impact evaluation – qualitative analysis Impact evaluation – difference-in-differences Economic evaluation – difference-in-differences
Methods using experimental or quasi-experimental design	Yes
Control group	Yes (in the difference-in-differences)
Causality of impact	Causality of impact was assessed using a quasi-experimental design. However, the evaluation results found that the impact of the intervention in the pilot areas was not statistically significant compared with non-pilot sites.
Enabling factors and strengths of the evaluation process	<ul style="list-style-type: none"> • Co-deciding phase • Quasi-experimental design
Barriers to and weaknesses of the evaluation process	<ul style="list-style-type: none"> • Costs of measurement to service providers • Data management and availability • Heterogeneous data from different sites <p>LASARS may have worsened actual accessibility of treatment</p>
Governance of measurement	
Independence of evaluator	Yes
Evaluator	LASARS University of Manchester; Rand Europe; Birkbeck, University of London; and User Voice

1.13.4. Effectiveness

Achievement of outcomes

No information is publicly available about the achievement of the outcomes relating to payments in the pilot sites. This was confirmed by all our interviewees.

Nevertheless, the commissioner from Lincolnshire who was interviewed provided some performance measures. These mainly showed a gradual increase in performance. However, the commissioner noted: “If you run the same reports for non-pilot areas, they also show a similar increase. This is probably due to the change in the government stance in 2010 to move services from a maintenance to a recovery-focused model more than the

fact that PbR was effective. If I remember correctly, prior to the pilot we were in the middle of the national table for performance and at the end of the pilot we were still in the middle. Given the additional resources required for the PbR this makes it **more effort for the same performance.**"

Benefits and drawbacks of SOC compared with TF

Most of the stakeholders agreed that the main benefit of the pilot, compared with previous services, was its greater **recovery orientation**. Particularly with regard to drug treatment, the outcomes-based scheme allowed a focus on full abstinence rather than drug substitution.⁵⁶¹ The provider interviewed confirmed that this service was provided before the launch of the SOC scheme, but that the service itself changed profoundly as a result of PbR. In particular, before the 2010 National Drug Strategy, drug and alcohol services did not **focus on full recovery**, and did not emphasise **local authorities' involvement**. These features of the programme were seen as the main added value of the pilot.

The perception of users was also in line with the provider's perspective. In fact, the 152 users interviewed as part of the process evaluation of the pilots confirmed a positive view of the pilot among clients, due to perceived **higher recovery-focused care** by key workers.⁵⁶²

On the other hand, the process evaluation noted that this orientation towards full recovery and the emphasis on the attainment of targets came at the expense of a **focus on the service users' overall experience of the treatment process**⁵⁶³, which runs the risk of simplifying a complex social problem. Moreover, because patients not re-presenting⁵⁶⁴ after they had been discharged from the programme was evaluated as part of the outcomes, providers tended to retain patients who might be at risk re-presenting themselves for treatment, which led to some participants refusing treatment entirely.⁵⁶⁵

Among other benefits, the pilot allowed greater **flexibility** in service delivery. The increased focus on full recovery led to some services developing new approaches and improving areas that had historically been considered weak. Evidence from the process evaluation, collected via interviews with practitioners and service managers, suggested that providers had expanded the range of services they offered and had sharpened their focus in areas of previously inadequate provision.⁵⁶⁶

Scalability and replicability

No attempts were made to replicate or scale the scheme. However, in some of the pilot areas a PbR model was maintained, although the contract and design of the schemes changed. For instance, in Lincolnshire, **the new programme tracks group outcomes for cohorts of participants rather than for each individual**. This substantially reduces the administrative burden. The commissioner explained:

"The cohort based system requires a lot less resources than individual payments, due to the level of data required. In the UK we have the National Drug Monitoring System (NDTMS). It monitors performance of substance misuse services across the country. Every provider needs to comply with the minimum data set, and reports are provided monthly or quarterly to commissioners and providers. The cohort system

⁵⁶¹ Interview with the evaluator

⁵⁶² Interview with the evaluator

⁵⁶³ Donmall et al. (2017), p.50

⁵⁶⁴ Not re-presenting refers to drug users not presenting themselves back for treatment.

⁵⁶⁵ Mason et al. (2015), p.2

⁵⁶⁶ Donmall et al. (2017), p.67.

uses this data to monitor performance, so we can look at our providers but also make comparisons with other similar areas to monitor how well our services are achieving outcomes. This does not require any additional resources from us locally, other than reading and interpreting the data. The individual PbR system used in the pilot tracked individual clients with a combination of identifying criteria; this data was then tracked for entering treatment, engaging with sessions, changes in problematic behaviours, ceasing of injecting, etc. It even tracked discharges who may have presented in areas outside of Lincolnshire to monitor longer-term recovery. This was incredibly resource-intensive and the algorithms required to generate the reports made my eyes water... So, to sum up: with cohort PbR, if 100 clients are discharged, I need to monitor one set of numbers, with individual PbR, I'd need to monitor 100 sets of numbers. If you then [multiply] that by 7-8 outcomes, it starts adding up.”

1.13.5. Efficiency

Intervention costs

According to the evaluation, which compared pilot and non-pilot areas (note: the majority of non-pilot areas funded interventions through traditional financing, but some also experimented with PbR contacts – see Section 1.13.2), **treatment costs per client increased significantly following the introduction of PbR.** Treatment costs for primary drug users were initially lower in the pilot areas. These costs increased in non-pilot areas by 2013-14, but increased even more (by 11%) in the pilot areas. There were similar increases in treatment costs for primary alcohol users, but the results were less stable.⁵⁶⁷

The PbR funding regime introduced unexpected costs such as higher-than-anticipated clinical, managerial and data monitoring expenses, as well as set-up and transitional costs. Moreover, costs also increased due to a greater focus on patients facing complex and multiple issues.⁵⁶⁸

It is not possible to compare costs between the Drug and Alcohol Recovery Pilots and the service provided by the Lincolnshire local authority prior to the pilots, because there is limited information on intervention under the latter. With regard to the Drug and Alcohol Pilots, as shown in the figure below, the total costs varied widely between Drug Addiction Teams (DATs): while most teams spent under GBP 5 million on patients with drug addiction between April 2012 and March 2014, there were a few instances in which costs exceeded GBP 25 million. Similar variation can be observed among DATs working with patients with alcohol addiction.⁵⁶⁹ Per-capita costs also varied substantially, although most DATs spent less than GBP 2,000 per client.⁵⁷⁰

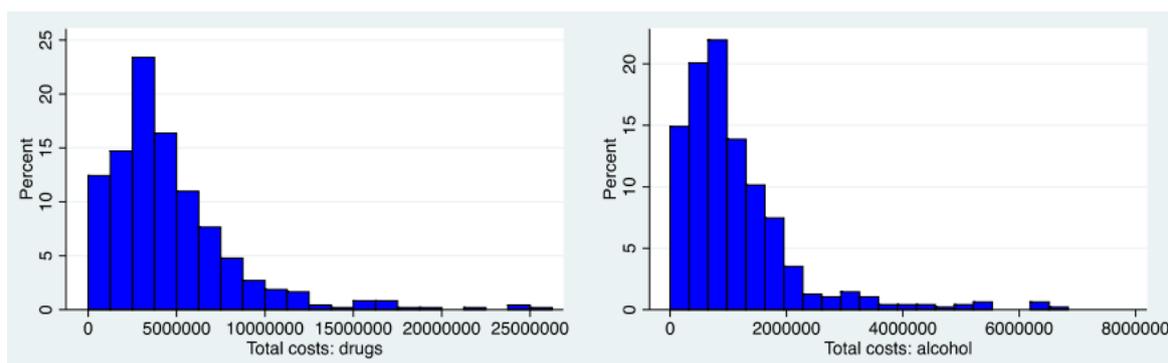
⁵⁶⁷ Donmall et al. (2017), p. 119.

⁵⁶⁸ Donmall et al. (2017), p.51, p. 87.

⁵⁶⁹ Donmall et al. (2017), p. 142. Please note that the total cost is not available.

⁵⁷⁰ Donmall et al. (2017), p. 141. Please note that the average cost per client is not available.

Figure 10. Total costs for DATs



Source: Donmall et al. (2017), 142.

Based on the impact assessment, **the evaluator interviewed argued that costs associated with the PbR scheme far exceeded its benefits.**⁵⁷¹ Furthermore, the Lincolnshire LA commissioner claimed that PbR schemes are able to provide cost savings only in the long term, while generating higher costs in the short term due to investment in data management structures. This is why it is necessary to have a sufficient overall budget at the start of the scheme, in order to compensate for uncertainties and restrictions during its first steps.⁵⁷²

Operational costs

No comparison is possible between the operational costs of the PbR scheme and its TF equivalent in the case of Drug and Alcohol Pilots, due to limited information on the TF equivalent. Therefore, below we describe the information on the PbR programme only. Nevertheless, **based on the interviews, the operation of the PbR scheme appears to have been more costly than traditional contracts**, mostly due to the heavy administrative burden associated with outcome tracking and reporting.

Set-up costs differed widely between the heterogenous sites at which the Drug and Alcohol Pilots were implemented (see the table below). In most cases, the costs incurred by commissioners were one-off and related to the establishment of databases, LASARS and transfer of undertakings (TUPE) costs. Unfortunately, the total cost of the pilot at each site is not available, so we cannot estimate what share of the total contract value set-up costs amount to. However, it is important to note that there was a six-month delay in issuing payments to providers, mostly due to operational burden involved in setting up the contract.⁵⁷³

Table 59. Set-up costs for the Drug and Alcohol Pilot sites

Site	Systems	LASARS	Misc.	Total
A	N/A	N/A	N/A	N/A
B	35	219,765	1,119	36,119
C	N/A	N/A	N/A	N/A
D	0	233,513	0	233,513

⁵⁷¹ Interview with the evaluator (2020).

⁵⁷² Interview with the commissioner (2020).

⁵⁷³ Interview with the commissioner (2020).

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

Site	Systems	LASARS	Misc.	Total
E	0	569,412	0	569,412
F	70	0	0	70
G	0	0	0	0
H	0	0	454,812	454,812

Source: Adapted by Polimi based on Donmall et al. 2017, 63.

To provide a brief description of the scheme’s set-up: the Department of Health decided to experiment with PbR in Drug and Alcohol treatment programme in a way that was consistent with the new Drug Recovery Strategy 2010. Previously, the commissioning of drug treatment services had focused on retention in treatment as the principal measure of effectiveness, and service providers were paid using ‘block’ or ‘activity’ contracts. The pilot in eight areas was a pillar of the 2010 National Drug Strategy, which prioritised the recovery of service users from their dependence on drugs or alcohol. After the central government departments and Drug and Alcohol Action Teams (DAATs) across England selected the eight areas for piloting, a **co-deciding phase** began, involving a variety of local stakeholders. By issuing invitations to this co-decision procedure, the commissioner tried (but mostly failed) to widen and diversify the marketplace of providers,. Negotiations between commissioners and providers were lengthy because the commissioner pushed for conditioning 100% of payment on outcomes, whereas providers appeared more risk adverse, arguing for a balance between outcome- and output-based funding. Providers were worried by the lack of details about the outcomes sought, and by uncertainties surrounding the financial implications of the PbR. Providers stressed the perceived lack of evidence informing the development of PbR, anxieties about the emerging models, and the inconsistency of the PbR scheme with public health literature and scientific results.

While information on **management costs** is not available for all the sites, the audit service commissioned by Lincolnshire LA to ensure tariffs were not being manipulated cost GBP 253,400 per annum.⁵⁷⁴ The commissioner noted that PbR requires dedicated teams of analysts to monitor and oversee the intervention.

1.13.6. Design of the scheme

Design features

Most of the **negative feedback** concerned the complex nature of the programme’s evaluation. Not only was the set-up and use of LASARS costly for reporting individual outcomes, but the number of indicators monitored was at times overwhelming (see Section 1.13.4).⁵⁷⁵

Furthermore, although during the set-up process commissioners pushed for 100% outcome-based payments, after the pilots, commissioners emphasised **the importance of not having contracts 100% based on outcomes**, but instead implementing a gradual approach towards PbR. The commissioner interviewed also stressed that the **PbR approach is not suitable when the available budget for the intervention is limited**: “I would still recommend services not be commissioned with a full individual PbR system; there are some benefits with a cohort based approach, but only if capacity allows for

⁵⁷⁴ Interview with the commissioner (2020).

⁵⁷⁵ Donmall et al. (2017), p. 122; ICF (2015), p. 104

performance improvements. If services are short on funding, I would recommend a block contract approach, as PbR can have a negative effect on performance if no efficiency savings are possible, as providers will not commit to spend funding they are not assured of getting or if targets are too high.”

Impact on third-sector organisations and social enterprises

The service provider interviewed (who represented a charity) was quite critical towards the scheme, and would not recommend the use of PbR. The intervention was deemed to have been less effective due to the budgetary uncertainty linked with the use of PbR. Such uncertainty affects the process of hiring staff : unwilling to take the risk, the provider might hire fewer staff, making intervention success more difficult to achieve. However, the provider noted that **PbR could be exploited as a premium mechanism**, rather than as a modality upon which the entire financing would depend. The provider suggested that it could be a mechanism under which the amount of payment relating to the outcomes achieved would be paid in addition to the amount of money that providers require to cover their costs.

1.14. Transforming Rehabilitation (UK)

1.14.1. Background

In 2014, the UK government's Transforming Rehabilitation (TR) reform changed the way that **adult probation services** were organised and delivered. The **main goals that guided the development of the Transforming Rehabilitation programme**, the resulting PbR scheme designed by the Ministry of Justice (MoJ), were as follows:

- opening up the market to a diverse range of rehabilitation providers from the private and voluntary sectors;
- encouraging innovation, paying providers by results for reduced reoffending;
- extending statutory rehabilitation in the community to approximately 45,000 additional offenders a year who are released from short prison sentences of less than 12 months;
- reorganising the prison estate to provide continuous support from custody into the community (so-called 'Through the Gate' (TTG) support); and
- creating a new public-sector National Probation Service (NPS) to manage high-risk offenders.

The reform created 21 privately owned Community Rehabilitation Companies (CRCs) and the public-sector NPS. The CRCs are owned and run by the successful bidders in an open competition, and deliver services under contract from the National Offender Management Service. Eight, mainly private-sector, suppliers won the bids and worked under contracts managed by Her Majesty's Prison & Probation Service. The contracts were intended to run until 2021/2022.

CRCs manage those offenders who present a low or medium risk of serious harm in the community. These include offenders sentenced to Community Orders (COs), Suspended Sentence Orders, and those who are subject to licence conditions or supervision requirements. CRCs are requested to deliver innovative rehabilitative support and mentoring to offenders. CRCs should provide services in relation to education, substance misuse, housing and employment, but they have a lot of freedom in terms of carrying out

work to drive down reoffending rates.⁵⁷⁶ Meanwhile, the NPS directly manages offenders who pose a high risk of serious harm to the public.

The re-organisation also introduced the idea of 'resettlement prisons', under which prisoners are transferred to an establishment close to their home address for (at least) the last three months of their sentence. While in the resettlement prison, the offender is provided with rehabilitation and TTG resettlement services provided by the CRCs and the NPS.⁵⁷⁷

Payments to the CRCs include both fees for services provided and a PbR component. PbR represents around 10% of total predicted payments to the CRCs. The fee-for-service component primarily covers the CRC's operating costs for mandated activities (e.g. delivering the sentence of the court and licence conditions, TTG services, and any mandatory activity days required by the court).

Overall payment is based on a weighted annual volume of offenders assigned to each CRC, with a proportion of this payment dependent on the CRC's performance in reducing reoffending in comparison with a historical baseline level. Specifically, performance is measured in terms of a statistically significant reduction in reoffending compared with the 2011 baseline (pre-reform reoffending rate).⁵⁷⁸ The PbR payment was designed to pay the provider the profit component of their total cost model. Two reoffending measures are used to assess the performance of the CRCs (and the NPS):

- the binary rate: proportion of offenders who reoffend;
- the frequency rate: the average number of reoffences per re-offender.

Statistically significant points (SSPs) were set to identify the appropriate minimum changes to the reoffending rate that would need to take place before any PbR adjustments could be applied to the payment received by the provider⁵⁷⁹. The upper SSP ('threshold level') is the minimum reduction in reoffending rates expected by the MoJ; the lower SSP represents the level of reoffending rates that is significantly higher than the baseline and is therefore considered unacceptable ('Deduction Level').⁵⁸⁰

The unit payment for achieving the binary metric is GBP 4,000 per offender who desists from reoffending during the 12 months after they have been discharged⁵⁸¹. The unit payment for achieving the frequency metric is GBP 1,000 per reoffence avoided.

The payment arrangement in relation to the binary measure has been designed as follows (see also the figure below)⁵⁸²:

- If reoffending rates remain at their historical level or fall between the baseline and the threshold level (see above) in a given contract package area (the 'baseline'), providers receive no additional payments on top of the fees for service. These fees are paid in 12 equal payments, made monthly in arrears.
- If the reoffending rate is below the historical baseline and is:
 - at the threshold (point C in the figure below), some profits (the 'PbR Foundation Payment') are paid in addition to the fees for service. These do not equal the full level of profits the service provider would have received under a standard (non-PbR) contract;
 - improved beyond the threshold level (point B in the figure below), additional payments are made for each extra non-offender. Only at this point can the

⁵⁷⁶ User Voice. (2016). Transforming Rehabilitation – A service user's perspective.

⁵⁷⁷ MoJ. (2014). Target Operating Model: Version 3. Rehabilitation Programme.

⁵⁷⁸ NAO. (2017). Investigation into changes to Community Rehabilitation Company contracts.

⁵⁷⁹ MoJ. (2013). Rehabilitation Programme – Payment Mechanism Straw Man.

⁵⁸⁰ MoJ. (2014). Transforming Rehabilitation Programme Payment Mechanism.

⁵⁸¹ MoJ. (2013). Rehabilitation Programme – Payment Mechanism Straw Man, p. 13.

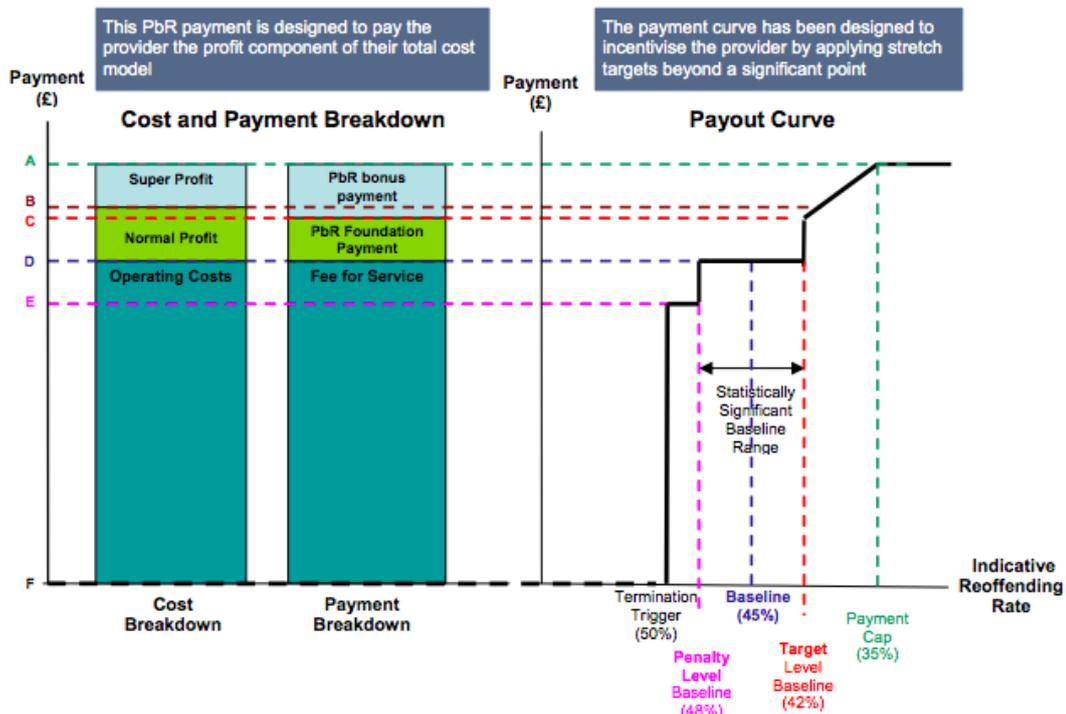
⁵⁸² MoJ. (2014). Transforming Rehabilitation Programme Payment Mechanism.

provider potentially reach the ‘full’ payment equivalent to the payment under a non-PbR model;

- improved substantially beyond the threshold level, the provider receives a bonus payment, up to the maximum allowed by the cap (point A in the figure below).

If the reoffending rates are above the historical baseline, **deductions** are made from fees for service payments.

Figure 11. payment arrangements for Transforming Rehabilitation



Source: MoJ. (2013). Rehabilitation Programme – Payment Mechanism Straw Man, 9.

In relation to the ‘frequency of reoffending’ metric, no payment or deduction thresholds were proposed. In relation to the frequency metric, providers would be rewarded for any reduction in reoffending according to this metric; likewise, financial deductions would be made for poor performance. A termination trigger was set at a confidence interval of 99% from the frequency reoffending baseline. There system includes a binary hurdle, according to which payments are only made for improvements in the frequency rate if the adjusted binary rate is lower than the 2011 baseline reoffending rate.

The contract year comprises four quarterly cohorts for which reoffending is measured. These four quarters are then aggregated to form an annual cohort, for which reoffending is measured and PbR potentially paid on the annual binary metric and the frequency metric.

The time lag for the binary reoffending metric to become measurable is 26 months. For the frequency metric, the time lag is 35 months. Therefore, the PbR Foundation Payment is designed to ease the provider’s cash flow. Moreover, to ensure that the MoJ retains the volume risk related to PbR, payments are adjusted to take the account of volumes. The first frequency payments were based on the 2015/2016 annual cohort consisting of the October to December 2015 and the January to March 2016 cohort periods only.⁵⁸³

⁵⁸³ MoJ. (2018). Final and Interim Proven Reoffending statistics for the Community Rehabilitation Companies and the National Probation Service.

While reoffending has decreased overall, the CRCs have not achieved the Ministry's targets. The MoJ expected the CRCs to reduce reoffending by 3.7 percentage points over the life of the contracts. However, by March 2017, mid-way through the reform, the overall reduction in the proportion of proven reoffenders since 2011 was 2.5 percentage points. Between 2011 and March 2017, there was a 22% overall increase in the average number of reoffences per reoffender, and just six out of the 21 CRCs consistently achieved their targets for reducing reoffending.⁵⁸⁴

Various **assessments** were implemented by the National Audit Office, Her Majesty's (HM) Inspectorate of Probation, and the House of Commons Justice Committee. In July 2018, the Justice Secretary acknowledged that the quality of probation services being delivered was **falling short of expectations** and announced that the Ministry would terminate its CRC contracts 14 months early in December 2020, at additional cost to the taxpayer. The CRC contracts have faced difficulties as a result of lower than expected numbers of cases, higher fixed costs, and an increasing trend in the frequency of reoffending, undermining the standard of services.⁵⁸⁵

Table 60. Summary of Transforming Rehabilitation

Personally targeted social service	Social exclusion
Target population	Offenders serving short sentences
SOC scheme type	Mixed PbR
Payment model	Frequency and Hybrid
Cost	GBP 2.3 billion maximum forecast payments to Community Rehabilitation Companies (CRCs) from 2014 to December 2020 (as of August 2018)
Commissioner(s)	Ministry of Justice
Social service provider(s)	21 Community Rehabilitation Companies: One CRC is owned by a consortium of voluntary, private and public organisations (Working Links); all others are owned by private companies (Sodexo, Purple Futures; The Reducing Reoffending Partnership, EOS Works Ltd, MTCNovo, Seetec).
Investor(s)	N/A
Intermediary	N/A
Evaluator(s)	<u>Linked to payments:</u> Prison, Probation and Reoffending Statistics Division of the MoJ <u>Not linked to payments:</u> National Audit Office (which commissioned a report prepared by User Voice); HM Inspectorate of Probation; House of Commons Justice Committee

1.14.2. Equivalent TF scheme

The National Probation Service for England and Wales is a **statutory criminal justice service**, responsible for the supervision of offenders in the community and the provision of reports to the criminal courts to assist them in their sentencing duties. The Transforming Rehabilitation Reform, which included a PbR component, was introduced between 2014 and 2015, with the aim of changing the scope and structure of community and prison-based probation and rehabilitative services. Therefore, it is possible to make comparisons between

⁵⁸⁴ MoJ. (2020). Final and interim proven reoffending statistics for Community Rehabilitation Companies and the National Probation Service. January to March 2018 and 2017/18 annual cohort; NAO. (2019). Transforming Rehabilitation: Progress Review.

⁵⁸⁵ NAO. (2019). Transforming Rehabilitation: Progress Review.

the intervention in place before the implementation of the reform, and the services delivered since 2015 under the TR reform.

Prior to June 2014, probation services in England and Wales were delivered by 35 self-governing Probation Trusts, each working under the direction of the National Offender Management Service.⁵⁸⁶ Probation trusts worked with offenders who were serving a community-based sentence, or prisoners who had been released from custody.⁵⁸⁷

The reform introduced a legislative provision that required offenders serving short custodial sentences (i.e. prison sentences of under 12 months) to receive 12 months of compulsory post-sentence supervision. Prior to this reform, probation services were not delivered to prisoners serving short sentences. In addition, a set of new services called ‘Through the Gate’ Resettlement Services were introduced. These new services were rolled out in prisons with the aim of preparing prisoners for release and resettlement, and increasing their prospects of leading a better life. In this way, offenders were given continuous support by one provider while transferring from custody into the community.

Table 61. Comparability of Transforming Rehabilitation with Probation Trusts

	Same as SOC	Comments
Personally targeted social service	Yes	Social exclusion
Target population	Yes/No	The reform required compulsory post-sentence supervision to be given to offenders serving short custodial sentences (i.e. prison sentences of under 12 months)
Number of participants	No	TF: In 2012, Probation Trusts were responsible for a caseload of around 225,000 offenders ⁵⁸⁸ SOC: CRCs in 2016 supervised approx. 140,750 offenders ⁵⁸⁹
Location	Yes	Both were rolled out nationally
Cost	No	SOC: actual payments from 2014 to 2015: GBP 0.1 billion; actual payments from 2015 to 2016: GBP 0.6 billion (data include the fee for service, payment by results, fee for work carried out on behalf of the National Probation Service) ⁵⁹⁰ TF: GBP 867 million in the period between 2011 and 2012; GBP 853 million from 2012 to 2013 ⁵⁹¹
Commissioner(s)	Yes	National Probation Service/MoJ

⁵⁸⁶ MoJ. (2013). Transforming Rehabilitation: A revolution in the way we manage offenders, Cm 8517.

House of Commons Justice Committee. (2013). Crime reduction policies: a coordinated approach? Interim report on the Government's Transforming Rehabilitation programme. Twelfth Report of Session 2013–14.

<https://publications.parliament.uk/pa/cm201719/cmselect/cmjust/482/48205.htm>

⁵⁸⁷ NAO. (2014). Probation: landscape review.

⁵⁸⁸ NAO. (2014). Probation: landscape review, p. 10.

⁵⁸⁹ NAO (2019). Transforming Rehabilitation: Progress Review – Transcript.

⁵⁹⁰ NAO (2019). Transforming Rehabilitation: Progress Review – Transcript, p.32 (Figure 32).

⁵⁹¹ NAO. (2014). Probation: landscape review, p. 12.

	Same as SOC	Comments
Social service provider(s)	No	SOC: 21 Community Rehabilitation Companies TF: 35 self-governing Probation Trusts

1.14.3. Outcomes measurement

The outcomes of Transforming Rehabilitation were assessed using the reduction in the reoffending rate in each of the areas managed by a Community Rehabilitation Company. This rate has been computed using data taken from many administrative IT systems, then compared against a historical baseline (namely the most recently available rate in a given area before the reform).

Other evaluations were carried out by the National Audit Office, User Voice and Her Majesty's Inspectorate of Probation, as outlined below.

Measurement of outcomes relating to payment mechanisms

Measurement of the outcomes achieved was conducted using a non-experimental method, namely a **pre-post comparison only of users receiving the treatment**, based on secondary data.⁵⁹²

Evaluation of the outcomes upon which payment was conditioned was carried out using a historical baseline comparison, based on statistics regarding proven reoffending. The process of data collection and analysis is outlined below.

The **data required to measure proven reoffending came from a range of sources** (including prison data, probation data and criminal records from the Police National Computer); and from a number of agencies (Her Majesty's Prison and Probation Service, the Youth Justice Board and the Home Office). These figures were derived from existing administrative IT data systems. The process then involved matching the data on prison discharges and court order commencements with the Police National Computer database.⁵⁹³

The data source for offender starts in each PbR cohort changed between the procurement process for CRC contracts and the measurement of outcomes for the first PbR cohort. It is estimated that changing the data source leads to an apparent increase in reoffending rate of around 0.5 percentage points.⁵⁹⁴

Data collection and analysis were organised in cohorts of offenders. A single cohort was constructed on a three-month basis in each area for the binary metric, and on a 12-month basis in each area for the frequency metric. The binary metric was also subject to an annual top-up calculation based on a 12-month cohort. An offender entered the cohort if they were released from custody or were sentenced to a community order or a suspended sentence order within the specified period of time. The final reoffending results for the CRC PbR offender cohorts were therefore based on a one-year proven reoffending measure.

Cohorts were not segmented according to the type of sentence or offender, although statistics for each type of sentence and offender were collected for monitoring purposes. Offenders who appeared multiple times in a cohort were only included once.

⁵⁹² MoJ (2016). Response to consultation on changes to proven reoffending statistics.

⁵⁹³ MoJ (2019). Guide to proven reoffending statistics.

⁵⁹⁴ MoJ (2016). Response to consultation on changes to proven reoffending statistics.

Offences were counted using the National Statistics ‘proven reoffending’ measure. One-year **proven reoffence** was counted as any offence committed within a 12 month follow-up period, which then attracted a court conviction or caution within that 12 month follow-up period (or within a further six month waiting period, to allow for cases to work their way through the courts). Offences were counted as proven reoffences if they met all of the following criteria: they were recordable; they were committed in England or Wales; they were prosecuted by the police.

Performance in terms of reducing reoffending, on both the binary and frequency measures, was assessed against a baseline re-conviction rate. The ‘baseline’ re-conviction rate was the expected level of reoffending in each area, as determined by the most recent pre-reform re-conviction data. The baseline used was from the year 2011. The **PbR binary results** were adjusted using the Offender Group Reconviction Score (OGRS) score to take into account the fact that differences in the mix of offenders can affect the likelihood of reoffending, and thus impact binary reoffending rates. OGRS scores estimate the probability of reconviction for an offender with a given history of offending being, based on certain differences in offender characteristics (age, gender and criminal history). This calculation was used to standardise the mix of offenders in each cohort of a given CRC, to the 2011 mix for the same area.

In 2018, this method was changed:

- The baseline year, against which CRC performance in relation to the frequency of reoffending is compared, was changed. From 2018, all CRCs were compared against a 2015/2016 baseline, except the Merseyside CRC, which retained the 2011 baseline.
- The 12-month offender cohorts were changed into three-month offender cohorts. The change resulted in a greater proportion of prolific offenders and hence higher reoffending rates.

The **target** was set to be the proportional difference between actual reoffences and the expected number of reoffences. Statistically significant points (SSPs) were set that identified the appropriate minimum changes in reoffending rates. The higher SSP represented the minimum improvement in reoffending rates expected by the MoJ (Target Level Baseline); the lower SSP represented the level of reoffending rates significantly worse than the baseline and, therefore, considered unacceptable (Penalty Level Baseline). The MoJ specified thresholds consisting of an average three-percentage-point reduction (or increase) in reoffending rate compared to the baseline, at which it could be reasonably confident that an underlying improvement (or deterioration) had occurred among an annual cohort of offenders). These thresholds were set on the basis of an 80% confidence interval (meaning that, based on historical variation, there is an 80% probability that measured reoffending would fall within the thresholds if there had been no changes in underlying performance). Thus, there is a 20% probability that reoffending would fall outside this range; 10% above and 10% below. This helped to give MoJ confidence when making a payment that it was rewarding a genuine achievement in terms of reduced reoffending.

Two reoffending measures were used to assess CRC and NPS performance: the **binary rate** (the proportion of offenders who reoffend) and the **frequency rate** (the average number of reoffences per reoffender).

Table 62. Indicators for each outcome domain

Outcome domain: reduced reoffending	
Binary rate	Proportion of offenders who reoffend, computed as the percentage of offenders within a quarterly cohort who are convicted of an offence within a 12-month period. Offences were counted using the Office of National Statistics ‘proven reoffending’ measure.

Outcome domain: reduced reoffending	
	The expected number of reoffenders was calculated according to the following formula: OGRS-adjusted baseline reoffending rate x number of offenders.
Frequency rate	The average number of reoffences per reoffender was computed by measuring the average number of reoffences per reoffender committed within an annual cohort within a 12-month period. Offences were counted using the Office of National Statistics 'proven reoffending' measure. The expected number of reoffences was calculated according to the following formula: expected number of reoffenders x baseline reoffence frequency.

Other evaluations

Impact assessment - User Voice and National Audit Office⁵⁹⁵

User Voice, a charity led by ex-offenders working to reduce reoffending through collaboration, undertook a distinctly peer-led consultation across England to explore how current service users who were in receipt of community supervision understood the changes brought about by the Transforming Rehabilitation agenda, and what it meant for them personally.

The charity used both qualitative and quantitative methods of investigation, namely, it conducted five focus groups (involving a total 45 service users) and received 251 responses to surveys peer-distributed across four CRCs, of which 44.6% were supervised by CRCs and 13.6% by NPS. The focus groups and surveys explored users' experiences of probation as a whole (both CRCs and the NPS), and their perceptions of change over the preceding 12 months. They mainly evaluated how users perceived the quality of services before and after the reform.

User Voice investigated the following outcome domains: *relationships with probation officers; perceived role of probation (supportive agency, law enforcement agency or both); number of probation officers; level of contact in terms of frequency and barriers to contact (ease of access to appointments and contact prior to being released from prison); care service provision such as emotional support; and help with housing and employment.* Specifically, service users were asked whether they felt there had been any changes since the Transforming Rehabilitation agenda had been implemented.

Impact assessment - HM Inspectorate of Probation

HM Inspectorate of Probation conducted an impact assessment using several focus groups with staff from service providers (case managers) and semi-structured interviews with service users in different areas (CRCs and NPS). The scope of each of the Transforming Rehabilitation inspections was similar, but the detailed focus varied in order to explore specific areas of practice as the Transforming Rehabilitation programme progressed. Each of the inspections visited a different set of Local Delivery Units. HM inspectorate of Probation implemented six rounds of inspection from **April 2014 to May 2016 (Early implementation of Transforming Rehabilitation)**. It collected users' views on: *assignment, induction and first appointments, managing the offender, delivering the*

⁵⁹⁵ User Voice. (2015). Transforming Rehabilitation – A service user's perspective.

sentence, and reviewing the work with the offender. It then analysed *breach and escalation of cases initially managed by the CRC*. **HMI Probation's Quality & Impact programme**⁵⁹⁶, which commenced in **April 2016**, was designed to examine probation work in discrete geographical areas. The inspection focused on assessing how the quality of practice contributed to achieving positive outcomes for service users, and evaluating what impact had been achieved, since the current impact provided evidence of progress towards the long-term desistance of individuals. In particular, HMI Probation sought to report on whether reoffending had been reduced⁵⁹⁷, whether the public was protected from harm⁵⁹⁸, and whether individuals had abided by the terms of their sentence⁵⁹⁹. In April 2018, the HM Inspectorate of Probation introduced a new inspection framework, the Quality & Impact Inspection Model⁶⁰⁰, under which all probation providers were to be inspected annually and assigned overall ratings.

Experience of voluntary sector organisations

TrackTR, a study performed by the charity Clinks, focused on the voluntary sector's role in Transforming Rehabilitation (published in 2015, 2016, 2018)⁶⁰¹. Results from its surveys include self-reported levels of involvement with new probation providers, from which the voluntary sector receives funding, and the extent to which Transforming Rehabilitation had impacted their current funding to deliver services.

In-depth informal conversations with providers and policy makers were held over the course of the project to better understand the data received from the voluntary sector. In addition, six in-depth interviews were undertaken with six case study organisations.

Barriers to and enablers of the overall evaluation process

The main factor enabling the measurement process of those outcomes which triggered the payment lied in it being based solely on data collected through the **administrative IT system**. This did not pose **any data collection burden on service providers**.

Several weaknesses of the measurement process, specifically in terms of the **design of the metrics**, were highlighted by the evaluator.

First, the design of the method and metrics did not ensure the attribution of outcomes and did not consider deadweight, as many intervening factors were not taken into account. These include cuts in police activities, change in police behaviour, changes in the profile of offenders (for instance those coming via the courts tend to be those who are more likely to reoffend, which is not within the control of the CRCs), and other effects specific to the area in which the intervention was implemented. The use of OGRS did not entirely take account of all differences between offenders to be considered, with many factors that affecting the likelihood of reoffending, such as the number of offenders with a drug addiction, still not included in the analysis. This is particularly important for the frequency measure, which is not adjusted for the likelihood of reoffending that a caseload presents.

⁵⁹⁶ HM Inspectorate of Probation. (2018). Quality & Impact inspection: The effectiveness of probation work by the London Community Rehabilitation Company.

⁵⁹⁷ Specific questions at p. 63 of HM Inspectorate of Probation. 2016. Quality & Impact inspection. The effectiveness of probation work in Durham.

⁵⁹⁸ Specific questions at p. 64 of HM Inspectorate of Probation. 2016. Quality & Impact inspection. The effectiveness of probation work in Durham.

⁵⁹⁹ Specific questions at p. 65 of HM Inspectorate of Probation. 2016. Quality & Impact inspection. The effectiveness of probation work in Durham.

⁶⁰⁰ <https://www.justiceinspectorates.gov.uk/hmiprobation/wp-content/uploads/sites/5/2016/02/Quality-Impact-model.pdf>

⁶⁰¹ TrackTR. (2016). Change & Challenges. The voluntary sector's role in Transforming Rehabilitation

TrackTR. (2018). Underrepresented, under pressure under resourced. The voluntary sector in Transforming Rehabilitation.

TrackTR. (2015). Early Doors. The voluntary sector's role in Transforming Rehabilitation.

The MoJ is considering the use of other measures to better estimate the performance of CRCs, such as the number of offenders obtaining access to housing and Universal Credit, as well as outcomes relating to successful employment or mental health programmes. An alternative approach to solve this issue, proposed by the evaluator, is to **compare a CRC's performance with the average performance of all other CRCs**. If the CRC reduced reoffending significantly, they would receive a bonus payment.⁶⁰²

Second, the metrics and the binary hurdle created a **perverse incentive for the provider to focus on a binary measure**. Indeed, the binary hurdles introduce the risk of 'cherry picking' and/or the 'creaming and parking' of offenders, because a service provider might focus on those who are least likely to reoffend. The binary hurdle also penalised better-performing providers who met the frequency but not the binary measures.

Third, a problem with the PbR frequency measure is that **the measure excludes non-offenders**, thus reducing the binary rate of reoffending could hurt the frequency rate.

Finally, an issue relating to **data availability** emerged. Indeed, it took two years for data on proven reoffending to become available, and changes in reoffending could not be directly attributed to the interventions of CRCs, as they were also influenced by services such as support with housing, employment and substance misuse. Furthermore, the 2011 baseline was developed years before the implementation. It is fair to say that the context in which the CRCs were operating was different, and this undermined the validity of the PbR measure (for example, the caseloads today are different from those in 2011).

⁶⁰² <http://www.russellwebster.com/gtdpaymentmech/>

Table 63. Summary of the measurement process in Transforming Rehabilitation

Measurement methodology	
Methods related to payment mechanisms	Reduction computed by comparing the reoffending rate in a specific area, in terms of both binary and frequency metrics, to a baseline based on the most recent reoffending rate available pre-reform.
Other evaluations	Surveys, focus groups, and semi-structured interviews to assess the quality of services and their effects on service users.
Methods using experimental or quasi-experimental design	No
Control group	Historical aggregate baseline
Causality of impact	Causality of impact cannot be fully attributed because intervening factors have not been properly assessed. Comparison is made against historical data, and the metrics are adjusted to take account of the characteristics of the offender.
Enabling factors and strengths of the evaluation process	- No requirement for service providers to collect data
Barriers to and weaknesses of the evaluation process	- Several issues relating to the design of the metrics - Deadweight and attribution issues are not addressed - Long time span to achieve final measurement of outcomes
Governance of measurement	
Independence of evaluator	No
Evaluator	Statistical division of the MoJ National Audit Office (public agency) User Voice (consultancy) Her Majesty's Inspectorate of Probation (public agency)

1.14.4. Effectiveness

Achievement of outcomes

In terms of outputs, the number of offenders supervised under Transforming Rehabilitation increased by 17% (37,000) between January 2015 and September 2018. Between January 2015 and September 2018, the number of offenders recalled to prison for breaching their licence conditions increased from 4,240 to 6,240 (47%). Over the same period, the percentage of offenders recalled to prison who had received short sentences of less than 12 months increased from 3% to 36%.

The overall target expected by the government was not achieved. Both the public-sector National Probation Service (NPS) and privately owned Community Rehabilitation

Companies (CRCs) failed to meet some of their performance targets. On average, CRCs met 53% of their quarterly contractual targets by September 2018.

Moreover, the quality of the probation work carried out by CRCs was found to be poor by the independent Chief Inspector of Probation. Out of 13 CRCs inspected between December 2016 and March 2018, HM Inspectorate of Probation rated nine (69%) negatively for the quality of their work in reducing reoffending and protecting the public; and for abiding by the sentence of the courts, it rated five negatively (38%). All 13 NPS areas were rated positively for abiding by the sentence of the court, and 10 (77%) positively for protecting the public and reducing reoffending. HM Inspectorate of Probation gave 'amber/green' ratings (indicating that practices mostly complied with standards) in just five of its 37 CRC audits between February 2017 and October 2018 (14%). By February 2019, it had rated eight CRCs as 'requires improvement' and one as 'inadequate' under its new inspection framework, introduced in April 2018. It had inspected three NPS regions, which were rated as 'good'.

On average, it appeared that the NPS, which managed the same type of offenders as former Probation Trusts, was performing better than CRCs, which managed the new services for short sentence offenders.

The most recent Report of the Chief Inspector of Probation (March 2019) stated:

- "80% of Community Rehabilitation Companies (CRCs) inspected by HMI Probation have been rated as 'inadequate' for the implementation and delivery of probation supervision."
- "37% of magistrates indicated that they had less confidence community sentences are an effective alternative to custody now than they were under previous arrangements (2016 Magistrates Association survey)."⁶⁰³

Possible explanations as to why the number of re-offenses has not decreased, while the number of reoffenders has declined, include the following⁶⁰⁴:

- The observed reduction in reoffending rates was possibly due to police cuts. The police were charging fewer offenders and the crimes they committed were more likely to be more harmful than in the past. The police might have prioritised harmful and prolific offenders, and thus infrequent offenders of low-harm offences were not being picked up.
- The senior management teams in each CRC might have focused on reducing binary rates of reoffending and gave less attention to prolific offenders.
- The profile of offenders has changed, with prolific offenders representing a greater proportion of the cohort. Many of the macro trends point to a change in the types of offenders presenting at court and my experience suggests that probation services are unlikely to cause a large increase in frequency of reoffending on their own.⁶⁰⁵
- The probation profession has been diminished. There is a national shortage of qualified probation professionals and too much reliance on unqualified or agency staff. The critical relationship between the individual and the probation worker is not sufficiently protected in the current probation model.

Comparing the outcomes achieved by the TR reform to the situation before the reform, the reoffending rate for adult offenders serving court orders fell from 39.9% in 2003 to 34.1% in

⁶⁰³ <http://www.russellwebster.com/crcpbrjan18/>

⁶⁰⁴ <http://www.russellwebster.com/crcpbrjan18/>

⁶⁰⁵ http://www.russellwebster.com/crcpbrjan18/?elementor-preview&1518417048481#_ftn1

the year to March 2010.⁶⁰⁶ Nevertheless, the government considered it still ‘unacceptably high’.⁶⁰⁷

According to the independent Chief Inspector of Probation:

- 75% of the work undertaken by probation trusts was sufficient to keep the individual’s risk of harm to a minimum;
- 74% of the work undertaken by probation trusts was sufficient to make each individual less likely to reoffend;
- 79% of the work undertaken by probation trusts was sufficient to support effective compliance and enforcement;
- In 79% of high-risk-of-harm cases, all reasonable action was taken to keep risk of harm to others to a minimum;
- In 76% of all cases (and in 84% of high-risk-of-harm cases), breach action or recall was instigated on all occasions when required.⁶⁰⁸

In addition, The National Offender Management System performed its own performance assessment on the Probation Trusts. Performance was graded into one of four bands: 4 – exceptional performance; 3 – good performance; 2 – requiring development; and 1 – serious concerns. Performance was assessed in the three areas that best describe the work of probation: public protection, reducing reoffending, and sentence delivery. In the latest year for which Probation Trust Rating System results are available (2012-2013), all trusts were rated as having achieved at least level 3 – good performance. Five trusts achieved level 4 – exceptional performance.

However, the evaluator interviewed believes that the **overall failure of the reforms should not be attributed to the PbR mechanism, but rather to the negotiation process and to issues relating to payments.**

Benefits and drawbacks of SOC compared with TF

According to the evaluator interviewed, the key **benefit** of the PbR scheme Transforming Rehabilitation, compared with traditionally financed Probation Trusts, was that it **focused on the reoffending outcomes** rather than putting an output in place. The **reform helped to ‘shake up’ a system** that was stuck in its ways with regard to the services that were delivered, and during the bidding process a lot of new ideas emerged.

Conversely, the opinions of most stakeholders and the results of evaluations regarding the **potential of the scheme to innovate and enhance service quality** were negative overall.

It is clear from assessments by the UK National Audit Office⁶⁰⁹, HM Inspectorate of Probation⁶¹⁰ and other stakeholders including the House of Commons Justice Committee⁶¹¹, that in many areas the quality of the probation services delivered under the TR reform fell short of the MoJ’s expectations. Before the TR reform, the service provided by the Probation Trusts had been considered effective by both the HM Inspector for Probation and the HM Prison and Probation Service⁶¹². The Report of the Chief Inspector

⁶⁰⁶ MoJ. (2013). Transforming Rehabilitation A revolution in the way we manage Offenders.

⁶⁰⁷ NAO. (2014). Probation: landscape review.

⁶⁰⁸ HM Inspectorate of Probation. (2013). Annual Report 2012-13. Independent inspection of adult & youth offending work.

⁶⁰⁹ NAO. (2019). Transforming Rehabilitation: Progress review – Transcript; NAO. (2017). Investigation into changes to Community Rehabilitation Company contracts.

⁶¹⁰ HM Inspectorate of Probation. (2019). Report of the Chief Inspector of Probation.

⁶¹¹ House of Commons Committee of Public Accounts. (2019). Transforming rehabilitation: progress review. Ninety-Fourth Report of Session 2017–19.

⁶¹² NAO. (2014). Probation: landscape review, p.15-16.

of Probation showed that CRCs had failed to innovate – and that, overall, they had not fulfilled commitments to transform services⁶¹³. Those owners who were willing to remodel services found probation difficult to reconfigure or re-engineer. Moreover, the Chief Inspector of Probation concluded that the Through the Gate (TTG) services, a peculiar service introduced by the TR reform, had consistently failed to meet offenders' resettlement needs. Two thirds of offenders had not received enough help with their accommodation, employment or finances before being released.

With regard to **user satisfaction**, many service users did not know whether they were managed by the National Probation Service (NPS) or the CRC.⁶¹⁴ Service users were asked whether they felt that any changes had occurred since the Transforming Rehabilitation agenda had been implemented. More than three-quarters (77%) said that they had not noticed any change in the overall service that they personally received. The respondents reported a **huge negative impact from changing their probation officer**. Some relationships had suffered because of Transforming Rehabilitation, with some service users experiencing multiple changes of officers, or officers internalising the stress of the change. Where changes had been noticed in service delivery, these were positive in relation to the provision of education, changing offender behaviour work and substance misuse, although some negative changes were discussed in relation to housing and employment. In a survey of the voluntary sector stakeholders involved in Transforming Rehabilitation,⁶¹⁵ 60% of voluntary organisations said that TR had a negative or a very negative impact on their service users.

⁶¹³ HM Inspectorate of Probation. (2019). Report of the Chief Inspector of Probation.

⁶¹⁴ User Voice (2016). Transforming Rehabilitation: The operational model from the service user's perspective.

⁶¹⁵ TrackTR. (2018). Underrepresented, under pressure under resourced. The voluntary sector in Transforming Rehabilitation.

Additional **drawbacks** of the PbR programme include:

- Estimation of expected volumes proved to be very difficult, and differences forced suppliers to ‘squeeze’ their supply chains to some extent and not provide opportunities for smaller voluntary sector organisations to get involved, as the Ministry had planned.
- **CRCs faced severe financial pressures**, which made them less willing to invest in failing services.
- The **black box** approach did not allow the MoJ to hold CRCs accountable when service quality fell short of expectations. This left the Ministry few levers with which to intervene.
- The **deductions** included in the PbR scheme, and the related threat of having payments clawed back, impacted the financial planning of providers and may have prompted a risk-averse strategy, with under-investment throughout the supply chain and a lack of improved reoffending outcomes.

In general, the reform raised a **great deal of scepticism from the public and among institutions** in charge of assessing its work (i.e. NAO or HM Inspectorate). One criticism was that MoJ began the reform before having properly assessed the results of previous pilots such as the Peterborough SIB and others. TR was originally sold on the basis of improving resettlement for short-term prisoners, but the complete failure of CRCs to provide resettlement support to this group has been one of the most damning and persistent criticisms of the programme.⁶¹⁶

Finally, an argument could be made that the domain of probation is not best suited to PbR payment mechanisms, because outcomes are affected by many intervening factors. Many of the steps necessary to reduce reoffending were not fully within the control of the CRCs. These included the availability of wider public services, as well as support with housing, employment, mental health and substance misuse.

Scalability and replicability

The reform was applied across all of England and Wales.

The Justice Secretary acknowledged that the quality of probation services being delivered fell short of expectations, and announced that the Ministry would terminate its CRC contracts 14 months early in December 2020, at an additional cost to the taxpayer of at least at least GBP 171 million.⁶¹⁷

1.14.5. Efficiency

Intervention costs

As shown in the table below, Probation Trusts received greater funding overall, but the cost per participant was lower than that for CRCs. The NAO concluded that Transforming Rehabilitation achieved poor value for money for the taxpayer, and CRCs faced collective losses of GBP 294 million, compared with expected profits of GBP 269 million.⁶¹⁸ Although the NAO evaluated the Probation Trusts as being effective, it did not conduct a value-for-money assessment.

⁶¹⁶ <http://www.russellwebster.com/crcpbrjan18/>

⁶¹⁷ NAO. (2019). Transforming Rehabilitation: Progress Review.

⁶¹⁸ <https://publications.parliament.uk/pa/cm201719/cmselect/cmpublicacc/1747/174706.htm>, p. 10 and p. 36

Table 64. Cost comparison between the Transforming Rehabilitation PbR and the Probation Trusts (GBP)

	Transforming Rehabilitation 2015-2016	Probation Trusts 2012-2013
Participants	114,378 ⁶¹⁹	224,823 ⁶²⁰
Total cost	600,000,000 ⁶²¹	853,000,000 ⁶²²
Cost per participant	5,245	3,794

Source: elaborated by PPMI, based on NAO Reports in 2019 and 2014.

Note: participants reported for Probation Trusts reflect 2012 calendar year rather than the 2012-2013 financial year. The figures have not been adjusted for inflation.

Operational costs

Information is not available with regard to the overall operational costs and start-up costs of both the Transforming Rehabilitation PbR and the activities of the Probation Trusts. However, **the PbR's evaluation costs are likely to be similar to those of the TF intervention** because the CRCs have been inspected by HM Inspector for Probation in a similar way to that in which Probation Trusts had been before the reform. Please note, however, that evaluation costs here do not take into account the need to verify PbR-specific outcomes, as these would fall under the management costs.

For more context about the set-up of the intervention, please see the figure below. The National Probation Service came into existence on 1 June 2014. The Community Rehabilitation Companies were also set up at that time, as companies in public ownership. Staff who had previously been employed by the Probation Trusts were divided between the two new organisations, and all existing cases had to be divided as well. Shortly afterwards, the Community Rehabilitation Companies were transferred into private ownership, following a competitive bidding process. The successful bidders signed contracts for the CRCs in December 2014, and on 1 February 2015 began delivering probation services⁶²³.

⁶¹⁹ CRC caseload between October 2015 and September 2016 taken from the offender management statistics quarterly. Retrieved from <https://www.gov.uk/government/collections/offender-management-statistics-quarterly>.

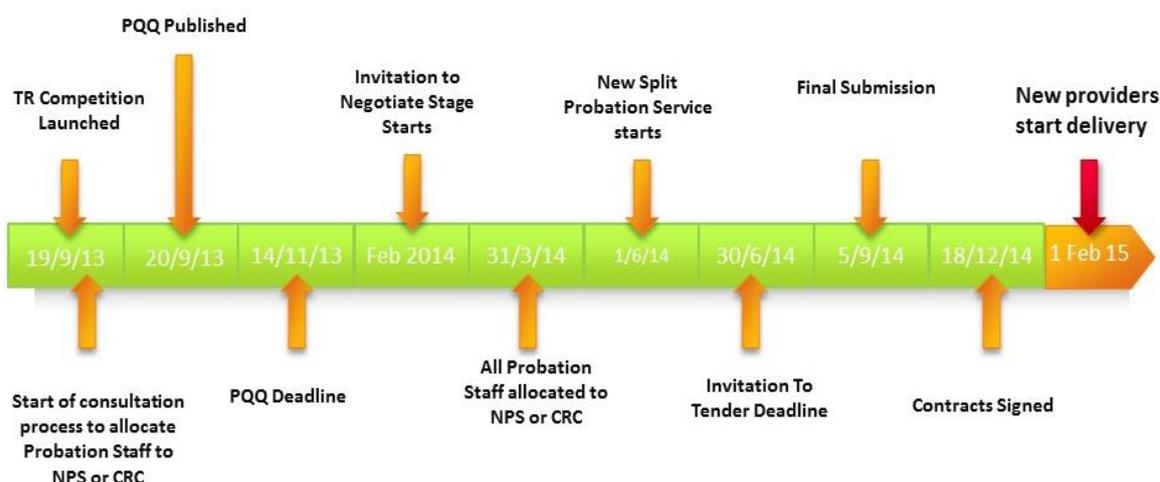
⁶²⁰ NAO 2014, 10.

⁶²¹ NAO (2019). Transforming Rehabilitation: Progress Review – Transcript, p.32.

⁶²² NAO 2014, 12.

⁶²³ <http://www.russellwebster.com/transforming-rehabilitation-timeline/>

Figure 12. Set-up process for Transforming Rehabilitation



Source: Russelwebster.com.

1.14.6. Design of the scheme

Design features

One effective design feature of the Transforming Rehabilitation programme was the ‘**bedding-in**’ period that took place before the start of the payment-related outcomes measurement. The first frequency payments were based on the 2015/16 annual cohort, which consisted of the October to December 2015 and the January to March 2016 cohort periods only (even though services began in February 2015). Annual binary top-up targets were not set for the 2015/16 annual cohort. This bedding-in period allowed the providers time to establish new processes and begin the transformation process. It also ensured that the impact of the new legislation with regard to offenders sentenced to less than 12 months could start to build up prior to the start of PbR measurement.⁶²⁴ Nevertheless, the NAO suggested that the bedding-in period should have been even longer: “the MoJ should consider running the new scheme **entirely on a fee for service basis for the first two years**, while the new system beds down, in order to establish a new baseline performance level on which to base the payment by results system.”⁶²⁵ The commissioner interviewed agreed with this recommendation, arguing that fee-for-service funding at the start of the intervention would have mitigated cash flow concerns.

Instead of funding the intervention entirely on a fee-for-service basis, the TR commissioner issued the **Foundation Payment**, which was meant to help providers ensure a minimum level of cash flow, and to help them cover on-going costs. However, the payment required disproportionate monitoring procedures to ensure value for money. The commissioner has therefore considered removing the Foundation Payment from the payment mechanism and explored alternative approaches to mitigating working capital constraints for providers.

The MoJ also considered ways to address **volume risk**, in order to safeguard itself against having to pay out excessive profits. For example, if participant volumes increased, providers would have been able to reach the **maximum payment cap** with a lower percentage change in reoffending rates.⁶²⁶ However, providers assumed all of the risk associated with

⁶²⁴ MoJ. (2014). Transforming Rehabilitation Programme Payment Mechanism.

⁶²⁵ NAO (2019) Transforming Rehabilitation: progress review, Session 2017–19, HC 1986, para. 1.4.

⁶²⁶ MoJ. (2014). Transforming Rehabilitation Programme Payment Mechanism.

falls in the volumes of offenders. This turned out to be the reality, which had a strong negative impact in terms of financing on the ability of CRCs to transform their businesses early in the programme. The reductions in volume varied greatly, from 6% to 36%.⁶²⁷ This was made worse by the fact that the **MoJ significantly overestimated the ability of CRCs to reduce their costs** to match any fall in income. It originally assumed that 20% of CRCs' costs were fixed, but in June 2017, the average fixed costs reported by CRCs amounted to 77%.⁶²⁸ This meant that CRCs were placed under significant financial strain when volumes of work were lower than expected, as they could not reduce their costs by as much as they needed to. The contracts also left the MoJ unable to hold CRCs accountable for poor services.⁶²⁹ Moreover, to overcome issues relating to the differences between predicted and actual volumes of caseloads, the commissioner suggested relying on estimates based on the previous year's actual figures, adjusting them year-on-year.⁶³⁰

Finally, various stakeholders recommended changing the way in which outcomes were measured and the associated payments, especially with respect to **narrowing the zone of statistical uncertainty** and **amending or removing the binary outcome hurdle**. For details, please see Section 1.14.3.

Impact on third-sector organisations and social enterprises

Overall, the **PbR approach** to commissioning offender rehabilitation services appears to have had a **negative impact on voluntary and third-sector organisations**. Her Majesty's Prison and Probation Service (HMPPS) acknowledged that the lower than anticipated volumes experienced by CRCs, coupled with their different mix of cases, had meant suppliers had 'squeezed' their supply chains to some extent and not provided opportunities for smaller voluntary-sector organisations to get involved, as the Ministry had planned. HMPPS said that it had learned from the rehabilitation reforms that the greater the risk it transfers in its contracts in terms of volumes and costs, the harder it is for such organisations to play a part. As of October 2018, just 159 (11%) of the 1,443 voluntary-sector organisations working in the criminal justice sector were providing services directly to CRCs. The NAO found that many voluntary and third-sector organisations were unable to provide the financial guarantees required to tender for a contract on their own.⁶³¹

1.15. Troubled Families – Phase II (UK)

1.15.1. Background

Troubled Families is a programme targeting families with multiple problems relating to crime, anti-social behaviour, truancy, unemployment, mental health problems and domestic abuse. Phase II of the programme aims to support families making "significant and sustained progress" against all the identified problems, or alternatively to move adults into situations of continuous employment.

The programme is managed by the Ministry of Housing, Communities and Local Government (MHCLG, formerly the Department for Communities and Local Government, DCLG), and delivered by upper-tier local authorities (LAs) in England and their partners. It involves a wide variety of interventions. Phase I of the programme was launched by the DCLG in 2012, and ended in 2015. The second phase of the programme, which

⁶²⁷ NAO (2016). Transforming Rehabilitation, p.9.

⁶²⁸ NAO 2019, 31.

⁶²⁹ NAO 2019.

⁶³⁰ MoJ. Transforming Rehabilitation Programme – Payment mechanism. Market Feedback and Development Considerations, p. 2.

⁶³¹ House of Commons Committee of Public Accounts. (2019). Transforming rehabilitation: progress review. Ninety-Fourth Report of Session 2017–19, p.14.

incorporates changes to both the financial scheme and the intervention logic, was launched in 2015 and is expected to run until 2021. The design of Phase II includes a more sophisticated mechanism to measure the achievement of outcomes and assess the effectiveness of the programme. Thus, only Phase II of the programme was selected to be part of this study.

One key **aim of the programme** is to tackle issues before they require costlier interventions. The programme aims to bring about widespread system change at local level, including the adoption of a 'whole family approach' and better multi-agency work. There is no centrally mandated way in which the programme should be implemented; instead, it emphasises flexibility at local level.⁶³²

Overall, the PbR model **intended to increase service responsiveness**. In 2011, Prime Minister David Cameron, during the launch of Phase I of the programme, cited in a public speech an estimate that, over the previous year, GBP 9 billion had been spent on the 120,000 families involved in the Phase I of the programme, due to the multiple interventions they received from different parts of the State. Of this GBP 9 billion, government figures estimated that GBP 8 billion of the money spent on these families had been spent "purely reacting to their problems".⁶³³ In line with this stance, the scheme operates according to the premise that public services have previously failed to target families who have multiple problems because those services operate in silos, and mostly in a reactive fashion.

At the core of the programme was the desire to achieve an overall shift in public expenditure from reactive service provision, based around responding to an accumulation of acute needs, towards earlier interventions targeting families in critical situations.⁶³⁴ This approach was intended to provide cost savings: the PbR is intended to allow reduced spending in the medium and long term on families affected by multiple problems.⁶³⁵

A second **reason** for choosing a PbR model relates to local authorities and their 'earned autonomy'. Aside from outcomes, local authorities were expected to assess their progress in joining up services and bringing about a 'whole family' approach to service delivery.

Specific actors help implement the programme. These include:

- Troubled Families coordinators from MHCLG, who oversee the way the programme is implemented by local authorities.
- Troubled Families key workers, who work alongside families to help them deal with their various needs, acting as a single point of contact. Key workers are given a high degree of flexibility in deciding how to help families.
- Troubled Families employment advisors. These are specialists from the Jobcentre Plus⁶³⁶, seconded by local authorities to provide employment advice to families.⁶³⁷

The PbR model involves both an **upfront attachment fee** to local authorities and a **reward payment** for each family that shows "significant and sustained improvement" across a set of criteria, or for adults moved into "continuous employment"⁶³⁸.

Local authorities receive an upfront attachment fee of GBP 1,000 for each family with whom they agree to work, and an GBP 800 results payment for each family with whom they

⁶³² Bate, A, Bellis, A. & Loft, P. (2020). The Troubled Families Programme (England), Briefing Paper, Number 07585, 22 January 2020, House of Commons Library, 3.

⁶³³ These figures were later criticised, with detractors arguing that the GBP 9 billion had been extrapolated from a smaller sample of families with at least five types of disadvantages and children with multiple behavioural problems, which may not be the case across the whole population of the Troubled Families programme (Bate et al. 2020, 13).

⁶³⁴ DCLG. (2019). National Evaluation of the Trouble Families Programme Final Synthesis Report, 7.

⁶³⁵ Bate et al. 2020, 13.

⁶³⁶ Jobcentre Plus is the part of the Department for Work and Pensions, which delivers working-age support services in the United Kingdom.

⁶³⁷ Bate et al. 2020, 22.

⁶³⁸ Bate et al. 2020, 3.

achieve an outcome (significant and sustained progress, or gaining continuous employment). In addition, each authority receives an annual **Service Transformation Grant** (most local authorities receive GBP 200,000 each year) to support local delivery of the programme. The opportunity to claim results is normally offered on a six-month basis.⁶³⁹

The programme has gradually been **moving towards its targets** in recent years: by 2019, 39% of treated families had made “sustained and significant progress”, and 5.3% of adults had moved into continuous employment, representing an overall success rate of 45%.⁶⁴⁰ The difficulties in obtaining continuous employment may be due to the fact that families in the programme face significant barriers to work such as childcare responsibilities for young children or mental health issues.⁶⁴¹ In January 2020, a further GBP 165 million in funding was assigned to the Troubled Families Programme, which was also **extended until 2021**.

Table 65. Summary of Troubled Families (Phase II)

Personally targeted social service	Social exclusion Specific problems
Target population	400,000 families To be eligible, each family must have at least two of the following six problems: 1. Parents or children involved in crime or anti-social behaviour. 2. Children who have not been attending school regularly. 3. Children of all ages who need help;; are identified as in need; or are subject to a Child Protection Plan. 4. Adults out of work or at risk of financial exclusion, or young people at risk of joblessness. 5. Families affected by domestic violence and abuse. 6. Parents or children with a range of health problems. (400,000 families equates to around 6.5% of all families in England, and the substantial discretion offered to local authorities in interpreting and applying the criteria means that almost any family who comes into contact with, or is referred to, a non-universal service could fall into the category). ⁶⁴²
Payment model	Binary
SOC scheme type	Mixed PbR
Cost	GBP 920 million (roughly GBP 200 million for the first year and GBP 700 million for the other four years) ⁶⁴³
Commissioner(s)	Ministry of Housing, Communities and Local Government (MHCLG), formerly the Department for Communities and Local Government (DCLG) ⁶⁴⁴
Social service provider(s)	Local Authorities throughout England In 33% of cases, local authorities contracted the service to the voluntary sector. ⁶⁴⁵
Investor(s)	N/A
Intermediary	N/A

⁶³⁹ DCLG. (2015). Financial Framework for the Expanded Troubled Families Programme, 38.

⁶⁴⁰ MHCLG (2019), Building Resilient Families: Third annual report of the Troubled Families Programme 2018-19, March 2019, 56,57

⁶⁴¹ MHCLG (2019), 57

⁶⁴² Bate et al. 2020, 5-6.

⁶⁴³ Bate et al. 2020, 3.

⁶⁴⁴ Bate et al. 2020, 3.

⁶⁴⁵ IPSOS Mori (2019) Troubled Families Program TFC Staff Survey Report Final, 7.

Evaluator(s)	<ul style="list-style-type: none"> - Technical Advisory Group: Dr Jo Blanden (University of Surrey); Mike Daly (Department for Work and Pensions); Dr Susan Purdon (Bryson and Purdon Social Research); and Professor Peter Urwin (University of Westminster). - Ipsos MORI⁶⁴⁶
---------------------	---

1.15.2. Equivalent TF scheme

In 2018, 14 local authorities moved to the **Earned Autonomy** model.⁶⁴⁷ The Earned Autonomy model involves more traditional upfront funding, meaning that some local authorities moved away from PbR. The Earned Autonomy model is implemented in some of the same areas of Troubled Families, with an identical target population, the same services, and the same main objectives as Troubled Families. We therefore compare the Troubled Families PbR scheme with the traditionally funded Earned Autonomy model.

The intention of Earned Autonomy was to “spur faster service transformation and drive high-quality support to families both during the lifetime of the programme and beyond”⁶⁴⁸. Until 2020, Earned Autonomy has been adopted by 14 local authorities. Interested LAs had to apply, and their applications were assessed by the MHCLG to be included in the new model⁶⁴⁹. LAs entering the Earned Autonomy model moved away from the full PbR arrangement that characterised Troubled Families, allowing them to access more upfront investment to embed better ways of working. The remaining local authorities remained within the Troubled Families PbR arrangement.⁶⁵⁰

Under the Earned Autonomy model, LAs continue to provide the same service but with lower risk involved, given that the **payment is provided totally upfront**, thus no longer depending on results.⁶⁵¹ The amounts of payments are based on the track records of the LAs involved. Every LA has to demonstrate how it will transform the service offering.⁶⁵²

Earned Autonomy has an especially positive impact with regard to the relationship between LAs and central government department. Troubled Families Coordinators in Earned Autonomy areas felt that the new funding model provided room to fund major initiatives quickly. One of the coordinators stated: “Earned autonomy is like a grown-up conversation between the Government and the local authority to say, okay, we know your programme is effective, you don’t have to jump through so many hoops to prove you have turned this family around, we kind of trust you on that”.⁶⁵³

Table 66. Comparability between Troubled Families PBR and the Earned Autonomy model

	Same as SOC	Comments
Personally targeted social service	Yes	Social exclusion Specific problems
Target population	Yes	Families with multiple problems

⁶⁴⁶ DLG (2019) National evaluation of the Troubled Families Programme 2015-2020: Findings Evaluation overview policy report, 25.

⁶⁴⁷ Bate et al. 2020, 23-24.

⁶⁴⁸ Ministry of Housing, Local Communities and Local Government (MHCLG) (2018). Supporting disadvantaged families. Annual report of the Troubled Families Programme 2017-18, 21.

⁶⁴⁹ Confirmed both during the interview with the commissioner, and interview with the provider.

⁶⁵⁰ Bate, A, Bellis, A. & Loft, P., The troubles Families Programme (England), Briefing Paper, Number 07585, 22 January 2020, House of Commons Library, pp. 19,20

⁶⁵¹ Interview with the provider

⁶⁵² Interview with the provider.

⁶⁵³ IPSOS Mori (2019a). Troubled Families Programme. Qualitative Case Study Report, Phase 2, Wave 2, 62.

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

	Same as SOC	Comments
Number of participants	Yes	The local areas involved in Earned Autonomy were previously involved in the Troubled Families scheme, and the same local target population was involved.
Location	Yes/No	14 local authority areas shifted from the Troubled Families programme to the Earned Autonomy model: <ul style="list-style-type: none"> - Barking and Dagenham - Brighton and Hove - Bristol - Cheshire West and Chester - Durham - Ealing - Islington - Kent - Leeds - Liverpool - Sheffield - Staffordshire - Westminster
Cost	Unknown	Detailed breakdowns of costs were not available
Commissioner(s)	Yes	MHCLG
Social service provider(s)	Yes	The same providers that previously delivered services locally under the Troubled Families scheme within the 14 local areas

1.15.3. Outcomes measurement

The Troubled Families PbR scheme involved a system of data reporting and monitoring at local level regarding the outcomes upon which payment was conditioned. Aside from this, the scheme was also subject to a qualitative process evaluation, an impact evaluation, and an economic evaluation, while surveys were conducted involving the key stakeholders in the programme. The impact and the economic evaluations were carried out by the Office of National Statistics together with the Ministry of Housing, Communities and Local Government (MHCLG). The surveys and the process evaluation were carried out by IPSOS Mori⁶⁵⁴.

Measurement of outcomes relating to payment mechanisms

The achievement of outcomes upon which the payment was conditioned were monitored through self-reported data provided from local authorities. This approach does not rely on an experimental or quasi-experimental design, nor on the presence of a control group.

Internal Programme Auditors and Local Authority Troubled Families Coordinators jointly agreed on the evidential expectations required to claim results (i.e. to confirm the achievement of the “significant and sustained improvement”) within their Troubled Families Outcomes Plan. This Plan details the targets for outcomes relating to payment indicators –

⁶⁵⁴ Bate et al. (2020).

which, as outlined below, are set locally.⁶⁵⁵ Following the definition of local outcomes, appropriate data are collected by the local authorities.

The outcomes assessed fall into the following main domains, which are consistent with the six programme headlines⁶⁵⁶:

- crime and anti-social behaviour;
- poor health;
- domestic violence and abuse;
- children who need help;
- poor school attendance;
- unemployment.

During the phase I of the programme, LAs could claim payment for a family once the family had been (broadly speaking) “turned around”. After a general review of the programme during Phase II, the payment to local authorities was conditioned on achieving two specific outcomes. Local authorities can claim funding for a family when they have:

- achieved *significant and sustained progress* against all problems identified at the point of engagement and during the intervention; or
- an adult in the family has *stopped receiving benefits and is in continuous employment*.
- “Continuous employment” is defined according to two criteria:
- Movement off work benefits, or satisfying an earnings threshold, if in receipt of Universal Credit (Troubled Families Employment Advisors provide information about earnings thresholds to local authorities).
- The sustainment of continuous employment for a period of time (the length of time an adult must remain in work depends on the type of benefit they were receiving previously)⁶⁵⁷.

Local authorities also have the possibility of adding local indicators.

⁶⁵⁵ Department for Communities and Local Government (DCLG) (2015). Financial Framework for the Expanded Troubled Families Programme.

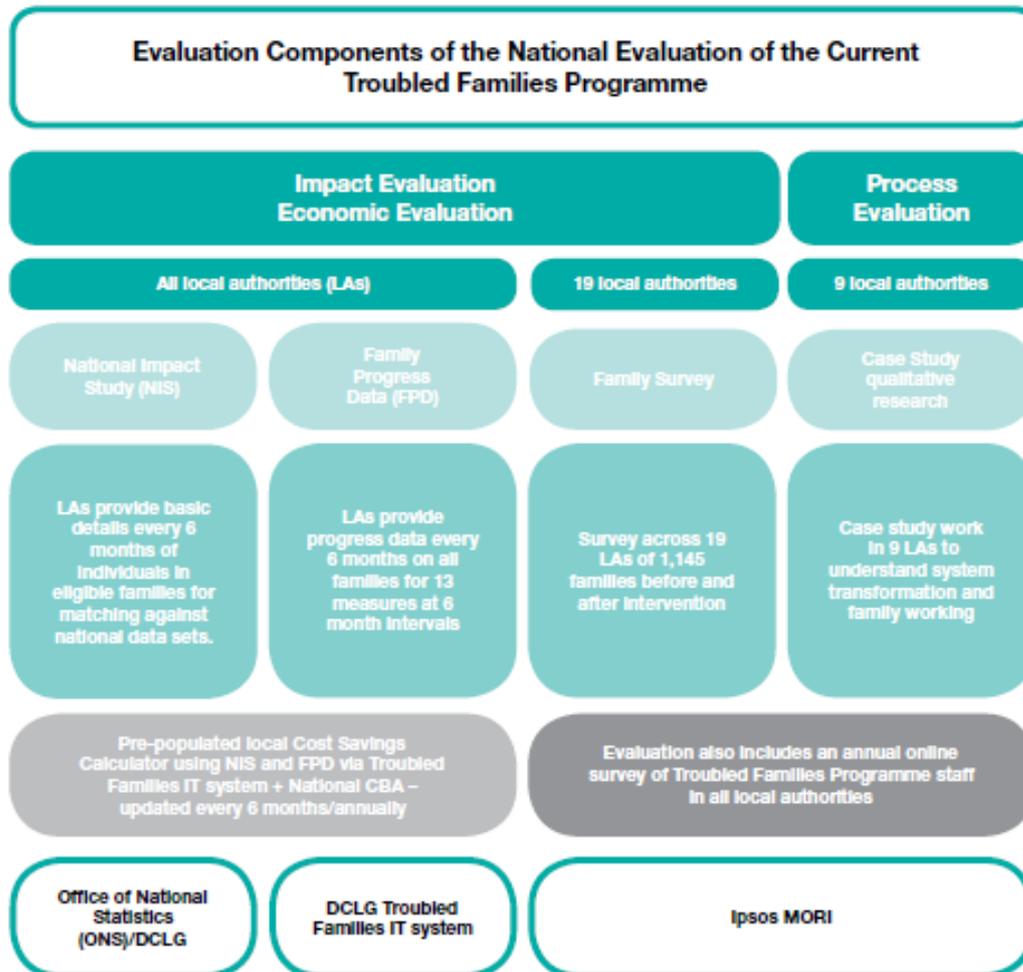
⁶⁵⁶ Bate et al. (2020).

⁶⁵⁷ Bate et al. (2020).

Other evaluations

The overall evaluation design of Phase II is summarised in the following figure.

Figure 13. Evaluation design for Troubled Families



Source: MHCLG (2019) National evaluation of the Troubled Families Programme 2015-2020: Findings Evaluation overview policy report, 25.

Impact evaluation – propensity score matching

The main evaluation of the impact of the programme, the National Impact Study, was carried out through a quasi-experimental design using the propensity score matching approach⁶⁵⁸. The use of propensity score matching provided a good level of internal validity for inferences, reducing selection bias. Nonetheless, the internal validity of a propensity score matching design appears much weaker than that of a randomised controlled trial, or other types of quasi-experimental designs such as difference-in-difference.

Data for the evaluation were provided directly to the MHCLG by the local authorities.

Data were collected by LAs, including basic details about individuals in eligible families to allow proper matching in the propensity matching design. The details provided by the local authorities about the families enrolled in the programme were matched with data held in administrative datasets. These datasets include the Police National Computer (held by the

⁶⁵⁸ Bate et al. (2020).

Ministry of Justice); the National Pupil Database (held by the Department for Education); as well as the Work and Pensions Longitudinal Study and the Single Housing Benefit Extract (both held by the Department for Work and Pensions).

Data were also collected for a comparison group, made up of families who were not enrolled in the programme.

The impact evaluation also relies on so-called Family Progress Data, which involves the collection of data that are not held in national administrative datasets. These are provided by LAs, about those families eligible for or enrolled in the Troubled Families programme.

LAs were asked to submit National Impact Study data and Family Progress data to the Office for National Statistics every six months. The identities of families in the National Impact Study data were then matched to individuals in administrative datasets to create a reliable control group.

Impact evaluation – Family and staff surveys

As part of the impact evaluation, three main surveys were carried out to register the perception of key stakeholders in the programme: families, staff, and employment advisors⁶⁵⁹.

More specifically, the Family Survey is a quantitative longitudinal survey of families in receipt of help from the programme in 19 local authorities. The Staff Survey is an online survey involving Troubled Family co-ordinators, key workers or local practitioners. A third survey is addressed to Troubled Family Employment Advisors.

The Staff and Employment Advisor surveys were run through just one round of data collection; analysis of the data from the Family Survey was based on a comparison of family outcomes at the beginning of the treatment and two years later. This allowed significant differences to be identified using statistical testing such as t-tests or McNemar's tests. The absence of a pre-post comparison in the Staff and Employment Advisers' surveys, coupled with the lack of significance testing, weakens the validity of the evaluation design.

The survey data were complemented by interviews and investigated outcomes related to the following domains:

- employment;
- education;
- health and wellbeing;
- crime and anti-social behaviour;
- domestic violence and abuse;
- service experience and satisfaction;
- condition of young people.

The main indicators exploited in the analysis for the outcomes are presented in the table below⁶⁶⁰.

⁶⁵⁹ Ibid.

⁶⁶⁰ IPSOS Mori (2019b) - Troubled Families Programme, Family Survey.

Table 67. Troubled Families outcome indicators

Outcome domain	Indicators
Employment	<p>Current status of main carer</p> <p>Current working status of main carer</p> <p>Workless household: no-one in household currently working</p> <p>Main carer's job applications and steps towards finding a job</p> <p>Active steps towards finding work in the last four weeks</p> <p>Perception of employment perspectives in the next year</p> <p>Proportion of households receiving fiscal benefits and tax credits</p> <p>Level of household income</p> <p>Household's ability to keep up with bills and regular debt repayments over the last two years</p> <p>Proportion of families falling behind in rent payments</p> <p>Perception of family financial management</p> <p>Proportion of families having at least one type of credit or product loan.</p>
Education	<p>Highest level of qualification in the household</p> <p>Children's behaviour at school (main carers called into school to talk about children behaviour)</p> <p>Concerns about children's attendance at school or college</p> <p>Proportion of children in need/children in a protection plan/looked after LA</p> <p>Proportion of children with special educational needs</p> <p>Need for more support with children with special education needs.</p>
Health and wellbeing	<p>Main carer's visits to a general practitioner or a family doctor in the last six months</p> <p>Number of A&E visits in the last six months</p> <p>Health status perception</p> <p>Proportion of families with at least one person with a long-standing health condition</p> <p>Life satisfaction perception</p> <p>Perception of mental health</p> <p>Probability of feeling under strain, losing sleep over anxiety and losing confidence in themselves</p> <p>Overall level of wellbeing</p> <p>Proportion of respondents feeling relaxed/optimistic/dealing well with problems</p> <p>Presence of barriers to healthy eating</p> <p>Scores on the Alcohol Use Disorders identification test</p> <p>Frequency of alcohol drinking</p> <p>Intensity of alcohol drinking (number of alcoholic drinks on a typical drinking day)</p> <p>Episodes of excessive drinking (Female >6 drinks, males >8 drinks) in last six months</p> <p>Proportion of carers smoking cigarettes</p> <p>Proportion of carers taking prescription drugs</p>
Crime and anti-social behaviour	<p>Contact(s) with the police in the last six months</p>

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE
PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

Outcome domain	Indicators
	<p>Proportion of carers or members of the family cautioned or convicted for committing a crime</p> <p>Presence of police action as a result of an involvement in crime</p> <p>Episodes of using force or violence at home in the last six months</p> <p>Actions taken to stop anti-social behaviour (evictions, warning letters, possession orders...)</p>
Domestic abuse and violence	<p>Happiness perception in family relationships</p> <p>Proportion of carers regretting marrying their partner</p> <p>Presence of non-sexual abuse by a partner</p> <p>Presence of carers experiencing abuse from their partner in the last six months</p>
Service experience and satisfaction	<p>Proportion of families remembering keyworker's identity</p> <p>Frequency of contacts with keyworkers</p> <p>Identity of the family member interacting with the keyworker</p> <p>Perception of helpfulness and clarity of the keyworker</p> <p>Perception of the relationship with the keyworker</p> <p>Attitudes towards help received from keyworkers</p> <p>Attitudes towards the future</p>
Condition of young people	<p>Episodes of "being in trouble" at school in the last six months</p> <p>Proportion of young people reporting unauthorised absence from school</p> <p>Work status of the children</p> <p>Steps towards finding work</p> <p>Work readiness</p> <p>Work aspirations for the next year</p> <p>General perception of health status</p> <p>Perception of health and mental wellbeing</p> <p>Perception of likelihood of being bullied</p> <p>Proportion of children having an alcoholic drink</p> <p>Frequency of alcohol consumption</p> <p>Episodes of heavy drinking in the last four weeks</p> <p>Proportion of children smoking cigarettes</p> <p>Proportion of children who report trying at least one street drug</p> <p>Perception of feeling safe at home</p> <p>Contact with the police in the last six months</p> <p>Police action against children in the last six months</p> <p>Reported episodes of involvement in a crime in the last month</p> <p>Proportion of children carrying a weapon</p> <p>Proportion of children receiving actions used to discourage anti-social behaviour in the last six months</p>

Economic evaluation – propensity score matching

An economic evaluation of the Trouble Families programme is on-going and uses a quasi-experimental design that relies on the propensity score matching approach. The economic evaluation involves the monetisation of key indicators relating to programme outcomes in order to build a cost-benefit analytical design.

In the economic evaluation, LAs collect data about the programme's set-up and management costs. They also collect data about families' use of reactive and universal services, such as visits to Accident & Emergency (A&E), incidents of anti-social behaviour, and out-of-work benefits. Using a Cost-Saving Calculator provided by the Ministry of Housing, Communities and Local Government (MHCLG), authorities were able to monetise such data in terms of savings.

The cost savings calculator assesses the fiscal costs of families' service use by attaching a unit cost to each reported incident. Unit costs were derived from New Economy Manchester's Unit Cost Database, which provides estimates of the costs of all incidents of interest. The Database enabled a proper comparison for the majority of outcomes.

The evaluation assessed costs relating to outcomes associated with the six programme headlines mentioned above.

Process evaluation

A process evaluation was carried out using a qualitative approach to assess the 'service transformation' objective of the programme.

The evaluation followed a case-study methodology. Cases were investigated through in-depth interviews with the staff who delivered the programme, as well as with the families receiving services. Data from the case studies were analysed using a 'thematic analysis' approach based on existing and pre-defined Thematic Frameworks, which were created following early (pilot) interviews.

Data management was conducted using the Framework approach within the software NVivo, which supports rigorous and comprehensive within-case and thematic analysis. This approach contributed to making the coding processes more systematic and less subjective.

Barriers to and enablers of the measurement process

After harsh criticism concerning the Phase I evaluation, the government's political commitment to improving the robustness of evaluation, moving from Phase I to Phase II, was a key enabling factor, fostering the **variety of measures, data sources and methodologies** exploited to assess Phase II of the programme.

Barriers to the measurement process relate to the strong **heterogeneity between sites** and the frequency with which data was missing, particular in relation to the economic evaluation.

In addition, difficulties in the creation of reliable control groups weaken the robustness of the evaluation. With regard to the impact evaluation, its design is weakened by the fact that some families in the comparison group may have been receiving services related to the TF Programme. Aside from this, the heterogeneity between different treatments and between different local authorities complicates the comparison and the matching.

Table 68. Summary of the measurement process in Troubled Families (Phase II)

Measurement methodology	
Methods relating to payment mechanisms	Data self-reported by local authorities
Other evaluations	Impact evaluation – surveys of families, staff, and employment advisers Impact evaluation – propensity score matching Economic evaluation – propensity score matching Process evaluation (service transformation) – case study
Methods using experimental or quasi-experimental design	Yes
Control group	Yes (in the propensity score matching)
Causality of impact	Causality of impact was assessed in the impact evaluation using a quasi-experimental design.
Enabling factors and strengths of the evaluation process	<ul style="list-style-type: none"> - Political commitment for the evaluation - Wide variety of measures and data sources - Good degree of methodological robustness Quasi-experimental character of impact evaluation through the use of propensity score matching
Barriers to and weaknesses of the evaluation process	<ul style="list-style-type: none"> - Data management and availability (frequency of missing data, particularly for the economic evaluation) - Heterogeneous data from different sites Difficulties in creating reliable control groups for health and school attendance outcome
Governance of measurement	
Independence of evaluator	Yes
Evaluator	Local authorities for outcomes relating to payment Technical Advisory Group: <ul style="list-style-type: none"> - Dr Jo Blanden (University of Surrey); - Mike Daly (Department for Work and Pensions); - Dr Susan Purdon (Bryson and Purdon Social Research); - Professor Peter Urwin (University of Westminster) Ipsos MORI Office of National Statistics MHCLG

1.15.4. Effectiveness

Achievement of outcomes

As mentioned above, by 2019, a total of 171,890 families achieved the agreed outcomes, accounting for a success rate of 45%.⁶⁶¹ At the present time, no impact evaluation has been carried out to compare outcomes in the Earned Autonomy areas with those of the other areas, taking into account all the intervening factors. Although the evaluation by MHCLG (2020) contains data about the achievement of the outcomes, it must be noted that the bare numbers do not provide sufficient grounds to compare the effectiveness of the two models. The comparison does not take into account all intervening factors relating to the different LAs in which services were provided. For instance, as mentioned above, the switch from the PbR to the EA model was decided after a bidding process in which LAs were evaluated based on their track record during the PbR period.

Although we do not have all details about this evaluation, our interviews highlighted how the switch to the EA model was a sort of ‘upgrade’ for local authorities that were able to track outcomes more efficiently. Partly because the EA model was chosen for better-performing LAs, those LAs operating under the EA model continue to outperform those in which contracts are issued on a PbR basis, as illustrated in the table below.

Table 69. Troubled Families Outcomes, by funding model

	Families on the programme, March 2020	Families achieving successful outcomes up to 5 April 2020	The share of successful families
Earned Autonomy LAs	57,750	53,331	92%
PbR LAs	339,294	293,858	86%

Source: elaborated by Polimi, based on MHCLG (2020).

In terms of the impact of the two models, participants in the case studies mentioned that it was difficult to say whether Earned Autonomy had changed the demand for services as yet, particularly due to the number of other initiatives that were running alongside it, which could also be responsible for this impact.

User perceptions emerging from the analysis of case studies were also positive. Families’ views about their relationship with their key worker were largely positive: most remember their key worker by name and say that contact was frequent; most of the main carers recall agreeing a plan with their key worker and are **positive about their level of involvement** in this process; main carers felt that key workers were clear about the changes they needed to make as a family and **found them helpful and supportive; attitudes towards the help received were positive**, with most feeling that their key worker made a difference, particularly with regard to parenting⁶⁶².

⁶⁶¹ DCLG (2019), Building Resilient Families: Third annual report of the Troubled Families Programme 2018-19, March 2019, p.56,57

⁶⁶² IPSOS Mori (2019a).

Benefits and drawbacks of SOC compared with TF

Earned Autonomy areas were included alongside some PbR areas into the process evaluation. Thus, results from the case studies provide some qualitative insights about the added value of the two models, as well as their main benefits and drawbacks.

The evaluations recognised that the PbR programme had contributed to **innovating the delivery of services**. Although all services had been provided before the introduction of the Troubled Families scheme, they were reconfigured as a result of it, particularly the 'front door' component, meaning that keyworkers actually work alongside families to solve their problems.⁶⁶³ The main innovation coming with the SOC scheme was to **unify all services into a broad scheme with a whole-family approach**.⁶⁶⁴

Providers also noted that the service provided under the PbR brought added value to the provision of services to families, in particular through the development of partnerships, the overall cultural transformation of the services, and the development of early help services that tackled urgent social needs immediately.⁶⁶⁵ The subsequent introduction of the Earned Autonomy model was also positively perceived according to the case studies. Local authorities' perceptions of EA were positive, because becoming an Earned Autonomy area was widely seen as an endorsement of a local programme, and an opportunity not only to build on existing initiatives but to speed up progress on them.⁶⁶⁶

Most stakeholders agreed that the PbR programme allowed a degree of autonomy to local authorities. One local commissioner interviewed stated that autonomy in setting and assessing their outcomes, as well as flexibility in setting priorities and programming, allowed LAs to change interventions on the basis of the needs and progress of families.

In addition, the infrastructure for the measurement of outcomes allowed (according to one provider interviewed) data to be collected and share between different agencies working on social services, thus enabling the overall **development of data-based public services and improvements to the data management culture**.

One of the main drawbacks of the programme, mentioned by stakeholders and identified in the evaluations, relates to **uncertainty over budgets in the PbR model**. In this sense, stakeholders emphasised that the shift to the Earned Autonomy model allowed them to overcome this uncertainty in relation to budgeting. Overall, the results of the case study evaluation are **fairly positive about the Earned Autonomy model**. First, upfront funding provided local authorities with greater certainty compared with the PbR model, thus enabling their plans to materialise more quickly. Troubled Families Coordinators, for instance, noted that under the PbR scheme their ability to quickly fund major initiatives was fairly limited, and earning autonomy in this area was seen as an endorsement of the local programme and an incentive to support a major service transformation. One coordinator stated: **"Earned Autonomy is like a grown-up conversation between the Government and the local authority** to say: 'okay we know your programme is effective, you don't have to jump through so many hoops to prove you have turned this family around, we kind of trust you on that."⁶⁶⁷

Scalability and replicability

The PbR scheme has not been scaled or replicated. However, it is worth noting that the number of beneficiaries and LAs involved in the programme increased between Phases I

⁶⁶³ Interview with the provider

⁶⁶⁴ Interview with the provider

⁶⁶⁵ Interview with the provider

⁶⁶⁶ IPSOS Mori (2019a), 61.

⁶⁶⁷ IPSOS Mori (2019a), 61- 65

and II. In addition, , the government announced in January 2020 that the programme would continue for another year, until March 2021.⁶⁶⁸

1.15.5. Efficiency

Intervention costs

No detailed information is available about the funding allocated to Troubled Families in each LA, making it impossible to compare the costs in LAs that continued to pursue the PbR approach with those that transitioned into the Earned Autonomy model. According to the stakeholder interviewed, however, the allocated funding was very similar.

We also note that the Troubled Families PbR has been deemed cost-effective. The total public benefit for the 2017/18 cohort was estimated to be GBP 651 million, or an average GBP 5,245 per family. This suggests every GBP 1 spent on the programme delivered GBP2.28 of economic benefits. The total fiscal benefit for the 2017/18 cohort was estimated to be GBP 432 million, or an average of GBP 3,484 per family. This suggests that every GBP 1 spent on the programme delivered GBP 1.51 of fiscal benefits, although not all of these are cashable, particularly in the short term.⁶⁶⁹

Operational costs

Given that the components of the Troubled Families PbR and the Earned Autonomy model track the same outcomes using the same processes, we would expect their operational costs to be similar. Nevertheless, **the service providers interviewed noted that the reporting burden under the Earned Autonomy model was lower, due to payments being provided upfront and the payment scheme being less conditional.** Although set-up, implementation and evaluation costs incurred by the commissioner are not available, providers underlined the necessity for Local Authorities to hire data analysts and staff for reporting if they are to engage in a programme such as Troubled Families.

1.15.6. Design of the scheme

Design features

Perhaps the most efficient and effective design feature of the Troubled Families programme is its **identification and pro-active reaching out** to families facing multiple problems. As mentioned in the DLCC report, “the efficiency lies in the process through which families are reached in the programme, because identifying problems at the earliest point and supporting families with keyworkers helps prevent more serious costs later on and reduces the impact on other services, even if it appears to increase costs in the short term.”⁶⁷⁰ Similarly, stakeholders appreciated the programme’s focus on **the family as a whole**, which was reflected in the outcomes measured. However, the same approach, both in terms of pro-active reach and outcomes considered, is now also implemented through the Earned Autonomy model (TF funding) – although arguably, the PbR approach helped to bring this about by transferring some of the risk associated with the new intervention from the commissioner to the providers.

⁶⁶⁸ Bate et al. (2020), p. 42

⁶⁶⁹ MHCLG (2019), Building Resilient Families: Third annual report of the Troubled Families Programme 2018-19, March 2019, p.48-49.

⁶⁷⁰ MHCLG (2019) Building Resilient Families: Third annual report of the Troubled Families Programme 2018-19, March 2019, p. 47

While interviewees noted that data collection and management costs, particularly those regarding health data, were time-consuming and economically costly, they praised the annual **Service Transformation Grant** – up to GBP 200,000 in funding that local authorities receive in each year to help develop appropriate data systems and support local delivery of the programme. With the new funding provided for the extension of the programme, it was decided to double the size of the Service Transformation Grant to local authorities, in order to increase local investment in transformations that improve the reach and capacity of family support.⁶⁷¹

Nevertheless, providers stressed that greater flexibility for local authorities in setting priorities and programming is a necessary condition in relation to the use of PbR schemes, allowing LAs to change programme priorities on the basis of the progress of families taking part in the programme.

Impact on third sector organisations and social enterprises

The service provider interviewed noted that uncertainty stemming from PbR funding caused problems with hiring staff. The provider mentioned that such uncertainty is solved through the Earned Autonomy Model.⁶⁷²

⁶⁷¹ MHCLG (2020). Improving families' lives: Annual report of the Troubled Families Programme 2019-2020, p. 9.

⁶⁷² Interview with the provider

Annex 2: Methodology

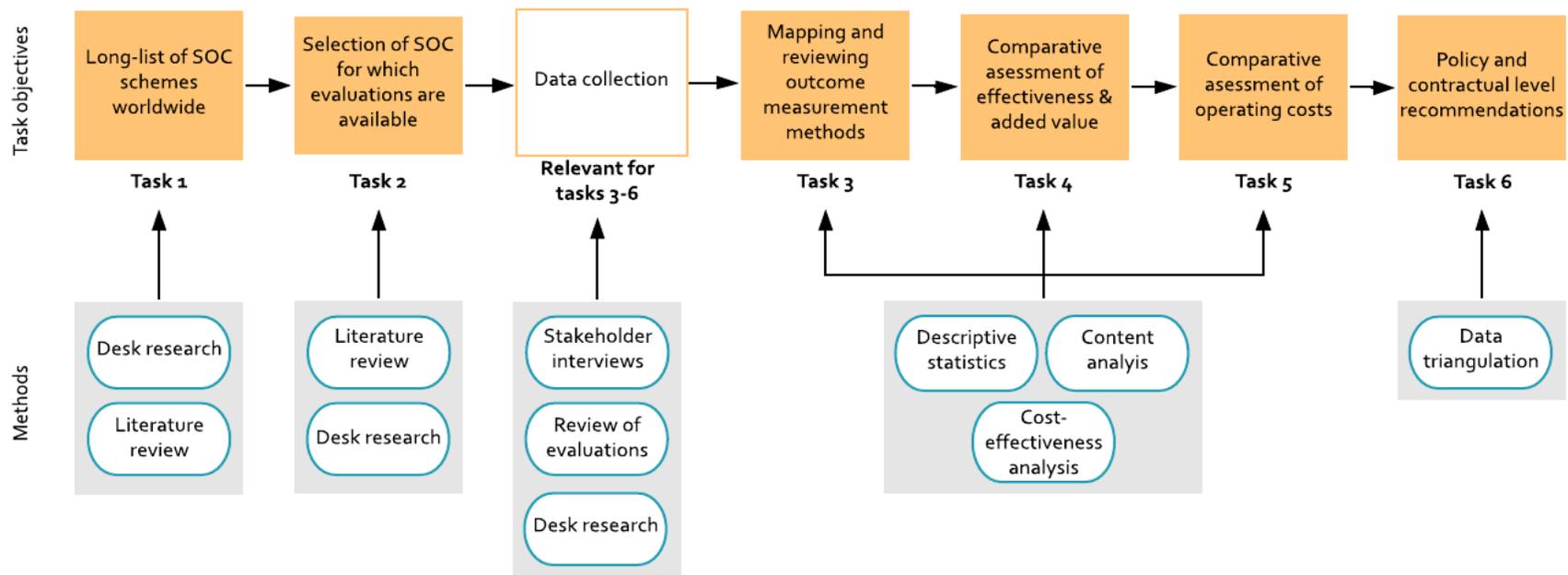
Annex 2 details the methodology applied during the study, the reasons for the methods chosen and their limitations. The study comprises six consecutively linked tasks, requiring a mixed methods approach. This relied on extensive desk research and a literature review; interviews with relevant stakeholders; content analysis; and quantitative methods, such as descriptive statistics and cost-effectiveness analysis.

In Task 1, we selected a long list of diverse SOC schemes and models, to which we applied the selection criteria had developed. The results of Task 1 are presented in Annex 3. In Task 2, we worked with the list of diverse SOC operations we had compiled in Task 1 to assess for which schemes we could find evaluations.

Subsequent tasks were based on the sample of SOC operations selected in Task 2. In Task 3, we mapped and reviewed the outcomes measurement methods used in the 15 SOC selected cases. In Task 4, we implemented comparative assessments of the effectiveness and added value of interventions delivered under SOC and non-SOC programmes. Task 5 assessed intervention and operating costs (administrative, human resource, financial) of both SOC and non-SOC interventions. Finally, in Task 6, we triangulated the different sources of information collected to prepare policy and contractual-level recommendations.

While Task 1 and Task 2 were mostly based on the initial desk research, to gather information for Task 3, Task 4, Task 5 and Task 6 we conducted a cross-country data collection programme. This included specific desk research in national languages and interviews with different types of stakeholders involved in each of the SOC schemes analysed. The data collection was guided by standardised templates developed to collect comparable and comprehensive information (see Annexes 4 and 5). This information then fed into the mapping, comparative analyses and development of recommendations. The figure below demonstrates how all tasks and methods relate to one another and result in the project deliverable. In the subsequent sections, we detail the methods used in each task, and conclude by outlining the limitations of the study.

Figure 14. Study design



Source: PPMI.

1.1. Task 1: Long-list of completed and diverse SOC's

Task 1 focused on the identification of completed SOC operations. To developing a diverse list of completed SOC schemes, we pursued a three-fold search strategy: a review of the main repositories and papers on SOC schemes; additional Internet searches; and leads provided by our country experts and broader networks.

The review of repositories, prominent papers and other sources, mapping or analysing SOC schemes, focused primarily on:

- Social Finance (due to its extensive SIB database⁶⁷³, which equipped us with a good first insight into the scope and scale of interventions around the globe).
- The OECD, whose Social Impact Investment Initiative aims to provide an overarching insight into four areas of social impact acceleration: financing, innovation, policy, and data⁶⁷⁴. This resource is also filled with multiple examples from around the world, covering a wide range of SOC interventions.
- Instiglio. Like Social Finance, this is a good source of information on results-based financing projects and impact bonds worldwide, including the SIB & Development Impact Bond map and Instiglio projects⁶⁷⁵.
- Pay for success US⁶⁷⁶, which provides a map of projects implemented in the US.
- The Brookings Impact Bonds project,⁴⁰ which has analysed a large number of SIBs to identify their potential to address a wide range of intractable social challenges in high-, middle- and low-income countries.

We employed these online sources as our first 'go-to' sites due their already quite exhaustive lists of social impact interventions, as well as other information on social impact investing, social service innovation and similar topics of interest.

Additional internet searches aimed to ensure an all-encompassing identification of all SOC interventions. We employed a **focused and systematic search strategy** using internet search engines (primarily Google). First, to ensure an exhaustive search, we defined a set of keywords covering the variety of different terms denoting SOC schemes (see examples in the box below). Second, we used the Boolean search operators AND, OR, NOT to construct focused search strings using our set of keywords alongside specific country names (to check each country and ensure a global reach). This was also carried out in national languages if English searches did not return a sufficient number of results.

Search keywords used

- **English:** "social outcomes contracting", "social impact bonds", "results-based financing", "pay-for-success", "pay-for-performance", "payment-by-results", "development impact bonds", "SIBs", "DIBs", "PBR", "RBF", "PFS", "PFP", "SOC", "social impact investing", "social impact investment", "social service innovation", "social impact contract", "social finance", "social return on investment AND United Kingdom", "social impact bond" AND "evaluation" AND "Portugal" (and other country-specific searches). "outcomes funds"; "outcomes funds" AND "EU states" (and specific country names: Germany, France, Italy, Netherlands, Portugal, Austria, Switzerland); "results fund" AND "evaluation", "social outcomes" AND "EU"

673 Social Finance. (n.d.). Retrieved from <https://sibdatabase.socialfinance.org.uk/>

674 OECD Social Impact Investment Initiative. (n.d.) Retrieved from <https://www.oecd.org/dac/financing-sustainabledevelopment/development-finance-topics/social-impact-investment-initiative.htm>

675 Instiglio. (n.d.). Retrieved from <https://www.instiglio.org/en/projects/> & <https://www.instiglio.org/en/sibs-worldwide/>

676 Pay for Success. (2019). Projects. Retrieved from: <https://www.payforsuccess.org/projects>

⁴⁰ Available at: <https://www.brookings.edu/series/impact-bonds/>

(and specific country names), “social outcomes funds” AND “EU” (and specific country names), “Payment by results” AND “Germany” (and other country names), “social impact fund” AND “Germany” (and other country names). “Output based commissioning”/ “output-based procurement”/ “outcome commissioning” AND “Germany” (and other country names). “Pay for success”/ “pay for performance”/ “results based financing”/ “performance based financing” AND “Germany” (and other country names).

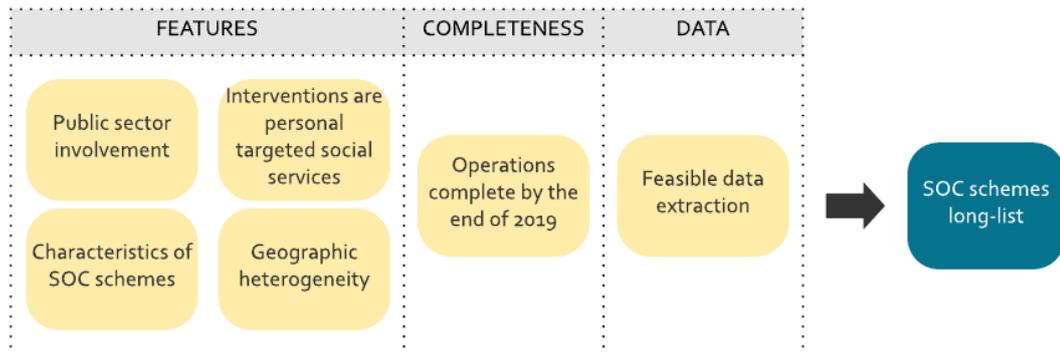
- **French:** “contrat de résultats sociaux”, “projets d’investissement à impact social” AND “évaluation”, “investissement à impact social”, “projet de résultats sociaux”, “paiement par évaluation des résultats”, “Contrat à Impact Social”, “CIS”, “fonds de résultat” AND “évaluation”, “Fonds de Paiement au Résultat”, “Le FPR” AND “financer”, “contrats à impact social”.
- **German:** “Sozialer Ergebnis Fonds” (social outcome fund), “Fonds” AND “soziale Auswirkungen”, “Investitionen in soziale Auswirkungen“, “soziale Wirkungspartnerschaften“ (SWK).
- **Italian:** “valutazione dei fondi per i risultati sociali”, “social impact investments” and “italy”, “fondi di investimento ad impatto sociale”, “finanziamento a impatto sociale”.
- **Portuguese:** “projetos de impacto social”, “fundos de resultados sociais”, “fundo para a inovação social” (aka fis), “títulos de impacto social”, “fundo de impacto social”.

When reviewing each search result, we applied a set of **inclusion and exclusion criteria** to determine if SOC schemes identified should be included into our list of SOC operations. These criteria served as a tool to select diverse SOC operations and practices. Our selection criteria included the following:

- The selected SOC operations should have been **completed**, with **data available and feasible for extraction** (i.e. reports are available online, potential interviewees can be identified).
- SOC operations should include the main characteristics of SIB or PbR schemes as described in Chapter 1 of the main report. We limited the selection to schemes in **which public sector has been involved** and left purely private initiatives outside the scope of this exercise.
- We reviewed SOC schemes and models implemented **worldwide**.
- We aimed to cover, to the extent possible, all of the **various design models of existing SOC schemes** (e.g. binary, frequency or hybrid schemes; direct, intermediated or managed SIBs; individual SIBs and SIB funds).
- SOC operations should cover **issue areas** that are relevant for social services. We included only those schemes that covered the following personally targeted social services: **caring obligation, disabilities, labour market-related, crisis and emergency, social exclusion, and specific problems**.

The figure below summarises the selection criteria used during the identification, selection and compilation of SOC operations.

Figure 15. Identification/selection criteria



After applying our selection criteria, we also checked to make sure that no additional SOC schemes had been completed by carrying out the following:

- Checking relevant social innovation and outcomes funds. The latter denote a programmatic (long-term) approach to outcomes-based commissioning, since they allow several SOC schemes to be funded in parallel.
- Checking on leads through our international networks, including the European Venture Philanthropy Association (EVPA) and the Global Steering Group for Impact Investment (GSG), to identify any additional schemes and fill in gaps, especially in countries for which very little information was available in English.

As a result of this exercise, we finalised a list of **64** SOC operations (see Annex 3). The table below provides a brief description of each SOC scheme in this sample, including its issue area, country and SOC type. For simplicity, schemes that span more than one social service area were classified in the table as ‘integrated’, as they resemble a ‘one-stop shop’ type of integrated services delivery. More than half of identified SOC schemes focused on labour market-related social services (52%). The majority of completed SOC schemes were in Europe (78%) with SIBs being the most popular SOC type of scheme (75%), while a quarter of schemes used PbR models.

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

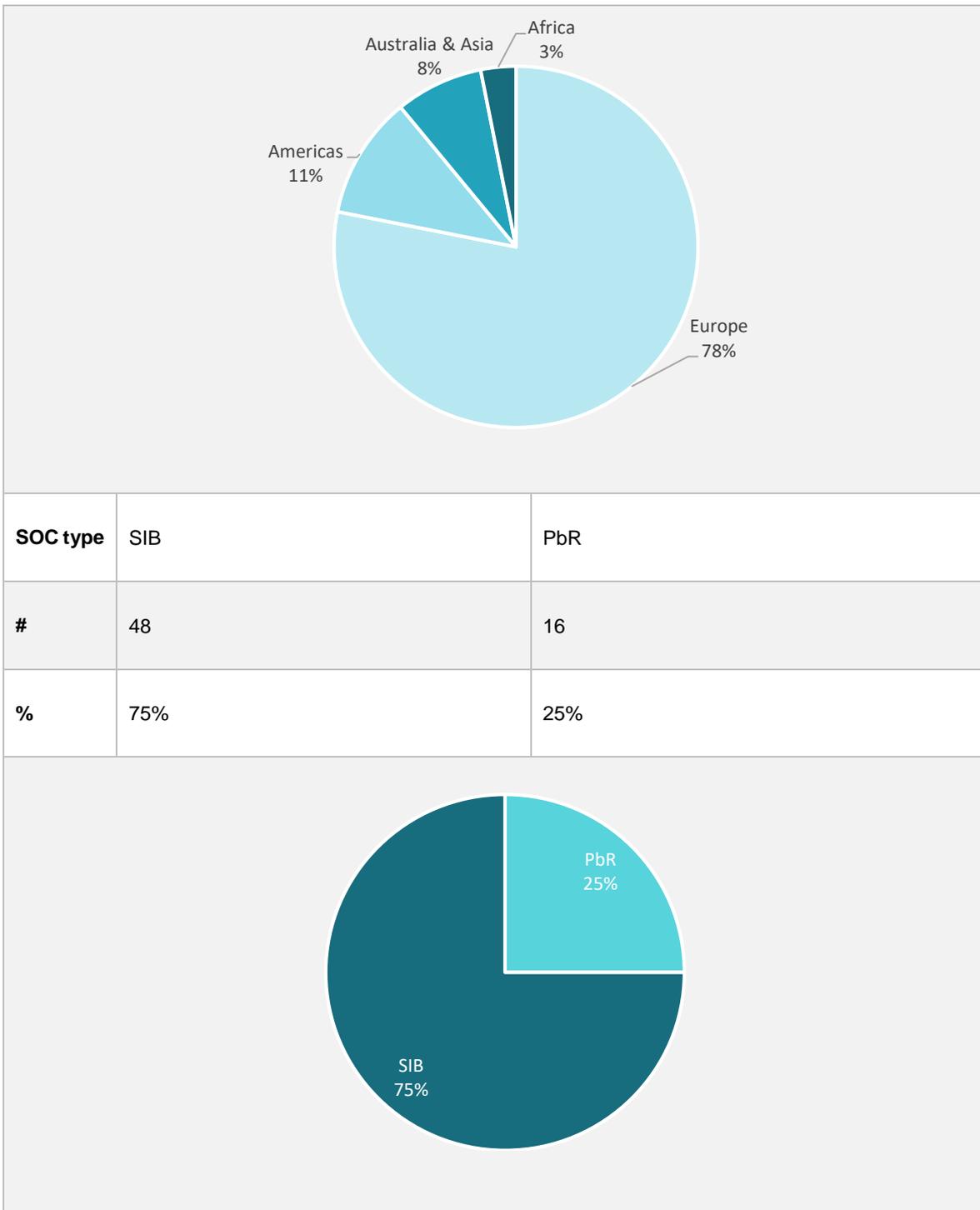
Table 70. Descriptive statistics of SOC cases identified for further analysis

Social service	Caring obligation	Disabilities	Labour market-related	Integrated	Crisis and emergency	Social exclusion	Specific problems
#	4	1	33	10	6	9	1
%	6%	2%	52%	16%	9%	14%	3%

A pie chart illustrating the distribution of 50 SOC cases across seven categories. The largest segment is 'Labour market related' at 51%, followed by 'Integrated' at 16%, 'Social exclusion' at 14%, 'Crisis and emergency' at 9%, 'Caring obligation' at 6%, and 'Disabilities' at 2%. A very small segment for 'Specific problems' is also indicated at 2%.

Country/region	Austria	Belgium	Finland	Germany	Netherlands	& Germany Netherlands	Ireland	UK	USA	Canada	Peru	Colombia	Mexico	Japan	South Korea	Australia	Morocco	Burkina Faso
#	1	1	1	1	6	1	1	38	2	1	1	2	1	3	1	1	1	1
%	78%								11%			8%			3%			

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS



Source: PPMI.

1.2. Task 2: Shortlist of SOC operations for which causalities of impact can be attributed

In Task 2, we used the results of Task 1 to develop a shortlist of SOC schemes to be analysed in subsequent tasks. The task focused on identifying specific schemes for which causalities of impact could be attributed. The presence of these attributes served as an additional 'filter' to select the final sample of SOC cases.

Rigorous monitoring and evaluation, necessary to establish the basis for payment, is a key feature of SOC, distinguishing them from traditionally financed (TF) social services.

Nonetheless, attributing causality of impact is a significant challenge for many SOC schemes. Sophisticated experimental or quasi-experimental designs – the gold standard for impact attribution – are very difficult to carry out in practice. More often, alternative evaluation approaches are employed, documenting ‘contribution’ rather than linear cause-and-effect⁶⁷⁷. We therefore applied a two-level definition of the attribution of causality of impact in our analysis:

- According to the **broader definition**, applied in Task 2, the attribution of causality of impact is the **existence** of reasonable outcomes/impact measurement (i.e. outcomes evaluation), linking the achieved outcomes to the intervention. We elaborate on its possible forms below, based on the guidelines of international organisations such as the UN and OECD.
- According to the **strict definition** (a positivist approach adopted by some researchers on SOC issues), attribution of causality of impact can be evaluated only after inspection of the methods applied (usually **experimental** or **quasi-experimental designs**). To establish such an attribution of causality, we need to carefully examine the indicators used, all of the intervening factors and effects considered by the evaluators such as deadweight (outcomes that would still have been achieved if there had been no intervention), drop-off, displacement, competing services (support provided by other providers outside the intervention either to the target or to control groups or both)⁶⁷⁸, spill-overs, attrition, or contamination⁶⁷⁹. To carry out this review, we needed to analyse evaluations in-depth. This was carried out during Task 3, so we moved this activity under Task 3 as well.

When assessing whether the attribution of the causality of impact was possible for individual SOC schemes in the **broader** sense, we considered the following:

- In SOC schemes, the monitoring and evaluation methodology should be developed from the **early stages** of the project design, in order to define the baseline, align with the intervention logic, and establish predictors for setting the targets and measuring future outcomes. We used the appointment of an (independent) evaluator before launching the scheme as the key proxy to identify this.
- There are three **design options** that address causal attribution and the level of this attribution: experimental designs (with a control group created through random assignment); quasi-experimental designs (with a comparison group designed through matching, regression discontinuity, propensity scores or other means); and non-experimental designs (which look systematically at whether the evidence is consistent with what would be expected if the intervention were producing the impacts, and also whether other factors could provide an alternative explanation)⁶⁸⁰. The desk research and literature review concerning the existing SOC schemes indicated that experimental or even quasi-experimental designs are rare⁶⁸¹. We therefore included some evaluations that used non-experimental designs into the shortlist as well, but prioritised experimental and quasi-experimental evaluations.
- Evaluations of social services are often based on a **mixture of quantitative and qualitative methods** (also taking into account the subjective views and experiences of stakeholders and beneficiaries). Using the broader definition, we considered both

⁶⁷⁷ Fraser, A., Tan, S., Kruithof, K., Sim, M., Disley, E., Giacomantonio, C., Lagarde, M. & Mays, N. (2018). Evaluation of the Social Impact Bond Trailblazers in Health and Social Care: Final Report. Policy Innovation Research Unit. Retrieved from: <https://piru.lshtm.ac.uk/assets/files/SIB%20Trailblazers%20Evaluation%20final%20report.pdf>

⁶⁷⁸ OECD. (2016). Understanding Social Impact Bonds. Retrieved from: <http://www.oecd.org/cfe/leed/UnderstandingSIBsLux-WorkingPaper.pdf>.

⁶⁷⁹ Instiglio (2017).

⁶⁸⁰ Regiers, Patricia. (2014). Overview: Strategies for Causal Attribution. UNICEF. Retrieved from: https://www.unicef-irc.org/publications/pdf/brief_6_overview_strategies_causal_attribution_eng.pdf

⁶⁸¹ Fraser et al. (2018).

approaches equally valid, if the evaluation methodology met other criteria for establishing causal attribution.

- Evaluations of outcomes can **take time to be realised**, thus making causality harder to assess and establish. Some SIB interventions are evaluated several times, to assess medium-term and long-term outcomes. In such cases, when a long-term analysis was not yet available, we further investigated the evaluations of medium-term outcomes (with the aim of not excluding many potentially rich and illustrative cases).

Process-wise, to accurately characterise the designs of SOC operations for the purposes of selection, we applied a four-step approach:

- First, we conducted **focused desk research** on each SOC scheme from the long-list (compiled in Task 1, see Annex 3), during which we collected their evaluation reports.
- If evaluation reports were inaccessible online (which was often the case in most countries other than the UK), we then screened related policy documents, online media publications (in most cases, SOCs had received quite a lot of media attention), expert articles, and relevant websites. Once we found hints that rigorous evaluations had been completed, we **contacted the case owners** for more details. While some case owners did not respond, others replied and either shared the evaluations with us or informed us that the evaluations were still to be completed.
- After having collected all the necessary information, we **scanned the evaluations** to review the identified SOC cases against the criteria described above. Given that the study aimed to compare SOC models with TF interventions, we also aimed to identify equivalent TF models for comparison, by: 1) exploring the control groups in evaluation reports if control-group design was available; and 2) engaging in additional desk research about TF interventions in the same geographical location as the SOC models.
- If the SOC scheme qualified according to the selection criteria, it was included in the short list of operations to be explored in further stages of the study. This list included the title of the intervention, the reasons why it was included, and a potential TF scheme for comparison.

Using this four-step approach, we identified a total of **41** SOC schemes for which evaluations were either available online or had been shared by stakeholders who had responded to our requests. After gathering these evaluations, we shortlisted the most relevant and interesting SOC cases. During the shortlisting process, we considered the design options, methods used, data richness and availability, as well as other factors such as the involvement of NGOs and organisations that work on social innovation, co-design, and – for the sake of identifying a comparable traditionally funded (TF) programme – whether similar TF programmes had been included for the purposes of providing a control group, or could be identified through desk research. We strived to maintain objectivity by including both successful and unsuccessful cases, in terms of whether or not they had achieved their targets.

The list of cases prepared is presented in the table below, which also includes the reasons for each scheme's inclusion in the study.

Table 71. Shortlisted SOC schemes

#	Programme	Scheme	Social service	Country	Duration	Reasons for shortlisting	Possible TF intervention
1	Perspektive:Arbeit – (economic and social empowerment for women affected by violence (Gewaltschutzzentrum Upper Austria; Frauenhaus Linz)	SIB	Crisis and emergency	Austria	2015 – 2018 (3years)	<ul style="list-style-type: none"> • NGO involvement • Clear results (unsuccessful) • Theory of change • Geographical diversity 	TBD
2	Eleven Augsburg – youth unemployment	SIB	Labour market-related	Germany	2013 – 2015 (2.5 years)	<ul style="list-style-type: none"> • Co-production (involvement of youth in recruitment efforts) • Multiple viewpoints considered • Two evaluations 	<ul style="list-style-type: none"> • If a similar programme could not be found, one Germany's active labour market policies (training for the unemployed, wage subsidies, direct provision of jobs) could be studied
3	DUO for a JOB – migrant unemployment	SIB	Labour market-related	Belgium	2014 – 2017 (3 years)	<ul style="list-style-type: none"> • Control group design • Contact for interviews available • Geographical diversity 	TBD
4	BOASs Werkt – cross-border unemployment	SIB	Labour market-related	Netherlands & Germany	2016 – 2018 (2 years)	<ul style="list-style-type: none"> • Control group design • Contact with an interviewee • Geographical heterogeneity 	TBD

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

#	Programme	Scheme	Social service	Country	Duration	Reasons for shortlisting	Possible TF intervention
5	Buzinezzclub, Rotterdam – youth unemployment	SIB	Labour market-related	Netherlands	2013 – 2017 (4 years)	<ul style="list-style-type: none"> • Contact with interviewee • Geographical heterogeneity • Abundance of secondary data (though not yet evaluated) 	TBD
6	JobPath – long-term unemployment	PbR	Labour market-related	Ireland	2015-	<ul style="list-style-type: none"> • Quasi-experimental research design: econometric impact evaluation (control group) 	TBD
7	Mental Health and Employment Partnership (MHEP) – unemployment among people with mental health issue and other groups with health conditions and disabilities	SIB	Labour market-related	UK	2016 – 2019 (3 years)	<ul style="list-style-type: none"> • Clear methodology (survey and interview-based) • Information about costs and comparison with traditional models 	<ul style="list-style-type: none"> • Blue Light Programme
8	Drug and Alcohol Recovery Pilots – treatment for alcohol and drug misuse	PbR	Social exclusion	UK	2010 – 2012 (1.5 years)	<ul style="list-style-type: none"> • Quasi-experiment: difference-in-differences method • Comparison with traditionally financed schemes 	<ul style="list-style-type: none"> • Pathways to Employment • Progress2Work & Progress2work-LinkUP
9	Provider-led Pathways to Work – unemployment among people claiming benefits for health reasons	PbR	Labour market-related	UK	2007 – 2011 (4 years)	<ul style="list-style-type: none"> • Quasi-experiment: difference-in-differences method • Quantitative data available 	<ul style="list-style-type: none"> • The Workstep employment programme • Work Preparation • Job Introduction Scheme

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

#	Programme	Scheme	Social service	Country	Duration	Reasons for shortlisting	Possible TF intervention
10	Transforming Rehabilitation – post-sentence probation services for offenders serving short sentences	PbR	Social exclusion	UK	2013 – 2020 (8 years)	<ul style="list-style-type: none"> A mix of primary and secondary, qualitative and quantitative data 	<ul style="list-style-type: none"> Part of the scheme used fee-for-service approach – could attempt comparison with that, but otherwise: Managing offenders on short custodial sentences, or Minerva project in Hull
11	Work Programme – unemployment services	PbR	Labour market-related	UK	2011-2017 (7 years)	<ul style="list-style-type: none"> Mixed-methods evaluations on the programme itself and on the commissioning model 	<ul style="list-style-type: none"> Flexible New Deal Pathways to Work pilot (Jobcentre led)
12	Youth Contract – youth unemployment	PbR	Labour market-related	UK	2012-2015	<ul style="list-style-type: none"> Quasi experimental research design: impact assessment based on matched control group Mixed-methods evaluation 	TBD
13	Troubled Families (Phase II) – services for families with multiple problems	PbR	Specific issues, labour market-related	UK	Phase I: 2012-2015 Phase II: 2015-ongoing	<ul style="list-style-type: none"> Phase II mixed methods evaluation 	<ul style="list-style-type: none"> The Respect Programme (2006-2010)
14	Adolescent Behavioral Learning Experience (ABLE) Programme at Rikers Island – recidivism among young offenders	SIB	Social exclusion	US	2012 – 2015 (3 years)	<ul style="list-style-type: none"> Quasi-experimental research design Geographical diversity 	TBD

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

#	Programme	Scheme	Social service	Country	Duration	Reasons for shortlisting	Possible TF intervention
15	The Benevolent Society Social Benefit Bond – intensive family support services	SIB	Caring obligation	Australia	2013 – 2018 (5 years)	<ul style="list-style-type: none"> • Two-stage evaluation • A comparison to the control group • Geographical diversity 	TBD

1.3. Data collection process

Having identified and finalised the list of schemes for which evaluations were accessible, we began our data collection process. Data collection was conducted for each of the schemes in the shortlist, as well as for their equivalent TF schemes. The process consisted of **desk research, interviews with relevant stakeholders, and data entry into a standardised data collection template**. These activities were conducted in English, as well as in the national languages of the countries in which certain selected SOC schemes were located.

Desk research

Our **desk research** began by looking at databases that compile information on SOC schemes, including Social Finance⁶⁸², Instiglio⁶⁸³, the projects database at Oxford University's Government Outcomes Lab⁶⁸⁴, and Pay for Success⁶⁸⁵, among others. We then implemented targeted searches for relevant studies and media messages on each scheme selected, using a similar methodology to that described under Task 1. The desk research was conducted in English as well as in national languages. We validated the information collected by reviewing the individual reports and evaluations of the selected schemes.

In parallel, we conducted desk research to identify **TF interventions** that were similar to each of SOC schemes on our short list. Depending on the existence of comparable TF interventions and the availability of data, we utilised three possible ways to match comparable SOC and TF schemes (see the figure below):

- A. Equivalent TF social interventions (in terms of goals and activities) that were in place before the introduction of a particular SOC scheme;
- B. TF social interventions that substituted successful SOC schemes;
- C. A broader approach to matching SOC and TF schemes implemented in similar contexts based on several similarity criteria, such as:
 - social problems that the intervention aims to solve, target group of beneficiaries;
 - same service provider;
 - amount of funding;
 - time span of the intervention;
 - presence of measured social outcomes objectives, possibly reported as changes on the direct beneficiaries;
 - macroeconomic and contextual conditions (e.g., institutional and legal background).

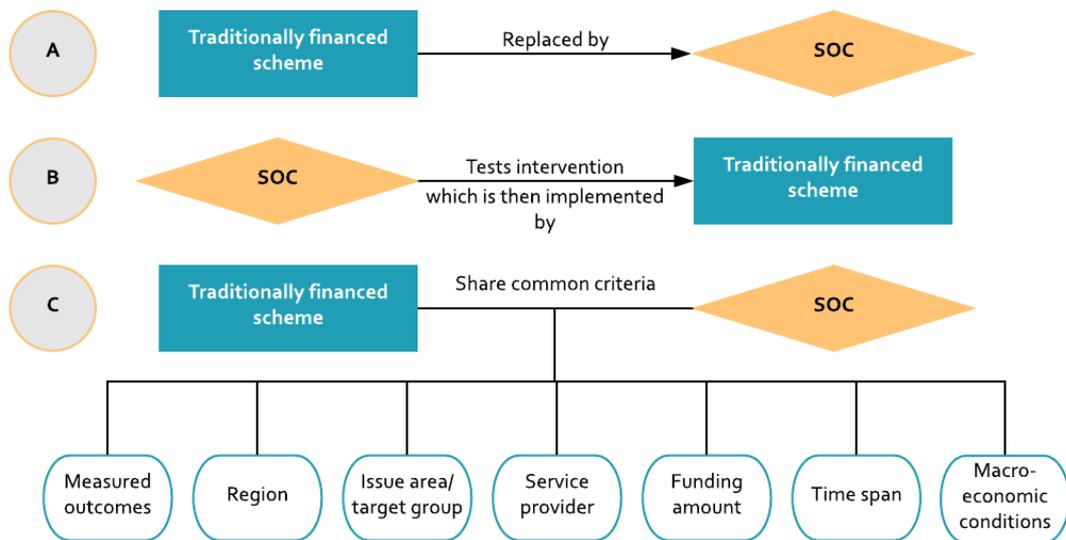
⁶⁸² Social Finance. (2019). *Impact Bond Global Database*. Retrieved from: <https://sibdatabase.socialfinance.org.uk/>.

⁶⁸³ Instiglio. (2019). *Impact Bonds Worldwide*. Retrieved from: <https://www.instiglio.org/en/sibs-worldwide/>.

⁶⁸⁴ Government Outcomes Lab (2019d). *Projects database*. University of Oxford Blavatnik School of Government. Retrieved from: https://golab.bsg.ox.ac.uk/knowledge-bank/project-database/?query=&contract_type=Social+Impact+Bond.

⁶⁸⁵ Pay for Success. (2019). *Projects*. Nonprofit Finance Fund. Retrieved from: <https://www.payforsuccess.org/projects/>.

Figure 16. Different options for matching SOC and TF schemes



Source: PPMI.

To identify and explore each TF scheme, we asked our interviewees for possible TF equivalents to the SOC programme in question, as well as looking for information in SOC evaluations (sometimes evaluators compared the SOC intervention with a real-life TF counterpart as part of the evaluation methodology). General Google searches were also used (with keywords relating to the policy field of a selected SOC scheme, target group, other descriptors of the intervention), as well as municipal or national policy documents and budgets (which were publicly available), and relevant project/ annual activity reports. If the researchers identified several TF schemes that could be used for comparison with a specific SOC intervention, we either selected the one with most relevant information available; or we compared relevant aspects of an SOC scheme with equivalent aspects of multiple TF schemes.

Interviews

Having assembled the background information on each scheme, we interviewed stakeholders who had contributed to their implementation. The key interviewees for each scheme included commissioners, service providers, investors and/or intermediaries as well as the evaluators of SOC schemes. Furthermore, for those SOC-TF pairs that had different commissioners, we also reached out to TF commissioners separately if insufficient information was available from our desk research. Specific persons or organisations relating to particular schemes were identified via desk research. Interviews were conducted both face-to-face (in cases when interviewees were based in the same city as our data collection team members, and prior to the Covid-19 restrictions) and by phone.

In total, **54** interviews were conducted. As the table below illustrates, we conducted most of the interviews with commissioners, service providers and evaluators. This is because these three stakeholder groups are common for both SIBs and PbR schemes. Unfortunately, we were not able to conduct interviews with every stakeholder group for every scheme. Despite reaching out to interviewees multiple times, the fact that much of the fieldwork was conducted during the COVID-19 pandemic meant that some stakeholders – commissioners and service providers in particular – were not available. Where interviews with these stakeholders could not be conducted, we reached out to other experts. These included representatives from the same commissioning authority who had not been directly involved

in the selected SOC programme but were nevertheless familiar with the case, as well as scholars who have studied the selected case.

Table 72. Number of Interviews completed with various stakeholder groups

Stakeholder Type	Number of Interviews
Commissioner	14
Service provider	14
Evaluator	13
Investor	6
Intermediary	4
Expert	3
Total	54

Questionnaires were tailored to **specific interviewees**, depending on their role in the selected schemes and level of involvement. For example, representatives from commissioning organisations were best positioned to provide detailed information about each of the schemes, as well as helping to identify similar TF schemes. Meanwhile, evaluators provided most of the information to cover Task 3. Our semi-structured questionnaires included both **factual and attitudinal** questions. Specific questions are outlined under each task in subsequent sections.

The information collected through desk research and interviews was entered into a **standardised data collection template** according to the guidelines developed by PPMI. This was provided to all researchers involved in data collection, along with methodological guidelines for data collection and entering data into the template. The template allowed us to make comparisons between SOC and TF schemes, which was particularly relevant for Tasks 4 and 5, as well as comparisons **across different SOC schemes**, which was necessary for Tasks 3-6.

1.4. Task 3: Mapping and review of outcomes measurement methods

Task 3 consisted of mapping and reviewing the outcomes measurement methods used in the SOC schemes identified in Task 2. The framework used to map the evaluation methods is presented in Chapter 2 of the main report. The data collection process used to implement Task 3 is described above, but evaluation reports, interviews with the commissioners and evaluators of the schemes were especially important. Interview questions are presented below.

Interview questions relevant to Task 3

- Could you tell us more about the rationale for choosing the measurement method adopted to evaluate the scheme?
- How were the particular indicators and targets chosen in order to measure the intervention outcome(s)?

- What intervening factors (if any) were considered in the evaluation? In other words, were there other factors contributing to the achieved outcome that were not measured?
- Which were the enabling factors and barriers encountered during the evaluation? Which factors were related to the method used and which ones to other issues (for example, stakeholder communication)?
- What is the main strength and weakness of the measurement method adopted to evaluate the scheme? Why?
- Would you measure outcomes differently now that the project has ended? If so, why? In case of more than one outcome, which one?

1.5. Task 4: Effectiveness and added value

Task 4 assessed the effectiveness and added value of interventions delivered under the identified SOC schemes, in comparison to the interventions delivered under traditional financing models. The standard definition of *effectiveness* is the ability of an action to accomplish specific objectives defined *ex-ante*⁶⁸⁶. Meanwhile, *added value* literally refers to an improvement or addition to something that makes it worth more; in other words, it includes all those features/consequences added to increase the value of the intervention. Here, we focused on whether the use of SOC funding increased the innovativeness of the service provided, and whether it enhanced user satisfaction.

We operationalised this task using the interview questions presented in the box below (data on which were also collected through desk research). Furthermore, we collected statistical information on each SOC and TF scheme, namely: outcome achievement rates, number of participants, and information on user satisfaction from participant surveys in their evaluation reports.

Interview questions relevant to Task 4

To commissioner:

- Why did you choose to fund the social intervention through an outcomes-based contract?
- How did you select the service provider?
- Which are, in your opinion, the additional benefits of funding the intervention through an outcomes-based contract compared to a traditional procurement procedure?
- Was this specific social service already provided by the public sector?
- Were there any other government agencies or levels that benefitted from the intervention, but were not involved in funding it? Did you try to involve them in the contract? If yes, how did you go about this process?
- What benefits were delivered because the intervention was funded through an outcomes-based contract and not traditional procurement?
- Did the intervention result in any drawbacks because it was funded through an outcomes-based contract and not traditional procurement?

⁶⁸⁶ Better Regulation Toolbox.

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

- Have there been any attempts to replicate or scale the intervention delivered under SOC? If so, was the intervention replicated using SOC or traditional funding mechanisms (public procurement)? Why? What were the challenges in the attempt to replicate the intervention? How were they overcome?
- [For schemes carried out in the European Union only] Are you aware of the changes that resulted from the public procurement Directive 2014/24/EU? How did they affect your work, including the ability to fund projects through outcome-based contracts?

To service provider:

- Has the intervention funded by the SOC scheme been implemented before? How, in terms of processes and funding? Was it financially sustainable?
- [if not implemented before] Do you think the intervention would have been implemented anyway without the SOC scheme? How?
- Why did you engage in the SOC scheme?
- Do you think the service funded by the SOC scheme is able to meet the needs of the beneficiaries better than other available services? Why yes or why not?
- Which duties in terms of reporting did you have to accomplish?

To investor:

- [For UK schemes only]: Did you use a tax exemption on your investment? Did it motivate you to participate in the project? Would you have participated if the tax exemption did not exist?

1.6. Task 5: Analysis of efficiency

In Task 5, we collected information to assess both the intervention costs and operational costs of each SOC intervention, and compared them with the equivalent costs of the comparable TF interventions. The operationalisation of these costs and the methods used to estimate them are summarised in Chapter 5. Below we outline the indicators that were collected to estimate these costs, as well as relevant interview questions to relevant stakeholders. Information for this task was collected through desk research, review of evaluation and audit reports, Freedom of Information requests to the commissioning departments, and interviews.

Table 73. Quantitative indicators gathered for Task 5

Indicator	SOC	TF
Overall cost	Yes	Yes
Total cost per beneficiary	Yes	Yes
Public costs per beneficiary	Yes	Yes
Cost per successful beneficiary	Yes	Possibly ⁶⁸⁷
Set-up cost (in terms of money and/or working hours/days)	Yes	Possibly

⁶⁸⁷ If outcome measures had been set, and evaluations conducted.

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

Indicator	SOC	TF
Intermediary fees	Yes	Possibly
Legal fees	Yes	Possibly
Working hours required to coordinate between the stakeholders	Yes	Yes
The number of organisations involved	Yes	Yes
The number of people involved	Yes	Yes
Evaluation costs	Yes	Possibly
Working hours required to oversee the evaluation	Yes	Possibly

Source: PPMI.

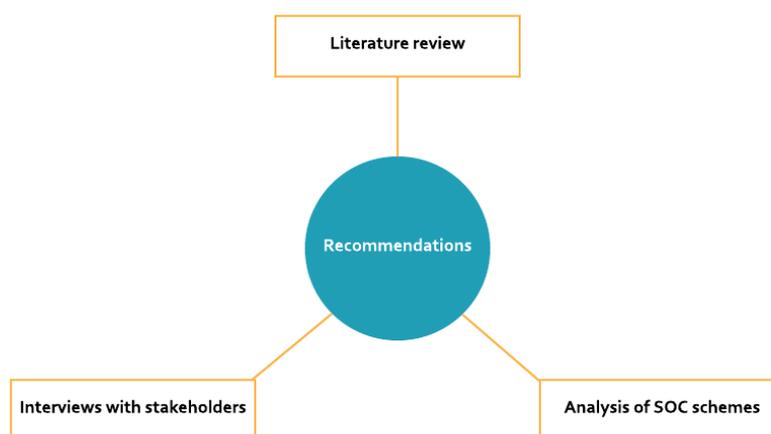
Interview questions relevant to Task 5

- In your own accounting, how did you keep and estimate the costs of the scheme (e.g., what categories of costs did you take into account)?
- Did you estimate costs per beneficiary/ costs per successful beneficiary? Do they differ from traditionally financed interventions?
- Do you consider that the SOC intervention has been cost efficient? Why?
- Are the set-up and design costs of the SIB/PbR model included in the overall price of the [name of the SOC intervention]? What are these costs? How would you assess them?
- Could you tell me more about the process of setting up the [SOC scheme]? Was the process easier or more difficult compared to the scheme funded through public procurement? Why?
- How much oversight did interventions implemented under each scheme require? Was the related number of working hours higher or lower under the SOC/SIB/PbR model compared to traditional procurement? Why?
- What about the time spent coordinating with different stakeholders under each scheme? Was the communication between all parties involved more time-consuming under one funding scheme than the other? Why or why not?
- What about human resources required to carry out each scheme? Were they equivalent? Why or why not? Did you need a person working full-time on coordinating the process?
- Has the intervention delivered through traditional procurement received an evaluation? If so, what was the cost of that evaluation? How does it compare to the evaluation performed for the [SOC scheme] content- and cost-wise?
- What are the [SOC scheme] design features that make it particularly efficient or inefficient?

1.7. Task 6: Policy recommendations

Task 6 of this study consisted of formulating recommendations at policy and contractual levels regarding outcomes measurement methods, areas/target groups for which the most impact/innovation could be achieved, and the most effective and efficient design models for SOC schemes. In order to provide sound policy recommendations, we built on the study findings and used **data triangulation**. We validated data from the literature review that was specific to policy and contractual recommendations against our analysis of SOC schemes, as well as the viewpoints of relevant stakeholders, as illustrated in the figure below.

Figure 17. Data triangulation to produce recommendations



Source: PPMI.

In order to validate our insights against the opinions of the stakeholders interviewed, we also asked interviewees directly about the **lessons learned** from deploying SOC schemes. While many of the questions regarding policy-level recommendations were covered during earlier tasks, here we focus specifically on the contractual level. This also yielded useful insights for Chapter 6 on the design of SOC schemes. The questions used are presented in the box below.

Interview questions relevant to Task 6

- To the commissioner:
- Did you contract someone directly or was evaluation sub-contracted by another entity? Why did you choose that option?
- Looking back on the process, would you specify anything additional in the contract?
- Do you think some social areas and/or target groups are better for SOC interventions than others? Why?
- What factors were crucial for success of the scheme? What factors (institutional, legal, resources, knowledge, etc.) were crucial for creating the scheme in the first place? Without which of them you would not have been able to succeed?
- Generally, how would you summarise the lessons learned?

1.8. Limitations

The study involves a number of limitations.

- First, our study results are heavily dependent on the information derived from the 15 SOC cases and their TF pairs, which makes them difficult to generalise to other SOC schemes. Nevertheless, we engaged in a review of the literature (see Task 6) in order to assess the extent to which our findings are in line with findings of other scholars.
- Second, where we observe different levels of outcomes being achieved by similar SOC and TF programmes, or where we find evidence that the costs per outcome differ between the two models, we cannot claim that these differentials are strictly because of dissimilarities in the funding mechanisms. This is due to the presence of various intervening factors (different locations, services provided, stakeholders involved, macroeconomic conditions, etc.) that might have affected the effectiveness of the two programmes.
- Third, while we attempted to collect equivalent information on all SOC schemes, at times such information was not available. This is particularly true with regard to Task 4, where the same outcomes were not measured for the SOC and TF pairs, as well as Task 5, where operational or intervention costs were not recorded or publicly available. Where possible, we attempted to fill the knowledge gaps with qualitative information from interviews, or by using alternative indicators (for example, comparing the programmes' outreach in terms of the number of participants engaged instead of comparing programme outcomes in Task 4). Nevertheless, some knowledge gaps remain. These are indicated throughout the report.

Annex 3: Long list of SOC schemes

#	Programme	Scheme	Social Service	Country	Duration	Source/s
EUROPE						
1	KOTO-SIB – fast employment and integration of immigrants	SIB	Labour market-related	Finland	2016 – 2019 (4 years, with a follow up until the end of 2022)	https://tem.fi/artikkeli/-/asset_publisher/uuden-kokeilun-tavoitteena-tyollistaa-tuhansia-maahanmuuttajia https://golab.bsg.ox.ac.uk/news-events/blogs/using-social-impact-bonds-integrate-refugees-finland-will-it-work/ https://www.nbforum.com/nbreport/sib-funds-in-a-nutshell/
2	Perspektive:Arbeit – economic and social empowerment for women affected by violence (Gewaltschutzzentrum Upper Austria; Frauenhaus Linz)	SIB	Crisis and emergency	Upper Austria region, Austria	2015 – 2018 (3 years)	https://www.alpine-space.eu/projects/alpsib/en/about/sib-good-practices/first-sib-in-austria
3	Eleven Augsburg - youth unemployment	SIB	Labour market-related	Augsburg, Germany	2013 – 2015 (2.5 years)	https://www.eleven.ngo/blog/begleitevaluation-sib-augsburg-veroeffentlicht
4	DUO for a JOB – migrant unemployment	SIB	Labour market-related	Brussels, Belgium	2014 – 2017 (3 years)	https://www.duoforajob.be/en/home/
5	Buzinezzclub – youth unemployment	SIB	Labour market-related	Rotterdam, Netherlands	2013 – 2017 (4 years)	https://apolitical.co/solution_article/rotterdam-gets-young-people-off-benefits-social-impact-bond/ http://www.popvox.nl/en/buzinezzclub/
6	Buzinezzclub – youth unemployment	SIB	Labour market-related	Eindhoven, Netherlands	June 2016 – 2019 (3 years)	https://eindhovennews.com/news/2016/10/buzinezzclub-should-help-young-people-into-work/
7	Buzinezzclub – youth unemployment	SIB	Labour market-related	The Hague, Netherlands	2014 –2016 (2 years)	N/A

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

#	Programme	Scheme	Social Service	Country	Duration	Source/s
8	Buzinezzclub – youth unemployment	SIB	Labour market-related	Drechtsteden, Netherlands	2015 – 2017 (2 years)	https://www.dordrecht.net/nieuws/2017-03-29-19957-buzinezzclub-stopt-wegens-succes.ht-ml
9	BOAS Werkt – cross-border unemployment	SIB	Labour market-related	Enschede, Netherlands and Rheine, Germany	2016 – 2018 (2 years)	https://medium.com/social-finance-uk/worlds-first-binational-social-impact-bond-fosters-dutch-german-collaboration-c5805dccb04b http://www.mondovisione.com/media-and-resources/news/first-cross-border-social-impact-bond-helps-dutch-unemployed-find-german-jobs/
10	The Colour Kitchen – youth unemployment	SIB	Labour market-related	Utrecht, Netherlands	2015 – 2019 (4 years)	https://www.societyimpact.nl/second-dutch-social-impact-bond-launches-monday-april-20th/ https://www.abnamro.com/en/newsroom/press-releases/2015/major-financial-boost-for-youth-employment-in-utrecht.html
11	Werk na detentie ('Work after prison')	SIB	Social exclusion, labour market-related	Netherlands	2016 – 2018 (2.5 years)	https://www.brookings.edu/blog/education-plus-development/2016/06/09/the-netherlands-leads-again-in-social-innovation-with-announcement-of-fifth-social-impact-bond/
12	Short-Sentenced Offenders The One Service	SIB	Social exclusion	Peterborough, United Kingdom	2010 – 2017 (7 years)	https://www.rand.org/randeurope/research/projects/social-impact-bonds.html https://www.gov.uk/government/publications/final-results-for-cohort-2-of-the-social-impact-bond-payment-by-results-pilot-at-hmp-peterborough http://www.russellwebster.com/disappointing-outcomes-for-peterborough-and-doncaster-prison-pbr-pilots/ https://golab.bsg.ox.ac.uk/knowledge/project-database/hmp-peterborough-one-service/
13	DWP Innovation Fund Round 1: Perth & District YMCA	SIB	Labour market-related	Perthshire and Kinross, United Kingdom	2012 – 2015 (3.5 Years)	https://www2.gov.scot/Topics/Built-Environment/Housing/investment/innovationfund/innovation https://www.socialfinance.org.uk/sites/default/files/publications/portuguese-social-investment-taskforce.pdf (p.26) https://golab.bsg.ox.ac.uk/knowledge/project-database/dwp-innovation-fund-round-i-scotland-perthshire-kinross-living-balance/
14	DWP Innovation Fund Round 1: BEST Network	SIB	Labour market-related	West Midlands, United Kingdom	2012 – 2015 (3.5 Years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/dwp-innovation-fund-round-i-west-midlands/
15	DWP Innovation Fund Round 1: Nottingham Futures	SIB	Labour market-related	Nottingham, United Kingdom	2012 – 2015 (3.5 Years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/dwp-innovation-fund-round-i-nottingham/

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

#	Programme	Scheme	Social Service	Country	Duration	Source/s
16	DWP Innovation Fund Round 1: Career Connect	SIB	Labour market-related	Greater Merseyside, United Kingdom	2012 – 2015 (3.5 Years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/dwp-innovation-fund-round-i-greater-merseyside-new-horizons/
17	DWP Innovation Fund Round 1: Links for Life	SIB	Labour market-related	East London, United Kingdom	2012 – 2015 (3.5 Years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/dwp-innovation-fund-round-i-stratford-other-parts-east-london/
18	DWP Innovation Fund Round 1: Think Forward/ Tomorrow's People	SIB	Labour market-related	East London, United Kingdom	2012 – 2015 (3.5 Years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/dwp-innovation-fund-round-i-east-london-tomorrows-people/
19	DWP Innovation Fund Round 2: Dyslexia Action & Include	SIB	Labour market-related	Cardiff and Newport, United Kingdom	2012 – 2015 (3.5 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/dwp-innovation-fund-round-ii-wales-cardiff-newport/
20	DWP Innovation Fund Round 2: Prevista	SIB	Labour market-related	West London, United Kingdom	2012 – 2015 (3.5 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/dwp-innovation-fund-round-ii-west-london-prevista/
21	DWP Innovation Fund Round 2: Teens and Toddlers	SIB	Labour market-related	Greater Manchester, United Kingdom	2012 – 2015 (3.5 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/dwp-innovation-fund-round-ii-greater-manchester-teens-and-toddlers/
22	DWP Innovation Fund Round 2: Adviza	SIB	Labour market-related	Thames Valley, United Kingdom	2012 – 2015 (3.5 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/dwp-innovation-fund-round-ii-thames-valley-energise/
23	Fair Chance Fund: P3, The Y and YMCA Derbyshire	SIB	Crisis and emergency, labour market-related	Leicestershire, United Kingdom	2015 – 2018 (3 years)	https://www.gov.uk/government/news/23-million-to-help-homeless-turn-around-their-lives https://golab.bsg.ox.ac.uk/knowledge/project-database/fair-chance-fund-leicestershire-ambition-east-midlands/
24	Fair Chance Fund: St. Basil's	SIB	Crisis and emergency, labour market-related	Birmingham, United Kingdom	2015 – 2018 (3 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/fair-chance-fund-birmingham-rewriting-futures/

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

#	Programme	Scheme	Social Service	Country	Duration	Source/s
25	Fair Chance Fund: Fusion Housing	SIB	Crisis and emergency, labour market-related	West Yorkshire, United Kingdom	2015 – 2018 (3 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/fair-chance-fund-west-yorkshire-fusion-housing/
26	Fair Chance Fund: Local Solutions	SIB	Crisis and emergency, labour market related	Liverpool, United Kingdom	2015 – 2018 (3 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/fair-chance-fund-liverpool/
27	Fair Chance Fund: P3 & CCP	SIB	Crisis and emergency, labour market-related	Gloucestershire, United Kingdom	2015 – 2018 (3 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/fair-chance-fund-gloucestershire-aspire-gloucester/
28	Fair Chance Fund: depaul UK	SIB	Crisis and emergency, labour market-related	Manchester, Rochdale, Oldham and Royal Borough of Greenwich, United Kingdom	2015 – 2018 (3 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/fair-chance-fund-manchester-rochdale-oldham-royal-borough-greenwich-depaul-uk/
29	Fair Chance Fund: Home Group	SIB	Crisis and emergency, labour market-related	Newcastle, United Kingdom	2015 – 2018 (3 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/fair-chance-fund-newcastle-home-group/
30	Youth Engagement Fund: Futureshapers Sheffield	SIB	Social exclusion	Sheffield & Greater, Manchester, United Kingdom	2015 – 2018 (3 years)	https://www.gov.uk/government/news/new-social-impact-bonds-to-support-public-services https://golab.bsg.ox.ac.uk/knowledge/project-database/youth-engagement-fund-sheffield-futureshapers-sheffield/
31	Youth Engagement Fund: Career Connect	SIB	Social exclusion	Greater Merseyside, United Kingdom	2015 – 2018 (3 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/youth-engagement-fund-greater-merseyside-career-connect/
32	Youth Engagement Fund: Prevista Ltd	SIB	Social exclusion	London, United Kingdom	2015 – 2018 (3 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/youth-engagement-fund-london-prevista/
33	Youth Engagement Fund: Teens and Toddlers	SIB	Social exclusion	Greater Manchester, United Kingdom	2015 – 2018 (3 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/youth-engagement-fund-greater-manchester-teens-and-toddlers/

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

#	Programme	Scheme	Social Service	Country	Duration	Source/s
34	The Step Down Programme, Core Assets	SIB	Caring obligation	Birmingham, United Kingdom	2014 – 2018 (4 years)	https://www.cypnow.co.uk/cyp/news/1148468/social-impact-bond-funds-birmingham-fostering-scheme https://www.bridgesfundmanagement.com/bridges-ventures-backs-foster-placement-service-commissioned-birmingham-city-council/
35	Mental Health and Employment Partnership	SIB	Labour market-related	Haringey, Staffordshire and Tower Hamlets, United Kingdom	2016 – 2019 (3 years)	https://golab.bsg.ox.ac.uk/knowledge/case-studies/mhep/ https://www.socialfinance.org.uk/projects/mental-health-and-employment-mhep https://golab.bsg.ox.ac.uk/knowledge/project-database/mental-health-employment-partnership-haringey-staffordshire-tower-hamlets/
36	DCLG London Rough Sleepers, St. Mungo's Broadway	SIB	Crisis and emergency	London, United Kingdom	2012 – 2015 (3 years)	https://data.london.gov.uk/dataset/chain-reports http://www.russellwebster.com/roughsleepersib-final/ https://golab.bsg.ox.ac.uk/knowledge/project-database/dclg-london-rough-sleepers-street-impact/
37	DCLG London Rough Sleepers, Thames Reach	SIB	Crisis and emergency	London, United Kingdom	2012 – 2015 (3 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/dclg-rough-sleepers-thames-reach/
38	Single Homeless Prevention Service Thames Reach	SIB	Crisis and emergency	London Borough of Brent, United Kingdom	2017 – 2019 (2 years)	https://golab.bsg.ox.ac.uk/knowledge/project-database/single-homeless-prevention-service-shps-brent/
39	Drug and Alcohol Recovery Pilots	Pbr	Social exclusion	United Kingdom	2010 – 2012 (1 year)	https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/97806/pbr-lessons-learnt.pdf https://www.gov.uk/government/news/outcome-definitions-agreed-for-drugs-recovery-pilot-bids
40	Supporting People Pilots	Pbr	Crisis and emergency	Birmingham; Cheshire West and Chester; Derbyshire; Islington; Kent; Lewisham; Sheffield; Stockport; Southend-on-Sea; and Torbay, United Kingdom	2011 – 2014 (2 years)	http://www.russellwebster.com/supporting-people-and-payment-by-results/

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

#	Programme	Scheme	Social Service	Country	Duration	Source/s
41	HMP Doncaster Pilot	Pbr	Social exclusion	Doncaster, United Kingdom	2011 – 2014 (3 years)	Http://www.russellwebster.com/what-did-we-learn-from-the-doncaster-prison-pbr-reoffending-pilot/ http://www.russellwebster.com/disappointing-outcomes-for-peterborough-and-doncaster-prison-pbr-pilots/
42	Work Choice – Specialist Disability Employment Programme	Pbr	Labour market-related	Across the United Kingdom	2010 – 2012 (2 years)	https://publications.parliament.uk/pa/cm201516/cmselect/cmworpen/363/363.pdf (p.15) https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/210684/846summ.pdf
43	Provider-led Pathways to Work	Pbr	Labour market-related	Across the United Kingdom	2007 – 2011 (4 years)	https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/221223/WP113.pdf http://www.psi.org.uk/pdf/2010/rrep638.pdf https://www.nao.org.uk/wp-content/uploads/2015/06/Payment-by-Results-Learning-from-the-Literature.pdf (p.72)
44	Transforming Rehabilitation	Pbr	Social exclusion	Across the United Kingdom	2013 – 2018 (5 years)	http://www.russellwebster.com/crcpbroct17/ http://www.russellwebster.com/endtr/
45	Diabetes Programme Group	Pbr	Disabilities	Liverpool, United Kingdom	2010 – 2013 (3 years)	https://www.fticonsulting.com/~media/Files/us-files/insights/reports/an-outcome-based-contract-diabetes-care-liverpool.pdf
46	Troubled Families	Pbr	Specific problems, labour market-related	United Kingdom (Phase I in 152 English local authorities; Phase II in 113 local authorities)	Phase I: 2012 – 2015 Phase II: 2015 – 2021	http://bit.ly/2Sm2fA0 http://bit.ly/2ukgoxu http://bit.ly/2spxfpl http://bit.ly/38mfxcc http://bit.ly/31MVNW2
47	New Homes Bonus	Pbr	Housing	England, United Kingdom	2011 – ongoing	https://www.gov.uk/government/publications/evaluation-of-the-new-homes-bonus https://researchbriefings.parliament.uk/researchbriefing/Summary/SN05724 https://www.nao.org.uk/wp-content/uploads/2013/03/10122-001-New-Homes-Bonus_HC-1047.pdf
48	The Work Programme	Pbr	Labour market-related	England, Scotland and	2011 – 2017	https://www.employment-studies.co.uk/resource/work-programme-evaluation-participant-experience-report

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

#	Programme	Scheme	Social Service	Country	Duration	Source/s
				Wales (Great Britain)		https://www.nao.org.uk/report/the-work-programme/ https://researchbriefings.parliament.uk/researchbriefing/Summary/SN06340
49	The Youth Contract	Pbr	Labour market-related	United Kingdom	2012 – 2015	https://www.gov.uk/government/publications/youth-contract-report https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/207488/rrep833.pdf
50	Jobpath	Pbr	Labour market-related	Ireland	2015 – ongoing	https://www.gov.ie/en/policy-information/9e575c-jobpath/
AMERICAS						
1	Adolescent Behavioral Learning Experience (ABLE) Program at Rikers Island	SIB	Social exclusion	New York City, United States	2013 – 2016 (3 years)	https://www.frbsf.org/community-development/files/rikers-island-first-social-impact-bond-united-states.pdf https://www.mdrc.org/news/announcement/mdrc-statement-vera-institute-s-study-adolescent-behavioral-learning-experience https://www.huffpost.com/entry/what-we-learned-from-the-_b_7710272?Guccounter=1 https://payforsuccess.org/project/nyc-able-project-incarcerated-youth
52	Recidivism and Employment	SIB	Labour market-related	Rochester, United States	2013 – 2019 (5 years)	https://socialfinance.org/focus-areas/criminal-justice/nys-pay-for-success-financing/ https://www.governor.ny.gov/news/governor-cuomo-announces-new-york-first-state-nation-launch-pay-success-project-initiative
53	Sweet Dreams Project: Vulnerable Mothers EGADZ	SIB	Caring obligation	Saskatoon, Canada	2014 – 2019 (5 years)	https://www.saskatchewan.ca/government/news-and-media/2016/june/21/sweet-dreams-project
54	Indigenous Livelihoods and Rainforest Protection	SIB	Labour market-related	Peru	2015 – 2016 (1 year)	http://www.common-fund.org/wp-content/uploads/2017/05/Verification_Report.pdf http://www.common-fund.org/wp-content/uploads/2017/05/Verification_Report.pdf
55	Empowerment of rural population in Colombia	Pbr	Labour market-related	Colombia	2017 – 2017 (7 months)	https://www.instiglio.org/en/rural-colombia/
56	Empleando Futuro	SIB	Labour market-related	Colombia	2017 – 2018 (21 months)	http://www.sibs.co/wp-content/uploads/2019/11/Agenda-aprendizajes-versionfinal.pdf

STUDY ON THE BENEFITS OF USING SOCIAL OUTCOME CONTRACTING IN THE PROVISION OF SOCIAL SERVICES AND INTERVENTIONS

#	Programme	Scheme	Social Service	Country	Duration	Source/s
57	Youth in Mexico	Pbr	Labour market-related	Mexico	2014 – 2015 (10 months)	https://www.instiglio.org/en/education-chiapas/
AUSTRALIA & ASIA						
8	Adoption, family care environments for children who require protective care	SIB	Caring obligation	Yokosuka City, Japan	2015 – 2016 (1 year)	https://www.nippon-foundation.or.jp/en/news/articles/2015/20151126-20952.html
59	Supporting employment for and long-term independence of young people	SIB	Labour market-related	Amagasaki City, Japan	2015 – 2016 (1 year)	https://www.nippon-foundation.or.jp/en/news/articles/2015/20151126-20952.html
60	Learning and Social Skill Support for Children at risk	SIB	Labour market-related	Yokohama City, Japan	2016 – 2018 (1.5 years)	NA
61	The Benevolent Society Social Benefit Bond	SIB	Caring obligation	New South Wales, Australia	2013 – 2018 (5 years)	https://emmatomkinson.com/2017/07/18/aussie-sibs/#SIB2
62	First SIB in Korea	SIB	Labour market-related, disabilities	South Korea	2015 – 2018 (3.5 years)	http://panimpact.kr/seoul-sib-fact-sheet/
AFRICA						
63	Employment in Morocco	Pbr	Labour market-related	Morocco	2015 – 2016 (1 year)	https://www.instiglio.org/en/employment-in-morocco/
64	Poverty in Burkina Faso	Pbr	Specific problems	Burkina Faso	2015 – 2015 (7 months)	https://www.instiglio.org/en/extreme-poverty-burkina/

References

- ABN Amro (2020). Social Impact Bond: investeren in een ander. Retrieved from <https://www.abnamro.com/nl/over-abnamro/in-de-maatschappij/duurzaamheid/financieren-en-beleggen/social-impact-bonds/index.html>
- Adam, S., Bozio, A., Emmerson, C., Greenberg, D. & Knight, G. (2009). A cost-benefit analysis of Pathways to Work for new and repeat incapacity benefits claimants. Retrieved from: <https://www.ifs.org.uk/publications/4231>
- Adebowale, V. (2004). A New Deal for all, report of the National Employment Panel's Working Group on New Deal 25 plus, London. Retrieved from: <https://webarchive.nationalarchives.gov.uk/20090410041007/http://www.workandpensions.gov.uk/ndpb/nep-pdfs/aneuwealforall.pdf>.
- Albertson, K., Fox, C., O'Leary, C. & Painter, G. (2018). Payment by Results and Social Impact Bonds: Outcome-based Commissioning in the USA and UK. Policy Press. Retrieved from: https://books.google.lt/books?id=1NFMDwAAQBAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false
- Allen, M. (2014). Local action on health inequalities: Reducing the number of young people not in employment, education or training (NEET). Public Health England & UCL Institute of Health Equity. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/356062/Review3_NEETs_health_inequalities.pdf.
- Anderson, J. & Phillips, A. (2016). What we learned from the nation's first social impact bond. *The Huffington Post*.
- ARTD Consultants (2014). *Evaluation of the Benevolent Society's Resilient Families service – Evaluation plan*.
- ARTD Consultants (2014). *Evaluation of the Resilient Families Service – Preliminary report*. NSW Treasury.
- ARTD Consultants (2016). *Evaluation of the Resilient Families service – Interim report*. NSW Treasury.
- ARTD Consultants (2017). *Evaluation of the Resilient Families service (social benefit bond pilot) – Evaluation stage 2 progress report*. NSW Treasury.
- Austrian Federal Government (2013). Arbeitsprogramm der österreichischen Bundesregierung 2013-2018 (Austrian Federal Government Work Programme 2013-2018).
- Austrian Federal Ministry of Labour, Social Affairs and Consumer Protection (2018). Press Statement (2018) "Erfolg des ersten österreichischen Social Impact Bond: PERSPEKTIVE:ARBEIT wird in den Regelbetrieb übergeführt" Tuesday, 27. November 2018
- Austrian Federal Ministry of Social Affairs, Health, Care and Consumer Protection (2019). Handbuch PERSPEKTIVE:ARBEIT – Anleitung zur Umsetzung des Projekts PERSPEKTIVE:ARBEIT – Ökonomisches und soziales Empowerment für gewaltbetroffene Frauen [Manual PERSPEKTIVE:ARBEIT – Guidance for the implementation of the project PERSPEKTIVE ARBEIT]
- Bardens, J. & Garton-Grimwood, G. (2013). Introducing "Payment by Results" in Offender Rehabilitation and other reforms. Retrieved from: <https://commonslibrary.parliament.uk/research-briefings/sn06665/>
- Bate, A, Bellis, A. & Loft, P. (2020). The Troubles Families Programme (England), Briefing Paper, Number 07585, 22 January 2020, House of Commons Library.

- Beale, I., Bloss, C. & Thomas, A. (2008). The longer-term impact of the New Deal for Young People. DWP Working Paper No 23. Retrieved from: <https://dera.ioe.ac.uk/7485/1/WP23.pdf>
- Becker, E., Hayllar, O. & Wood, M. (2010). Pathways to Work: programme engagement and work patterns. Research Report No 653. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/214423/rrep653.pdf.
- Begum, S. (2018). Transforming employment support for people with mental health conditions - An evaluation of two pilots in Camden.
- Benefits and Work (2009). Shaw Trust Loses Millions on Pathways. Retrieved from: <https://www.benefitsandwork.co.uk/news/1100-shaw-trust-loses-millions-on-pathways>
- Benevolent Society (2013). *Information memorandum*.
- Benevolent Society (2016). *The Benevolent Society Social Benefit Bond – Investor report*.
- Benevolent Society (2018). *Resilient families – Impact report 2013-2018*.
- Berlin, G.L. (2016). Learning from Experience: A Guide to Social Impact Bond Investing. MDRC.
- Big Issue Invest (2019a). Fund Management Social Enterprise Investment Fund II L.P. Annual Report. Retrieved from: <https://images.bigissueinvest.com/2019/01/SEIF-II-2018-TO-PRINT-V2.pdf>
- Big Issue Invest (2019b). "Mental Health and Employment Partnership case study", Webpage, 2019, <https://bigissueinvest.com/case-studies/mental-health-and-employment-partnership>
- Bogdanor, A. (2004). Not Working: Why Workfare Should Replace the New Deal. Policy Exchange. Retrieved from: <https://www.policyexchange.org.uk/wp-content/uploads/2016/09/not-working-apr-04.pdf>.
- Bonjour, D., Dorsett, R., Knight, G., Lissenburgh, S., Mukherjee, A., Payne, J. & White, M. (2001). *New Deal for Young People: national survey of participants: stage 2*. Retrieved from: <https://westminsterresearch.westminster.ac.uk/download/cecf05acdc672e54fdb15d4b099526298bcbecfa95ef393ed2c918a14704002/63038682/Bonjour%2C%20Dorsett%2C%20Knight%202001%20ESR67%20final.pdf>
- Boukema, R. (2017). Roadmapping A New Social Impact Bond In Rotterdam. Leiden University.
- Bromage, N. & Swersky, A. (2018). Outcomes-Based Commissioning in Mental Health & Employment. Presented at the Innovation in Outcomes-Based Commissioning, West Midlands Regional Conference.
- Brookings (n.d.). Retrieved from: <https://www.brookings.edu/series/impact-bonds/>
- Bryson, A., Knight, G. & White, M. (2000). *New Deal for Young People: national survey of participants: stage 1*. Retrieved from: [https://westminsterresearch.westminster.ac.uk/download/73855f2d6e10e2960f68e7d33a30186aa43ed80fba4ffa6407524c15dd4d5807/14984450/Bryson Knight White 2000.pdf](https://westminsterresearch.westminster.ac.uk/download/73855f2d6e10e2960f68e7d33a30186aa43ed80fba4ffa6407524c15dd4d5807/14984450/Bryson%20Knight%20White%202000.pdf).
- Bundesministerium für Soziales, Gesundheit, Pflege und Konsumentenschutz - Eingang Sozialministerium (2019). Social Impact Bond. Retrieved from: <https://www.sozialministerium.at/Themen/Soziales/Soziale-Themen/Soziale-Innovation/Social-Impact-Bond.html>
- Butler, D., Bloom, D. & Rudd, T. (2013). Using social impact bonds to spur innovation, knowledge building, and accountability, *Community Development Investment Review*, 9(1), 53-58.
- Cabinet Office (2016). Guidance Social Value Act: information and resources. Retrieved from: <https://www.gov.uk/government/publications/social-value-act-information-and-resources/social-value-act-information-and-resources>

- Carè R. & Ferraro, G. (2019). Funding Innovative Healthcare Programs Through Social Impact Bonds: Issues and Challenges, *China-USA Business Review*, 18(1), 1-15. doi: 10.17265/1537-1514/2019.01.001
- Carè, R. & De Lisa, R. (2019). Social Impact Bonds for a Sustainable Welfare State: The Role of Enabling Factors, *Sustainability*, 11(10), 2884. doi:10.3390/su11102884
- Carter B., Joshi A. & Remme, M. 2018. 'Scaling Up Inclusive Approaches for Marginalised and Vulnerable People' K4D Emerging Issues Report. Brighton, UK: Institute of Development Studies.
- CBO (2016). Mental Health and Employment Partnership (MHEP) Social Impact Bond.
- CBO (2017). Midpoint in-depth review report produced as part of the CBO Fund Evaluation. Retrieved from: https://www.tnlcommunityfund.org.uk/media/research-documents/social-investment/comissioning_better_outcomes_in_depth_review_190320_122442.pdf?mtime=20190320122441
- Centre for Social Justice (2020). Commissioning Excellence in Disability. Retrieved from: <https://www.centreforsocialjustice.org.uk/library/commissioning-excellence-in-disability>.
- Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions (2013). Towards Social Investment for Growth and Cohesion – including implementing the European Social Fund 2014-2020. European Commission. Retrieved from: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2013:0083:FIN:en:PDF>
- Dar, A. (2016). Work Programme: background and statistics. Briefing Paper Number 6340, House of Commons Library.
- Davelaar, M., Tan, S., & Spies, H. (2015). WP7 Case Study: Innovative Practice „Buzinezzclub Rotterdam“. Rotterdam. Verfügbar unter Retrieved from: http://www.citispyce.eu/sites/default/files/15%20Rotterdam_%20Buzinezzclub.pdf.
- DCLG (Department for Communities and Local Government) (2015). Financial Framework for the Expanded Troubled Families Programme.
- DEASP (2015). JobPath - Request For Tenders. Retrieved from: <https://www.gov.ie/en/collection/8fbd78-jobpath-request-for-tenders/>
- DEASP (2017). JobPath Performance Data Jobseekers Referred to JobPath July – September 2015. Retrieved from: <https://www.gov.ie/en/press-release/85249d-department-of-social-protection-publishes-first-performance-report-o/>
- DEASP (2018a). Jobseeker satisfaction with Satisfaction with JobPath service providers (Online research 2018). Shared with the research team by DEASP representative on 2020-05-28.
- DEASP (2018b). Jobseeker satisfaction with public offices research (Online research 2018). Shared with the research team by DEASP representative on 2020-05-28.
- DEASP (2019a). Working paper: Evaluation of JobPath outcomes for Q1 2016 participants. Retrieved from: <https://assets.gov.ie/36499/ffdce98cddc34addb05cf41a70aaf4e7.pdf>.
- DEASP (2019b). Satisfaction with JobPath service providers. Retrieved from: <https://www.gov.ie/pdf/?file=https://assets.gov.ie/75060/61bdf8e2-63d3-4c50-8011-b236921885c4.pdf#page=null>
- DEASP (2020). Surveys. DEASP. Retrieved from: <https://www.gov.ie/en/publication/96cc45-surveys-2015/>.
- Dekker, F. & Verhoeven, J. (2019). Vijf Jaar Social Impact Bonds: De Balans Opgemaakt. Start Foundation. Retrieved from: https://www.startfoundation.nl/FbContent.ashx/pub_1000/downloads/v190513095700/Vijf%20jaar%20Social%20Impact%20Bonds.pdf

- Deloitte (2015). 'Toelichting Rendementsberekening SIB Rotterdam' (Traject 1) 1-3. Rotterdam Buzinezzclub Social Impact Bond. Shared with the research team on 2020-05-29.
- Deloitte (2016). 'Toelichting Rendementsberekening SIB Rotterdam' (Traject 2) 1-3. Rotterdam Buzinezzclub Social Impact Bond. Shared with the research team on 2020-05-29.
- Deloitte (2018). 'Toelichting Rendementsberekening SIB Rotterdam' (Traject 3) 1-4. Rotterdam Buzinezzclub Social Impact Bond. Shared with the research team on 2020-05-29.
- Denoël, C. (2018). Relations entre entreprises d'économie sociale et pouvoirs publics. Le cas de l'accompagnement des demandeurs d'emploi en Région de Bruxelles-Capitale. *Master thesis*, UCLouvain. Retrieved from https://dial.uclouvain.be/memoire/ucl/fr/object/thesis%3A17421/datastream/PDF_01/view
- Department for Work and Pensions & Department of Health (2017). *Improving lives: the future of work, health and disability*. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/663399/improving-lives-the-future-of-work-health-and-disability.PDF
- Dermine, T. (2014). Establishing Social Impact Bonds in Continental Europe. *M-RCBG Associate Working Paper Series*, 26, Harvard Kennedy School, Mossavar-Rahmani Center for Business & Government. Retrieved from: https://www.hks.harvard.edu/sites/default/files/centers/mrcbg/files/dermine_final.pdf
- Dermine, T., Le Grelle, M. & Simonart, F. (2016). Social Impact Bonding. *Stanford Social Innovation Review*, 2016 (Fall), 61-62.
- DfE (2010). What works: re-engaging young people who are not in education, employment or training (NEET)? Summary of evidence from the activity agreement pilots and the entry to learning pilots. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/182022/DFE-RR065.pdf.
- DfE (2016). Youth Contract: delivery data: September 2012 to end of March 2016. Retrieved from: <https://www.gov.uk/government/publications/youth-contract-delivery-data>
- Dimitrijevska-Markoski, T. (2016). Social impact bonds: a new tool for governance of social programs-evidence from the UK, USA and Australia. *International Journal of Public Policy*, vol. 12(3/4/5/6), pages 261-275.
- Donmall, M. et al. (2017). Evaluation of the Drugs and Alcohol Recovery Payment by Results Pilot Programme Final Report. The University of Manchester.
- Donmall, M. et al. (2017). Evaluation of the Drugs and Alcohol Recovery Payment by Results Pilot Programme Intermediate report. University of Manchester.
- Dorsett, R. & Lucchino, P. (2016). The Work Programme: factors associated with differences in the relative effectiveness of prime providers. DWP ad hoc research report no. 26. Retrieved from: <https://www.gov.uk/government/publications/work-programme-factors-associated-with-differences-in-the-relative-effectiveness-of-prime-providers>
- Dudley, C., McEnhill, L. & Steadman, K. (2016). Is welfare to work, working well? Improving employment rates for people with disabilities and long-term conditions. The Work Foundation, Part of Lancaster University. Retrieved from: http://englishbulletin.adapt.it/wp-content/uploads/2016/06/405_Work-Programmes.pdf.
- DUO for a JOB (2015). *Rapport Annuel 2014*. Brussels. Retrieved from: <https://en.calameo.com/read/00509316677e7560696d6?authid=W3c7leRHP2iV>
- DUO for a JOB (2017). *Rapport Annuel 2016*. Retrieved from: <https://www.duoforajob.be/wp-content/uploads/2017/06/Rapport-Annuel-DUO-for-a-JOB-2016.pdf>

- DUO for a JOB (2018). *Rapport Annuel 2017*. Brussels. Retrieved from: <https://en.calameo.com/read/00509316698795b998034?authid=VjT1JtxIfY8u&page=51>
- DUO for a JOB (2019). *Rapport d'expérience. Le mentoring, la clé de l'insertion socio-professionnelle des jeunes issus de l'immigration?* Brussels. Retrieved from: <https://www.calameo.com/read/005093166cbb79844724e?authid=FYcbV2sgy2wo>
- DWP (2006). Departmental Report 2006. DWP. Retrieved from: <https://webarchive.nationalarchives.gov.uk/20070305141234/http://www.dwp.gov.uk/publications/dwp/2006/dr06/annexa/table10.asp>
- DWP (2008). Transforming Britain's labour market: ten years of the New Deal. Retrieved from: <https://webarchive.nationalarchives.gov.uk/20080205133910/http://www.dwp.gov.uk/welfare-reform/docs/PMNewDeal2-01-08.pdf>
- DWP (2011). The numbers of starts on New Deal for Young People and New Deal 25plus, and the number of times they have started each programme. DWP. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/223116/nd_repeaters_2011.pdf
- DWP (2013). Work Programme contract package areas and providers. Guidance. Retrieved from: <https://www.gov.uk/government/publications/work-programme-contract-package-area-and-prime-providers/work-programme-contract-package-area-and-prime-providers>
- DWP (2014). Commissioning Strategy. Retrieved from: <https://www.gov.uk/government/publications/dwp-commissioning-strategy-2014>
- DWP (2017a). Annual Report and Accounts 2016-2017. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/624319/dwp-annual-report-and-accounts-2016-2017.pdf.
- DWP (2017b). Work Programme National Statistics. Data up to December 2017. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/692566/work-programme-statistics-to-december-2017.pdf
- DWP (2018a). Work Programme National Statistics. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/692566/work-programme-statistics-to-december-2017.pdf
- DWP (2018b). Work Programme National Statistics background information note. Retrieved from: <https://www.gov.uk/government/publications/work-programme-official-statistics-background-information-note/work-programme-national-statistics-background-information-note>
- DWP (2018c). Work and Health Programme provider guidance, 1 February 2018. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/823529/whp-provider-guidance-chapter-13.pdf
- DWP (2018d). Work Programme statistics, Table 2_8. Retrieved from: <https://www.gov.uk/government/statistics/work-programme-statistical-summary-data-to-december-2017>.
- DWP (2020). Guidance DWP Employment Provision Commissioning Strategy 2020. Retrieved from: <https://www.gov.uk/government/publications/dwp-commissioning-strategy-2020/dwp-employment-provision-commissioning-strategy-2020>
- DWP (2020). Work Programme statistical summary: data to June 2020. Retrieved from: <https://www.gov.uk/government/publications/work-programme-statistical-summary-data-to-june-2020/work-programme-statistical-summary-data-to-june-2020>
- DWP. (2011). Jobcentre Plus Pathways to Work: Official Statistics. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/222922/jcppathways1011.pdf.

- DWP. (2011). Jobcentre Plus Pathways to Work: Official Statistics. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/222922/jcppathways1011.pdf.
- DWP. (2015). Youth Contract Official Statistics: April 2012 to May 2015, Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/455663/youth-contract-statistics-to-may-2015.pdf
- Dwyer, P.J. (2018). Punitive and ineffective: Benefit sanctions within social security. *Journal of social security law*. Retrieved from: <http://eprints.whiterose.ac.uk/138989/>
- ECORYS and ATQ. (2019). Commissioning Better Outcomes Fund Evaluation 2nd Update Report. Summary Report Targeted at Commissioners.
- Edmiston, D. & Nichols, A. (2017). Social Impact Bonds: The Role of Private Capital in Outcome-Based Commissioning. *Journal of Social Policy*. Retrieved from: <https://www.cambridge.org/core/journals/journal-of-social-policy/article/social-impact-bonds-the-role-of-private-capital-in-outcomebased-commissioning/83B92D884934604E21EE2A963EB4E11C>
- Erens, B., Roland, M. & Knapp, M. (2014). Payment by Results (PbR) Drug & Alcohol Recovery Pilot Programme: A note of advice to the Department of Health (DH) on the proposed evaluation. London: Policy Research Unit in Policy Innovation Research.
- Eurocities (2016). Rotterdam Buzinezzclub Social Impact Bond. Retrieved from: http://nws.eurocities.eu/MediaShell/media/Feb16_BuzinezzclubRotterdam.pdf.
- European Commission (2015). Better Regulation Toolbox [SWD (2015) 111]. Retrieved from: https://ec.europa.eu/info/sites/info/files/better-regulation-toolbox-2015_0.pdf
- European Commission (2018). Proposal for a Regulation of the European Parliament and of the Council establishing the InvestEU Programme. European Commission. Retrieved from: https://eur-lex.europa.eu/resource.html?uri=cellar:319a131d-6af6-11e8-9483-01aa75ed71a1.0002.03/DOC_1&format=PDF
- European Parliament, the Council and the Commission (2017). European Pillar of Social Rights. European Commission. Retrieved from: https://ec.europa.eu/commission/sites/beta-political/files/social-summit-european-pillar-social-rights-booklet_en.pdf
- European Social Fund (2014). Work Programme Provider Guidance: Chapter 12 – Performance Management. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/723312/work-programme-chapter-12.pdf.
- European Social Fund (2016). Work Programme Provider Guidance: Chapter 9 - Work Programme evidencing/validating payments. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/564883/work-programme-provider-guidance-chapter-9.pdf.
- European Social Fund (2017). Chapter 6 - Provider Assurance. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/824154/generic-guidance-chapter-6.pdf
- European Social Fund (2018). Work Programme Provider Guidance: Chapter 15: Standard Definitions for the Work Programme.
- EVPA (2017). SITRA – Impact Investing. Retrieved from: https://evpa.eu.com/uploads/publications/Policy_Nexus_FI_SITRA_2017.pdf
- Field, F. & White, P. (2007). *Welfare Isn't Working: The New Deal for Young People*. Reform, London.
- Finn, D. (2005). Welfare to work: New Labour's 'employment first' welfare state. *Benefits*, 13(2). Retrieved from: <https://core.ac.uk/download/pdf/29577128.pdf>

- Finn, D. (2013). Opening up the 'Black Box': What services are Work Programme providers delivering and how are they doing it. Flexwork Research Conference, Amsterdam, October 24-25 2013.
- Fixsen, A. (2013). 'Monitoring and Evaluating Scaling up of health system Interventions: Theory and Practice.' Draft manuscript. Retrieved from: <https://irh.org/resource-library/theory-and-practice-monitoring-evaluating-scale-up-of-health-system-innovations/>
- Fliegau, M. T., De Carlo, L., Herdrich, J., Hornauer, D., Jansen, T., Schetelig, K., Xu, G. (2015). Investition, Intervention, Impact. Der Soziale Wirkungskredit in Deutschland. Berlin.
- Foster, S., Metcalf, H., Purvis, A., Lanceley, L., Foster, R., Lane, P., Tufekci, L., Rolfe, H., Newton, B., Bertram, C. & Garlick, M. (2014a). Work Programme evaluation: Operation of the commissioning model, finance and programme delivery. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/425301/rr893-report.pdf.
- Foster, S., Metcalf, H., Purvis, A., Lanceley, L., Foster, R., Lane, P., Tufekci, L., Rolfe, H., Newton, B., Bertram, B. & Garlick, M. (2014b). Research Summary: Work Programme evaluation: Operation of the commissioning model, finance and programme delivery. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/425302/rr893-summary.pdf.
- Fox, C. & O'Leary, C. (2017). A structured review of Payment by Results and Social Impact Bonds evaluation evidence: working paper, Manchester: PERU. Retrieved from: [Review of PbR and SIB evaluation evidence FULL JUL 17.pdf](#)
- Fraser, A., Tan, S., Kruithof, K., Sim, M., Disley, E., Giacomantonio, C., Lagarde, M. & Mays, N. (2018). Evaluation of the Social Impact Bond Trailblazers in Health and Social Care: Final report. Policy Innovation Research Unit. Retrieved from: <https://piru.lshtm.ac.uk/assets/files/SIB%20Trailblazers%20Evaluation%20final%20report.pdf>
- Fraser, A., Tan, S., Lagarde, M. & Mays, N. (2018). Narratives of promise, narratives of caution: A review of the literature on Social Impact Bonds. *Social policy & Administration*. Retrieved from: https://onlinelibrary.wiley.com/doi/full/10.1111/spol.12260?casa_token=O7PsbmLmpUkAAAAA%3AvJIG5eIJ1Nc2qrdl9K0TaAjxizHpA4ThKOfupNds55Dxv_kwcEt3u51-s9GzITEVLG2cPG9ruVgKwg
- Frissen, R. (2014). Nieuwe Financieringsvormen Voor Publieke Waarde. NSOB Centrum voor Sociaal Ondernemerschap.
- Frissen, R. (2014). Nieuwe Financieringsvormen Voor Publieke Waarde. NSOB Centrum voor Sociaal Ondernemerschap.
- Frissen, R., Lunes, R. & Vermeer, F. (Society Impact Platform) en Revenboer, A. (Ernst&Young) (2013). Social Impact Bonds. Een verkenning naar de kansen van een innovatief financieringsarrangement.
- Frissen, R., Lunes, R. & Vermeer, F. (Society Impact Platform) en Revenboer, A. (Ernst&Young) (2013). Social Impact Bonds. Een verkenning naar de kansen van een innovatief financieringsarrangement.
- Gadenne V., Maglicic, M., Nolan D., Wright H. & Frerichs J. (2020). Individual Placement Support: A Social Impact Bond Model. The Behavioural Insights Team
- Galitopoulou, S. & Noya, A. (2016). Understanding social impact bonds. Organisation for Economic Co-operation and Development (OECD). Retrieved from: <http://www.oecd.org/cfe/leed/UnderstandingSIBsLux-WorkingPaper.pdf>
- Geijsen, T. (2019). Re-integratie naar werk. Ecorys.
- Geijsen, T. (2019). Re-integratie naar werk. Ecorys.

- George, A., Metcalf, H., Hunter, G., Bertram, C., Newton B., Skrine, O. & Turnbull, P. (2014). Evaluation of Day One mandate of prison leavers to the Work Programme. Retrieved from: <https://www.gov.uk/government/publications/work-programme-evaluation-operation-of-the-commissioning-model-finance-and-programme-delivery>
- Global Economy and Development at Brookings (2019). Brookings Impact Bonds Snapshot – 1 January 2019. Retrieved from: <https://www.brookings.edu/wp-content/uploads/2019/01/Impact-Bonds-Snapshot-January-2019.pdf>
- Global Economy and Development at Brookings (2020). Social and development impact bonds by the numbers. Retrieved from: <https://www.brookings.edu/research/social-and-development-impact-bonds-by-the-numbers/>
- Government Outcomes Lab (20 Social Finance 18). Presentation at the Innovation in Outcomes-Based Commissioning West Midlands Regional Conference.
- Government Outcomes Lab (2019a). Mental Health and Employment Partnership (Staffordshire, Haringey & Tower Hamlets). Retrieved from: <https://golab.bsg.ox.ac.uk/knowledge-bank/case-studies/mhep/>
- Government Outcomes Lab (2019b). Mental Health and Employment Social Impact Bond (Haringey & Barnet). Retrieved from: <https://golab.bsg.ox.ac.uk/knowledge-bank/indigo-data-and-visualisation/impact-bond-dataset-v2/INDIGO-POJ-0176/>
- Government Outcomes Lab (2019c). MHEP - Data Template. Retrieved from: <https://golab.bsg.ox.ac.uk/knowledge-bank/case-studies/mhep/>
- Government Outcomes Lab (2019d). *Projects database*. University of Oxford Blavatnik School of Government. Retrieved from: https://golab.bsg.ox.ac.uk/knowledge-bank/project-database/?query=&contract_type=Social+Impact+Bond.
- Griggs, J. & Evans, M. (2010). Sanctions within conditional benefit systems: A review of evidence. Joseph Rowntree Foundation.
- Gustafsson-Wright, E. & Gardiner, S. (2016). Using impact bonds to achieve early childhood development outcomes in low-and middle-income countries. *Brookings Institution*. Retrieved from: <https://www.brookings.edu/wp-content/uploads/2016/07/Impact-Bonds-for-ECDweb.pdf>
- Hadjivassiliou, K., Tassinari, A., Speckesser, S., Swift, S. & Bertram, C. (2015). Policy performance and evaluation: United Kingdom. *STYLE (Strategic Transitions for Youth Labour in Europe) Working Papers*. Retrieved from: <https://www.style-research.eu/wp-content/uploads/2015/03/STYLE-Working-Paper-WP3.3-Performance-UK.pdf>
- Haller, B. & Amesberger, H. (2018). Qualitative Evaluierung des Social Impact Bond PERSPEKTIVE:ARBEIT – Ökonomisches und soziales Empowerment von gewaltbetroffenen Frauen, Institut für Konfliktforschung, Wien.
- Haller, B. & Dawid, E. (2006). Kosten häuslicher Gewalt in Österreich, Institut für Konfliktforschung, Wien.
- Hayllar, O. & Wood, M. (2011). Provider-led Pathways to Work: the experiences of new and repeat customers in phase one areas, DWP Research Summary, HMSO, Norwich. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/197262/723summ.pdf
- HM Inspectorate of Probation (2013). Annual Report 2012-13. Independent inspection of adult & youth offending work. Retrieved from: <https://www.justiceinspectorates.gov.uk/probation/wp-content/uploads/sites/5/2014/03/hmi-probation-annual-report-2012-2013.pdf>
- HM Inspectorate of Probation (2016). Transforming Rehabilitation. Early Implementation 5. Retrieved from: <https://www.justiceinspectorates.gov.uk/hmiprobation/wp-content/uploads/sites/5/2016/05/Transforming-Rehabilitation-5.pdf>

- HM Inspectorate of Probation (2018). Quality & Impact inspection: The effectiveness of probation work by the London Community Rehabilitation Company. Retrieved from: <https://www.justiceinspectorates.gov.uk/hmiprobation/wp-content/uploads/sites/5/2018/03/London-CRC-QI.pdf>
- HM Inspectorate of Probation (2018). The quality of service user assessment (probation services). Research & Analysis Bulletin 2018/01. Retrieved from: <https://www.justiceinspectorates.gov.uk/hmiprobation/wp-content/uploads/sites/5/2018/09/2018-01-The-quality-of-service-user-assessment-probation-services-final.pdf>
- HM Inspectorate of Probation (2019). Report of the Chief Inspector of Probation. Retrieved from: <https://www.justiceinspectorates.gov.uk/hmiprobation/wp-content/uploads/sites/5/2019/03/HMI-Probation-Chief-Inspectors-Report.pdf>
- Holvast, L. (2018). Social Impact Bonds: The Holy grail by doing good and making profit? Realistic Evaluation of Social Impact Bonds in the Netherlands. Wageningen University.
- Holvast, L. (2018). Social Impact Bonds: the Holy grail by doing good and making profit? Realistic Evaluation of Social Impact Bonds in the Netherlands. Wageningen University.
- Home Office (2012). Drug and Alcohol Recovery Pilots: Lessons learnt from Co-Design. Retrieved from: <https://www.gov.uk/government/publications/payment-by-results-for-drug-and-alcohol-recovery-lessons-learnt-from-co-design>
- House of Commons Committee of Public Accounts (2010). Support to incapacity benefits claimants through Pathways to Work. First Report of Session 2010–11. Retrieved from: <https://publications.parliament.uk/pa/cm201011/cmselect/cmpublicacc/404/404.pdf>.
- House of Commons Committee of Public Accounts (2016). Transforming rehabilitation. Seventeenth Report of Session 2016–17. Retrieved from: <https://publications.parliament.uk/pa/cm201617/cmselect/cmpublicacc/484/484.pdf>
- House of Commons Committee of Public Accounts. (2019). Transforming rehabilitation: progress review. Ninety-Fourth Report of Session 2017–19. Retrieved from: <https://publications.parliament.uk/pa/cm201719/cmselect/cmpublicacc/1747/1747.pdf>
- House of Commons Justice Committee. (2013). Crime reduction policies: a co-ordinated approach? Interim report on the Government's Transforming Rehabilitation programme. Twelfth Report of Session 2013–14. Retrieved from: <https://publications.parliament.uk/pa/cm201719/cmselect/cmjust/482/48205.htm>
- House of Commons Justice Committee. (2018). Transforming Rehabilitation Ninth Report of Session 2017–19. Retrieved from: <https://publications.parliament.uk/pa/cm201719/cmselect/cmjust/482/482.pdf>
- House of Commons Work and Pensions Committee. (2010). Management and Administration of Contracted Employment Programmes. Fourth Report of Session 2009–10. Retrieved from: <https://publications.parliament.uk/pa/cm200910/cmselect/cmworpen/101/101.pdf>.
- House of Commons Work and Pensions Committee. (2011). Work Programme: providers and contracting arrangements. Volume I: Report, together with formal minutes, oral and written evidence. Retrieved from: <https://publications.parliament.uk/pa/cm201012/cmselect/cmworpen/718/718i.pdf>.
- House of Commons. (2012). Work and Pensions Committee Youth Unemployment and the Youth Contract Second Report of Session 2012–13, Retrieved from: <https://www.parliament.uk/business/committees/committees-a-z/commons-select/work-and-pensions-committee/inquiries/parliament-2010/youth-unemployment-and-the-governments-youth-contract/>
- Hudson, M., Phillips, J., Ray, K., Vegeris, S. & Davidson, R. (2010). The Influence of Outcome Based Contracting in Provider-led Pathways to Work, DWP Research Report 638, HMSO, Norwich.

<https://webarchive.nationalarchives.gov.uk/20130314010800/http://research.dwp.gov.uk/asd/asd5/rports2009-2010/rrep638.pdf>.

- Hutchinson, J., Gilbert, D., Papworth, R. & Boardman, J. (2018). Implementing Supported Employment. Lessons from the Making IPS Work Project. *International Journal of Environmental Research and Public Health*, 15(7): 1545.
- Indecon (2018). Indecon Review of Local Employment Services. Retrieved from: <https://www.dsfa.ie/en/pdf/IndeconReviewofLocalEmployment%20Services-Final%20Report.pdf>.
- inFocus. (2016). 7 Steps to effective impact measurement. inFocus Enterprises Ltd. Retrieved from: [https://www.undp.org/content/dam/ukraine/docs/DG/socinnov/7%20Steps%20to%20Effective%20Impact%20Measurement_v3_13.12.16%20\(1\).pdf](https://www.undp.org/content/dam/ukraine/docs/DG/socinnov/7%20Steps%20to%20Effective%20Impact%20Measurement_v3_13.12.16%20(1).pdf)
- Instiglio. (2017). A practitioner's Guide to Results-Based Financing. Retrieved from: https://www.payforsuccess.org/sites/default/files/resource-files/RBF_PractitionersGuidebook_Instiglio_18Oct2017.pdf
- Instiglio. (2019). *Impact Bonds Worldwide*. Retrieved from: <https://www.instiglio.org/en/sibs-worldwide/>.
- Instiglio. (n.d.). Retrieved from <https://www.instiglio.org/en/projects/> & <https://www.instiglio.org/en/sibs-worldwide/>
- IPSOS Mori (2019) Troubled Families Programme. TFC Staff Survey Report Final.
- IPSOS Mori (2019a). Troubled Families Programme. Qualitative Case Study Report, Phase 2, Wave 2.
- IPSOS Mori (2019b) – Troubled Families Programme, Family Survey.
- Jebabli, Jonas (2016). Social Impact Bonds: Collaboration Across Institutional Domains. Master Thesis. Vienna University of Economics and Business.
- Juvat (2015). Official Project Overview - The first Austrian social impact bond: Economic and social empowerment for women affected by violence.
- Juvat (2016). Abschluss erster deutscher Social Impact Bond: Das Modellprojekt in Augsburg. Fact Sheet 27.09.2016
- Kelly, E., McGuinness, S., Redmond, P., Savage, M. & Walsh, J. (2019). *An initial evaluation of the effectiveness of Intreo activation reforms*. Research Series, 81. Retrieved from: https://www.esri.ie/system/files/publications/RS81_0.pdf.
- Knight, G., Salis, S., Francavilla, F., Radu, D., Hevenstone, D., Mocca, E. & Tousley, B. (2013). Provider-led Pathways to Work: Net impacts on employment and benefits. Working Paper No 113. DWP. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/221223/WP113.pdf.
- Koekoek, R. & Zwietering, H. (2015). Kansen en uitdagingen in Nederland. ABN Amro.
- Koekoek, R. & Zwietering, H. (2015). Kansen en uitdagingen in Nederland. ABN Amro.
- Kohli, J., Besharov, D.J. & Costa, K. (2012). Social Impact Bonds 101: Defining an Innovative New Financing Tool for Social Programs. Retrieved from: <https://www.americanprogress.org/issues/general/news/2012/03/22/11238/social-impact-bonds-101/> for American Progress Center
- KOIS (n.d.). *DUO for a JOB: the first Social Impact Bond in continental Europe*. Retrieved from: <https://www.koisinvest.com/post/duo-for-a-job-the-first-social-impact-bond>
- Köppe, S. & O'Connell, P.J. (2017). 'Case Study on Intreo: The one-stop-shop for job seekers in Ireland'. *Case Studies on Innovation and Reform in the Irish Public Sector*. Department of

- Public Expenditure and Reform*, p. 4. Retrieved from: <https://researchrepository.ucd.ie/handle/10197/9318>
- Kump, N., Kavaš, D. & Črnigoj, M. (n.d.). Implementing Social Impact Bonds in Slovenia. Institute for Economic Research. Retrieved from: http://alpsib-project.eu/media/1059/discussion_paper_final_ier.pdf.
- Lane, P., Foster, R., Gardiner, L., Lanceley, L. & Purvis, A. (2013a). Research Summary: Work Programme Evaluation: Procurement, supply chains and implementation of the commissioning model. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/197712/832summ.pdf.
- Lane, P., Foster, R., Gardiner, L., Lanceley, L. & Purvis, A. (2013b). Work Programme Evaluation: Procurement, supply chains and implementation of the commissioning model. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/197710/rrep832.pdf.
- Lavelle, O. & Callaghan, N. (2018). Public Employment Services – Mapping Activation. IGEEES & Department of Public Expenditure and Reform. Retrieved from: <https://igees.gov.ie/wp-content/uploads/2018/07/12.-Public-Employment-Services.pdf>.
- Layard, R. (2001). Welfare to work and the New Deal (No. CEPOP15). Centre for Economic Performance, London School of Economics and Political Science. Retrieved from: [http://eprints.lse.ac.uk/47445/1/Welfare%20to%20work%20and%20the%20new%20deal%20\(lsero\).pdf](http://eprints.lse.ac.uk/47445/1/Welfare%20to%20work%20and%20the%20new%20deal%20(lsero).pdf)
- Lester, P. (2013). Review of Work Programme Minimum Performance Levels. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/209112/paul-lester-review.pdf.
- Lois, R.L. & Schlegel, K.G. (2019). Audit report Project M.O.R.E., Inc. Office of the Dutchess County Comptroller. Retrieved from: <https://www.dutchessny.gov/Audit-Report-Project-MORE-Inc.pdf>
- Lowe, S. (2015). JobPath: The proposed introduction of an employment programme in the Republic of Ireland. *The Public Sphere*.
- Lowe, S. (2015). JobPath: The proposed introduction of an employment programme in the Republic of Ireland. *The Public Sphere*.
- Maguire, S. & Newton, B. (2011). Activity agreement pilots–Trialling different approaches to re-engaging young people not in education, employment or training (NEET): Valuation of the 2009–10 extension. Research Report DFE-RR086). London: Department for Education. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/596141/Activity_agreement_pilots_2009_to_2010_extension_evaluation.pdf
- Maguire, S. (2015). Young people not in education, employment or training (NEET): Recent policy initiatives in England and their effects, *Research in Comparative and International Education*, 10(4), 525-536. Retrieved from: https://www.researchgate.net/profile/Arianna_Tassinari/publication/321726040_Plus_ca_change_Innovation_and_continuity_in_UK_youth_employment_policy_in_the_Great_Recession/links/5a2e863745851552ae7f34a4/Plus-ca-change-Innovation-and-continuity-in-UK-youth-employment-policy-in-the-Great-Recession.pdf
- Maguire, S., Newton, B., Fearn, H., Huddleston, P., Levesley, T., Miller, L., Oakley, J., Usher, T., Williams, C. & White, C. (2010). Activity Agreement Pilots – Evaluation Of The 2008-2009 Extension. Department for Children, Schools and Families. DCSF-RR201. Retrieved from: <https://webarchive.nationalarchives.gov.uk/20100419061329/http://www.dcsf.gov.uk/research/data/uploadfiles/DCSF-RR201.pdf>

- Maguire, S., Thompson, J., Hillage, J., Dewson, S., Miller, L., Johnson, C., Newton, B., Bates, P. & Page, R. (2009). Activity Agreement Pilots Process Evaluation, Retrieved from: <https://core.ac.uk/download/pdf/4160397.pdf>
- Marks, M.B. & Weaver, P.M. (2017). Are Social Impact Bonds a Viable Resource for Social Innovations? A Brief Discussion Paper. TRANSIT Working Paper. Retrieved from: http://www.transitsocialinnovation.eu/content/original/Book%20covers/Local%20PDFs/264TRANSIT_WorkingPaper%2011_Resourcing%20Strategies-pmw.pdf
- Mason, P., Fullwood, Y., Singh, K. & Battye, F. (2015). Payment by results: learning from the literature. A review prepared for the National Audit Office.
- Mason, P., Fullwood, Y., Singh, K. & Battye, F. (2015). Payment by results: learning from the literature. A review prepared for the National Audit Office. Retrieved from: <https://www.nao.org.uk/wp-content/uploads/2015/06/Payment-by-Results-Learning-from-the-Literature.pdf>
- Mason, T., Sutton, M., Whittaker, W., McSweeney, T., Millar, T., Donmall, M. & Pierce, M. (2015). The impact of paying treatment providers for outcomes: difference-in-differences analysis of the 'payment by results for drugs recovery' pilot, *Addiction*, 110(7), 1120-1128.
- McKay, K. (2013). Evaluating Social Impact Bonds as a new reentry financing mechanism: A case study on reentry programming in Maryland. Department of Legislative Services, Office of Policy Analysis.
- Meager, N., Newton, B., Sainsbury, R., Corden, A. & Irvine, A. (2014). Work Programme evaluation: the participant experience report. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/425081/rr892.pdf
- Millar, J. (2000). Keeping track of welfare reform. The New Deal programmes. Retrieved from: https://www.researchgate.net/publication/237112758_Keeping_track_of_welfare_reform_The_New_Deal_programmes
- Millner, R., Üstün, S., Moder, C. & Meyer, M. (2019). Social Impact Bond Pilotprojekt 'Perspektive:Arbeit'. Wien: WU Wien.
- Ministry of Housing Local Communities and Local Government (MHCLG) (2018). Supporting disadvantaged families. Annual report of the Troubled Families Programme 2017-18.
- Mirza-Davies, J. (2015). Youth Contract article. House of Commons Library. Retrieved from: <https://commonslibrary.parliament.uk/research-briefings/sn06387/>
- Misner, K. (2013). An interview with Kristin Misner of the NYC Mayor's Office about social impact bonds and the city's ABLE project. Think Justice Blog
- MoJ (2013). Rehabilitation Programme Market Engagement, Payment Mechanism Straw Man. Retrieved from: <https://www.justice.gov.uk/downloads/rehab-prog/payment-mechanism.pdf>
- MoJ (2013). Transforming Rehabilitation: A revolution in the way we manage offenders, Cm 8517. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/228580/8517.pdf
- MoJ (2013). Transforming Rehabilitation: A Strategy for Reform, Cm 8619, pp3 and 6. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/228744/8619.pdf
- MoJ (2013). Transforming Rehabilitation Programme – Payment mechanism. Market Feedback and Development Considerations. Retrieved from: <https://www.justice.gov.uk/downloads/rehab-prog/competition/payment-mechanism-feedback.pdf>

- MoJ (2014). Target Operating Model: Version 3. Rehabilitation Programme. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/387795/target-operating-model-3.pdf
- MoJ (2014). Transforming Rehabilitation Programme Payment Mechanism. Retrieved from: <http://data.parliament.uk/DepositedPapers/Files/DEP2014-0197/payment-mechanism-design-overview.pdf>
- MoJ (2019). Strengthening Probation, Building Confidence Consultation Response. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/802150/probation-reforms-impact-assessment.pdf
- MoJ (2020). Final and interim proven reoffending statistics for Community Rehabilitation Companies and the National Probation Service. January to March 2018 and 2017/18 annual cohort. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/882276/crc-nps-final-results-bulletin-apr20.pdf
- Mulheirn, I. (2013). Paying for results? Rethinking probation reform. Social Market Foundation. Retrieved from: <http://www.smf.co.uk/wp-content/uploads/2013/08/Publication-Briefing-Paying-for-Results-Rethinking-probation-reform.pdf>
- Municipality of Enschede (n.d.). Factsheet Social Impact Bond 'Werken in Duitsland' in het kort. Retrieved from: <https://www.enschede.nl/sites/default/files/factheet-sib-werken-in-duitsland--online-versie-gem-enschede.pdf>
- Murphy, M., Dulee-Kinsolving, A., Eustace, A. & Clarke, A. (2016). *IDSS Country Study Ireland*. Retrieved from: http://mural.maynoothuniversity.ie/12063/1/MM_IDSS%20Country%20Study.pdf.
- Murphy, M.P. & Hearne, R. (2019). Implementing marketisation: comparing Irish activation and social housing. *Irish Political Studies*, 34(3), 444–463. doi:10.1080/07907184.2019.1583215.
- Murphy, M.P. (2016). Low road or high road? The post-crisis trajectory of Irish activation. *Critical Social Policy*, 36(3), 1-21. doi:10.1177/0261018315626841.
- Nafilyan, V. & Speckesser, S. (2014). Estimating the long-term social benefits of a programme aiming to re-engage NEETs in education. Retrieved from: https://www.sheffield.ac.uk/polopoly_fs/1.679096!/file/2014_B1_2_Nafilyan.pdf
- Nafilyan, V. & Speckesser, S. (2014). The Youth Contract provision for 16-and 17 year-olds not in education, employment or training evaluation: Econometric estimates of programme impacts and net social benefits. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/358761/RR318B_-_The_youth_contract_provision_for_16-_to_17-year-olds_not_in_education__employment_or_training_impact_analysis.pdf.pdf
- NAO (2002). The New Deal for Young People. Retrieved from: <https://www.nao.org.uk/wp-content/uploads/2002/02/0102639.pdf>
- NAO (2010). Support to incapacity benefits claimants through Pathways to Work. Retrieved from: <https://www.nao.org.uk/report/support-to-incapacity-benefits-claimants-through-pathways-to-work/>
- NAO (2014). Probation: landscape review. Retrieved from: <https://www.nao.org.uk/wp-content/uploads/2014/03/Probation-landscape-review.pdf>
- NAO (2014). The Work Programme. Retrieved from: <https://www.nao.org.uk/wp-content/uploads/2014/07/The-work-programme.pdf>

- NAO (2015). Outcome-based payment schemes: government's use of payment by results. NAO. Retrieved from: <https://www.nao.org.uk/wp-content/uploads/2015/06/Outcome-based-payment-schemes-governments-use-of-payment-by-results.pdf>
- NAO (2016). Transforming Rehabilitation. Retrieved from: <https://www.nao.org.uk/wp-content/uploads/2016/04/Transforming-rehabilitation.pdf>
- NAO (2017). Investigation into changes to Community Rehabilitation Company contracts, HC 676. Retrieved from: <https://www.nao.org.uk/wp-content/uploads/2017/12/Investigation-into-changes-to-Community-Rehabilitation-Company-contracts.pdf>
- NAO (2019). Supporting Disabled People to Work. Retrieved from: <https://www.nao.org.uk/report/supporting-disabled-people-to-work/>.
- NAO (2019). Transforming Rehabilitation: Progress Review – Transcript. Retrieved from: <https://www.nao.org.uk/wp-content/uploads/2019/02/Transforming-Rehabilitation-Progress-review.pdf>
- NCVO/KnowHow (2018). Payment by Results. Retrieved from: <https://knowhow.ncvo.org.uk/funding/commissioning/commissioning-1/payment-by-results>
- Newton, B. (2009). Activity and learning agreement pilots: programme theory evaluation: working paper 5: activity agreements and small step progression. Retrieved from: <https://www.employment-studies.co.uk/resource/activity-and-learning-agreement-pilots-programme-theory-evaluation-working-paper-5>
- Newton, B., Meager, N., Bertram, C., Corden, A., George, A., Lalani, M., Metcalf, H., Rolfe, H., Sainsbury, R. & Weston, K. (2012). Work Programme Evaluation: Findings from the first phase of qualitative research on programme delivery. Research Report 821. Department for Work and Pensions. Retrieved from: <https://www.employment-studies.co.uk/resource/work-programme-evaluation-findings-first-phase-qualitative-research-programme-delivery>
- Newton, B., Speckesser, S., Nafilyan, V., Maguire, S., Devins, D., Bickerstaffe, T. & Marvell, R. (2014). *The Youth Contract for 16-17 year olds not in education, employment or training evaluation*. London: Department for Education. Retrieved from: <https://www.employment-studies.co.uk/resource/youth-contract-16-17-year-olds-not-education-employment-or-training-evaluation>
- Nice, K., Davidson, J. & Sainsbury, R. (2009). Provider-led Pathways: Experiences and views of early implementation, DWP Research Report 595, HMSO, Norwich.
- NICVA (n.d.). Payment by Results. Retrieved from: https://www.nicva.org/sites/default/files/d7content/attachments-articles/public_sector_reform_-_payment_by_results.pdf
- Noble, J. (2015). The transforming rehabilitation tier 1 tendering process: the voluntary sector perspective. NPC.
- Nonprofit Finance Fund (n.d.). Basics. Retrieved from: <https://www.payforsuccess.org/learn/basics/>
- NSW Office of Social Impact Investment (2020). *Evaluation of the resilient families service – final report*.
- Oakley, M. (2014). *Independent review of the operation of jobseeker's allowance sanctions validated by the Jobseekers Act 2013*. Department for Work and Pensions. Retrieved from: <http://www.haringeyadvice.org.uk/uploads/jsa-sanctions-independent-review-print.pdf>
- OECD (2016). Understanding Social Impact Bonds. OECD. Retrieved from: <http://www.oecd.org/cfe/leed/UnderstandingSIBsLux-WorkingPaper.pdf>.
- OECD Social Impact Investment Initiative. (n.d.) OECD. Retrieved from: <https://www.oecd.org/dac/financing-sustainable-development/development-finance-topics/social-impact-investment-initiative.htm>

- Office of the Comptroller and Auditor General (2018). Report on the Accounts of the Public Services 2017. Retrieved from: <https://www.audit.gov.ie/en/Find-Report/Publications/2018/2017-Annual-Report-Chapter-12-JobPath-Employment-Activation-Service.pdf>.
- Parsons, J., Weiss, C. & Wei, Q. (2016). Impact evaluation of the adolescent behavioral learning experience (ABLE) program. New York: Vera Institute of Justice. Retrieved from: <https://www.vera.org/publications/rikersadolescent-behavioral-learning-experience-evaluation>.
- Pay for Success (2019). *Projects*. Nonprofit Finance Fund. Retrieved from: <https://www.payforsuccess.org/projects/>.
- Pay for Success (2019). *Projects*. Retrieved from: <https://www.payforsuccess.org/projects>
- Phillips, A. & Olson, J. (2013). Rikers Island: the first social impact bond in the United States, *Community Development Investment Review*, 097-101.
- Platform 31 (2019). Do's and don'ts bij een Social Impact Bond. Retrieved from: <https://www.platform31.nl/nieuws/do-s-and-don-ts-bij-een-social-impact-bond>
- Policy, Programs and Strategy Statutory Child Protection (2014). *Intensive Family Preservation Program Guidelines*.
- PricewaterhouseCoopers International Limited (2004). New Deal 25+ Evaluation Report No. 9. PWC. Retrieved from: <https://www.communities-ni.gov.uk/sites/default/files/publications/del/new-deal-25plus-evaluation-report-no9.pdf>.
- Publishing Service of the UK Government. (n.d.). Payment by Results. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/249928/Payment-by-Results.pdf
- Rees, J., Whitworth, A. & Carter, E. (2014). 'Support for All in the UK Work Programme? Differential Payments, Same Old Problem', *Social Policy & Administration* Vol. 48, No. 2, 221-239.
- Regiers, P. (2014). Overview: Strategies for Causal Attribution. UNICEF. Retrieved from: https://www.unicef-irc.org/publications/pdf/brief_6_overview_strategies_causal_attribution_eng.pdf
- Riley, R. & Young, G. (2000). New Deal for Young People: Implications for Employment and the Public Finances. National Institute of Economic and Social Research. Retrieved from: https://www.niesr.ac.uk/sites/default/files/publications/140306_123853.pdf.
- Roy, M.J., McHugh, N. & Sinclair, S. (2018). A critical reflection on social impact bonds. *Stanford Social Innovation Review*. Retrieved from: https://ssir.org/articles/entry/a_critical_reflection_on_social_impact_bonds
- Rudd, T., Nicoletti, E., Misner, K. & Bonsu, J. (2013). Financing Promising Evidence-Based Programs: Early Lessons from the New York City Social Impact Bond. MDRC.
- Scheck, B. (2017). Begleitevaluation des ersten deutschen Social Impact Bonds. JuMP Jugendliche mit Perspektive.
- Scheltinga, P. & Nijhof, B. (2019). Concept Eindrapportage 'Monitoring Social Impact Bond Werken in Duitsland'. Kennispunt Twente.
- Scheuerle, T. & Nieveler, A. (2017). Implementing Social Impact Bonds in Germany – Challenges for Pay-for-Success Models in the German Welfare System. German discussion paper. Interreg Alpine Space AlpSib. European Regional Development Fund.
- Scholten, P. (2015). Social impact bonds. Van fenomeen naar praktijk. Start Foundation
- Scholten, P. (2015). Social impact bonds. Van fenomeen naar praktijk. Start Foundation
- Sida. (2015). A methodological introduction. Results based financing approaches (RBFA) – what are they? Sida. Retrieved from: <https://www.sida.se/contentassets/1b13c3b7a75947a2a4487e2b0f61267c/18235.pdf>

- SITRA. (2019). Social outcomes contracting and public procurement in various EU Member States. Retrieved from: <https://media.sitra.fi/2019/06/07135808/sitrasocsummary-findings-social-outcomes-contracting.pdf>
- Smans, V. (2015). *Social impact bonds (SIB) et mesure d'impact social. Etude de cas du SIB belge en collaboration avec Actiris et Duo for a Job* (Unpublished master's thesis), ICHEC, Brussels, Belgium. Retrieved from: https://issuu.com/fgffg/docs/mtafinance2016-smans_tfe
- Social Finance and Health and Employment Partnerships (2015). Launching a Social Impact Bond in Mental Health and Employment. Retrieved from: https://www.base-uk.org/sites/default/files/document-archive/9991-Conference%20workshops/c4_-_social_impact_bonds.pdf
- Social Finance. (2019). *Impact Bond Global Database*. Retrieved from: <https://sibdatabase.socialfinance.org.uk/>.
- Social Finance. (2019). Mental Health And Employment (MHEP). Retrieved from <https://www.socialfinance.org.uk/projects/mental-health-and-employment-mhep>.
- Social Finance. (n.d.). Retrieved from <https://www.socialfinance.org.uk/who-we-are/global-network>
- Stafford, B. (2007). New Deal for Disabled People: Third synthesis report–key findings from the evaluation. DWP. Vol. 430, Retrieved from: <https://www.york.ac.uk/inst/spru/pubs/pdf/rrep430.pdf>.
- Sumo, R., van der Valk, W., Duysters, G. & van Weele, A. (2016). Using performance-based contracts to foster innovation in outsourced service delivery. *Industrial Marketing Management*, 59, 12–24. doi:10.1016/j.indmarman.2016.05.029
- Tan, S., Fraser, A., McHugh, N. & Warner, M. (2019). Widening perspectives on social impact bonds. Taylor & Francis. Retrieved from: <https://www.tandfonline.com/doi/full/10.1080/17487870.2019.1568249>
- Tanner, E., Purdon, S., D'Souza, J. & Finch, S. (2009). Activity Agreement Pilots Quantitative Evaluation. Retrieved from: <https://dera.ioe.ac.uk/11342/1/DCSF-RR096.pdf>
- Tassinari, A., Hadjivassiliou, K. & Swift, S. (2016). Plus ça change...? Innovation and continuity in UK youth employment policy during the Great Recession. *Social Policies*, 3(2), 225-248.
- Tennant, R., Kotecha, M. & Rahim, N. (2010). Provider-led Pathways: experiences and views of implementation in phase 2 districts, DWP Research Report 643, HMSO, Norwich.
- TrackTR. (2015). Early Doors. The voluntary sector's role in Transforming Rehabilitation. Retrieved from: https://blogs.ncvo.org.uk/wp-content/uploads/2015/08/track-tr_early-doors.pdf
- TrackTR. (2016). Change & Challenges. The voluntary sector's role in Transforming Rehabilitation. Retrieved from: <https://blogs.ncvo.org.uk/wp-content/uploads/2016/05/Change-and-Challenge-Final.pdf>
- TrackTR. (2018). Under represented, under pressure under resourced. The voluntary sector in Transforming Rehabilitation. Retrieved from: https://www.clinks.org/sites/default/files/2018-10/clinks_track-tr_under_final-web.pdf
- UK Parliament (2017). Work Programme: Written question – 60438. Retrieved from: <https://www.parliament.uk/written-questions-answers-statements/written-question/commons/2017-01-17/60438>
- UNDP. (n.d.). Social and development impact bonds (results-based financing). Retrieved from: <https://www.sdfinance.undp.org/content/sdfinance/en/home/solutions/social-development-impact-bonds.html>
- USAID. (n.d.). Performance-based financing. Retrieved from: <http://www.fppinancingroadmap.org/learning/specific-topics/performance-based-financing>

- User Voice (2016). Transforming Rehabilitation: The operational model from the service user's perspective. Retrieved from: <https://www.nao.org.uk/wp-content/uploads/2016/04/User-Voice-report-Transforming-rehabilitation-The-operational-model-from-the-service-users-perspective.pdf>
- Van Es, B. (2015). Social impact bonds. Een nieuw instrument om jeugdwerkloosheid te bestrijden. Hogeschool Rotterdam.
- Van Es, B.C., Houben, O.J. & Nijeholt, M.P.L. (2016). Social Impact Bonds: an Innovative Way to Combat Youth Unemployment. *Journal of Positive Management*. Retrieved from: <https://apcz.umk.pl/czasopisma/index.php/JPM/article/view/JPM.2016.003>
- Van Reenen, J. (2004). Active labor market policies and the British New Deal for the young unemployed in context. In *Seeking a premier economy: The economic effects of British economic reforms, 1980-2000*. University of Chicago Press.
- Vennema, B. & Koekoek, R. (2013). Social Impact Bonds. Introductie van een nieuw financieringsmodel in Nederland. ABN Amro.
- Vennema, B. & Koekoek, R. (2019). 5 Jaar Social Impact Bonds In Nederland. Social Finance NL.
- Wiggan, J (2015). What variety of employment service quasi-market? Ireland's job path as a private power market. In: Irving, Z., Fenger, M. & Hudson, J. (eds), *Social Policy Review*. vol. 27, Bristol: Policy Press.
- World Health Organization (2009). Practical guidance for scaling up health service innovations. World Health Organization. Retrieved from: https://apps.who.int/iris/bitstream/handle/10665/44180/9789241598521_eng.pdf
- Zaidi, A. (2009). Welfare-to-Work Programmes in the UK and Lessons for other Countries. Policy brief. European Centre for social welfare policy and research, Vienna October.
- Zwetsloot, J. & van der Werff, S. (2018). Toelichting en handleiding SIB Gemeente Rotterdam. SEO economisch onderzoek.

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at:

https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696 or
- by email via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications at: <https://op.europa.eu/en/publications>. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: <http://eur-lex.europa.eu>

Open data from the EU

The EU Open Data Portal (<http://data.europa.eu/euodp/en>) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.



Publications Office
of the European Union