



Department
for Work &
Pensions



Government
Social Research

Effective contracting of employment and health services

Evidence Review

October 2024

DWP research report no. 93

A report of research carried out by RAND Europe on behalf of the Department for Work and Pensions.

Crown copyright 2024.

You may re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. To view this licence, visit

<http://www.nationalarchives.gov.uk/doc/open-government-licence/> or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email psi@nationalarchives.gov.uk.

This document/publication is also available on our website at:

<https://www.gov.uk/government/organisations/department-for-work-pensions/about/research#research-and-analysis-publications>

If you would like to know more about DWP research, email

socialresearch@dw.gov.uk

First published October 2024.

ISBN 978-1-78659-715-1

Views expressed in this report are not necessarily those of the Department for Work and Pensions or any other government department.

Acknowledgements

We would like to thank numerous colleagues from the Department for Work and Pensions, in particular Simon Marlow and Michael Conn, for their continuous support and guidance during this study.

We are also very grateful to the many researchers and service providers who generously offered their time and contributed to the interviews and workshop consultations.

Our thanks go also to colleagues at RAND Europe who provided valuable feedback on the design, implementation and results of this research.

The authors

Dr Julia Doyle

Dr Madeline Nightingale

Jessica Dawney

Michael Whitmore

Dr Chris van Stolk

Joanna Hofman

Summary

Context

Robust, quantitative evidence to determine the costs and benefits of contracting out employment services and health assessments and comparing these models to no provision or to traditional, in-house delivery by Public Employment Services (PES), is relatively scarce. In addition, existing studies do not allow us to fully understand why contracted services performed well (or not so well), if and how this differs according to context and contract design. As a result of this, evidence is also limited on how the process of contracting out could be improved to achieve better outcomes for individual clients, while keeping the costs of service reasonable for the public purse.

This study, commissioned by the Department for Work and Pensions (DWP) and conducted by RAND Europe, synthesises findings from the existing literature and qualitative interviews with commissioners, providers and other experts, to address the question:

What are the advantages and disadvantages of contracting out public employment services and health assessments and what is the optimum model for doing so?

This main question agreed with DWP captures a number of specific issues explored in this report around contracting employment services and health assessments. These two types of services represent different contracting models in the UK: (i) *payment by results (PbR)* contracts feature prominently in employment services (these were temporarily replaced with Cost Plus contracts introduced in response to the COVID-19 pandemic); (ii) *payment for service* contracts are commonplace in health assessments. The reviewed evidence focuses primarily on contracts utilising PbR mechanisms, which are common in the commissioning of employment services in the UK and internationally. There were no international examples of commissioning health assessments in the reviewed literature. This is reflected in the research findings but where possible and relevant, other contract models are discussed.

Findings

Contracting out allows to better adjust the volume of service provision to labour market conditions, and – at least in theory – to better tailor services to the clients' needs by using specialised service providers (empirical evidence is discussed below). We found that **PbR contracts in employment services may bring benefits in terms of programme efficiency and value for money**. However, there is **mixed evidence on the impact of PbR contracts on programme effectiveness**: some studies find that PbR contracts for employment programmes are associated with cost savings, whereas this appears not to be the case for other programmes. **The main costs associated with PbR contracts are linked with undesirable behaviour**

from providers such as creaming and parking (i.e., focusing their efforts on job seekers who are deemed closer to the labour market). Contracting out may also weaken the accountability of public services.

We examined risks associated with PbR contracts and found that, **for commissioners, these risks include overpaying for services, underpaying and market failure**. For providers, the risks are high upfront costs, underperformance and reputational concerns. In addition, contracting out may incur a risk of variable service provision for users/beneficiaries.

The process of contracting out is determined by specifications issued by the commissioners. This documentation formulates expectations of the purchaser towards the service provider(s), and it defines, and governs responsibilities and rights among the parties. Such documentation can attempt to regulate every aspect of services to be provided or leave some (or most) aspects in the provider's discretion (black box). Compared to detailed specifications, **black box provision is intended to allow providers greater flexibility and facilitate innovation**, but there is little empirical evidence this actually happens (see more below). Flexibility for providers results in greater service variability and may increase the risk of creaming and parking. Minimum service requirements are designed to assure basic service standards, but views are divided on the optimum level of specification.

In the UK, DWP commissions employment contracts to one or more (prime) contractors across geographical regions known as Contract Package Areas (CPAs). These prime contractors manage a network of subcontractors delivering services in the area. We found that **the prime model of contracting builds long-term relationships with large contractors** who can take on the risk of a PbR contract. However, there is **a risk of subcontractors not being able to benefit fully from contracts**. The number and size of CPAs differs across programmes, but **large CPAs may mean that providers lack local expertise**.

Criteria used to select providers for contracted out services often combined quality and price. In the examples identified in this evidence review, the weight of price relative to quality ranged from 20% to 40%. The quality aspects usually aim to capture relevant experience and expertise. The approach using the Price per Quality Point (PQP) metric is relatively new and it has not been discussed in the reviewed literature.

There is little evidence on how different types of contracts influence market competition in the short- and long-term. However, there is some evidence on the effect of contracting out on market competition: this evidence points to **some providers being 'pushed out' or forced to adapt to secure contracts and deliver services**. The dominance of prime contractors in the market has been well established.

Trade-offs between optimising short-term and long-term outcomes has been a process of 'trial and error' in the UK and internationally. One of the **key challenges is finding an optimal balance between upfront fees and outcome payments** to create incentives for providers and savings for commissioners. In the examples

identified in this evidence review, the proportion of payments linked to outcomes varied markedly by country and by programme ranging from 30% (e.g. CareerCompass and CareerAdvance in the US) to over 90% (e.g. the WP in the UK).

While **contract management is critical for ensuring the quality of service provision**, there is no agreed definition of service ‘quality’ or how to measure it. Performance measurement and monitoring aim to encourage innovation and limit gaming.

Incentives to encourage collaboration or information sharing among providers are limited. There is little evidence on sharing contract resources amongst providers and other stakeholders. Expertise and specialist knowledge of providers is key and the need for professional development of staff was noted.

Innovation in employment services may mean a variety of things including working with a new client group, introducing new forms of support or training or making organisational processes more efficient. Competition intends to encourage innovation and this needs to be considered early on in the commissioning process. However, **evidence for innovation taking place in PbR contracts is limited** and the reviewed literature points to opportunities for innovation often not being taken up.

Conclusion

Optimum contract design will depend on the aims and objectives of the programme, meaning it is not possible to draw overarching conclusions about the ‘best’ approach. There may be tensions and trade-offs between objectives associated with different elements of contract design. Multiple external factors may also affect delivery (e.g., economic performance throughout the contract duration may make it easier or more difficult for providers to achieve targets set up at the start). While there is no single recipe for a flawless PbR contract model, the evidence base is growing. The table below outlines main features of PbR contracts, associated levers through which their features could be optimised and how these levers may lead to adverse effects elsewhere. This framework may help commissioners better understand these dependencies and allow them to make better informed choices suited for their needs.

Optimising PbR contracts to achieve the desired outcomes

Benefits, costs and risks in PbR contracting	Assumption	Lever	Associated negative effects
Benefit: improved effectiveness and cost-efficiency	Focus on outcomes drives better results ('what gets measured, gets done')	Increased weighting for outcome payments	Gaming behaviour by providers Disadvantaging specialist providers / subcontractors by primes
Cost: high transaction costs	Reducing costs of commissioning through PbR increases VFM to commissioners	Less detailed contract specifications (black box) Fewer minimum service requirements	Variability in service provision Deteriorating service quality

		Longer, fewer and bigger contracts	Reducing market competition
Risk: gaming strategies	Reducing incentives to gaming minimises risks and costs to commissioners	Differentiate outcome payments by client group More minimum service requirements More detailed contract specifications (grey box)	Limiting flexibility of providers and potential for innovation
Risk: primes' monopoly	Market competition stimulates greater effectiveness and cost savings for commissioners	Smaller CPAs or two-stage competitions within existing CPAs Shorter contracts (Financial) incentives or contractual requirements for primes to work with smaller providers	Increased costs for commissioners
Risk: quality of services	Safeguards must be in place to ensure service users receive high quality provision	More minimum service requirements Strengthening monitoring arrangements or introducing a licencing / accreditation system Strengthening enforcement or breakout clauses	Increased costs for commissioners Reduced incentives (increased costs) for providers

More evidence is needed on health assessments (both internationally and in the UK) to better understand the state of play in contracting out these services and possible ways forward.

Stronger evidence on the actual costs incurred and benefits achieved of different models of provision (in-house, contracted out (including PbR, Cost Plus and payment for service contracts), partnership, grant funding arrangements) is needed to be able to compare these. On that basis it would be possible to draw more informed conclusion on which works better for whom, under what conditions, and why.

Contents

Acknowledgements	3
The authors	4
Summary	5
Contents	9
Abbreviations.....	11
Glossary	12
Boxes.	14
1. Introduction.....	15
1.1. Contracting out public services.....	15
1.2. Research objectives and questions	16
1.3. Methods.....	17
1.4. Structure of this report	18
2. Contract design	20
2.1. Costs and benefits associated with contracting out.....	21
2.2. Risks associated with contracting out.....	41
2.3. The level of specification in contracted services.....	51
2.4. Cost Plus model	59
2.5. Prime model	61
2.6. Contract Package Areas may be too large	66
2.7. Length of contracts ideally falls between five to seven years	69
3. Market competition	72
3.1. Selection criteria often combined quality and price.....	72
3.2. There is little evidence of how different types of contracts impact market competition in the short- & long-term.....	77
3.3. Trade-offs between optimising short-term and long-term outcomes has been a process of ‘trial and error’	80
4. Live running	84
4.1. Measuring service quality	84
4.2. Incentives to encourage collaboration or information sharing among providers are limited	91
4.3. There is little evidence on sharing contract resources amongst providers and other stakeholders.....	93
4.4. Expertise and specialist knowledge of providers’ staff is key and need for professional development of staff was noted	95
4.5. Encouraging innovation through PbR	98
5. Conclusions	103
5.1. The aims of this research	103
5.2. Summary of findings and key conclusions.....	103
5.3. Remaining evidence gaps and future research	108
References	109
Annex 1: Research framework	115
Annex 2: Search protocol	119
Databases	119
Search terms.....	119
Inclusion and exclusion criteria.....	120

- Snowball search 120
- Targeted hand search 120
- Search results 120
- Data extraction template 121
- Annex 3: Interview topic guide 123
 - Introduction/background 123
 - Contract design 123
 - Market competition 124
 - Live running 124
 - Conclusion, thanks and wrap up 124
- Annex 4: Synthesis workshop 125
 - List of workshop participants 125
 - Agenda 125
 - Breakout groups topics and facilitation questions 125
- Annex 5: Methodological challenges associated with estimating the costs and benefits associated with PbR contracts 127

Abbreviations

CAEHRS	Commercial Agreement for Employment and Health Related Services (framework)
CPA	Contract Package Areas
DWP	Department for Work and Pensions
ESA	Employment and Support Allowance
EU	European Union
FND	Flexible New Deal
ITT	Intention to Treat
IPES	Intensive Personalised Employment Support
JETS	Job Entry Targeted Support
JSA	Job Seeker's Allowance
KPI	Key Performance Indicator
NDDP	New Deal for Disabled People
NDYP	New Deal for Young People
NHS	National Health Service
NIR	Non-Intervention Rate
OECD	Organisation for Economic Co-operation and Development
PA	Personal Advisor
PbR	Payment by Results
PES	Public Employment Services
QSR	Quick Scoping Reviews
RAG	Red/Amber/Green rating system
RCT	Randomised Controlled Trial
SMEs	Small- and medium-sized enterprises
UK	United Kingdom
VFM	Value for Money
WCA	Work Capacity Assessments
WHP	Work and Health Programme
WP	Work Programme

Glossary

Black box	Service provision where providers have the freedom to decide what to do and they receive financial rewards for outcomes achieved (with the aim of encouraging innovation).
Contract Package Areas	Geographical areas across England and Wales with a small number of prime contractors managing a network of subcontractors delivering services in the area.
Cost Plus contract	An agreement to reimburse a company for expenses incurred in addition to a specific amount of profit, usually stated as a percentage of the contract's full price.
Creaming or cherry-picking	Undesirable behaviour from providers to focus on clients with good employment prospects (and to neglect those with complex needs).
Gaming	A strategy used by providers to exploit weaknesses in contracts / specifications to maximise economic gains (at cost to service quality).
Grey box	Specifying some aspects of service provision but allowing providers to tailor other elements, in contrast to a 'black box' in which some (or most) aspects are left in the provider's discretion.
Live running	Execution of the contracts for delivering employment services and health assessments.
Outcome payment	Financial remuneration awarded by the commissioner to providers for each service user reaching the pre-agreed criteria for claiming an 'outcome'.
Price cap	A limit on the price a provider can charge.
Parking	Undesirable behaviour from providers to withhold or reduce support for unemployed clients who are difficult to place.
Payment by Results (PbR)	A model of financing public services where payments depend (largely or wholly) on the results achieved by service providers, thus providing a financial incentive for providers to deliver effective (in the sense of achieving outcomes) services throughout the term of the contract.
Prime/subcontractors	Model of commissioning employment contracts in the UK to a small number of (prime) contractors per Contract Package Area (CPA); prime contractors manage a network of subcontractors delivering services in the area.
Quasi-market	A term commonly used in the literature to describe the market created by contracting out public services. In quasi-markets the recipient of the services and the purchaser are separate entities and providers do not necessarily compete for profit.
Restart Scheme	The Restart Scheme was launched in 2021 in England and Wales in 12 CPAs. It gives benefit claimants who have been out of work for at least 9 months enhanced support to find jobs in their local area. ¹
Specification	Documentation stating how the service should be provided:

¹ See more on the Restart Scheme [here](#).

	<ul style="list-style-type: none"> • Little specification means ‘black box’ provision where providers have the freedom to decide what to do (this should encourage innovation). • Detailed specification limits the freedom of the providers (and limits the potential for innovation). • Specifying some aspects of service provision but allowing providers to tailor other elements is referred to by DWP as ‘grey box’.
Upfront or service fee	Payment to providers for the delivery of some basic service to each user, usually early in the contract and regardless of the 'outcome'.
Work and Health Programme	The Work and Health Programme (WHP) replaced the Work Programme (WP) and was rolled out in England and Wales between November 2017 and April 2018. It aims to help disabled people, the long-term unemployed, and some priority groups (so called early access groups) to enter into and stay in work. It is delivered in 6 CPAs and devolved to some Local Government Partners and Devolved Deal Areas. ²
Work Programme	The Work Programme (WP) was a PbR welfare-to-work programme in Great Britain in the period 2011-2017. It was delivered in 18 CPAs. ³
Workfare programmes	Social assistance programmes that require recipients to participate in work or other activities (e.g., training).

² See more on the WHP [here](#).

³ See more on the WP [here](#).

Boxes

- Box 1. Body of evidence on contracting out employment services from the UK..... 24*
- Box 2. Health assessments: suitability of the PbR model..... 41*
- Box 3. Health assessments: risks in contracting out..... 50*
- Box 4. Health assessments: level of specifications 58*
- Box 5. Health assessments: Cost Plus model 60*
- Box 6. Health assessments: CPAs 69*
- Box 7. Health assessments: length of contracts 71*
- Box 8. Health assessments: service quality..... 86*
- Box 9. Health assessments: collaboration and sharing resources..... 95*
- Box 10. Health assessments: expertise and specialist knowledge 97*
- Box 11. Health assessments: innovation 102*

1. Introduction

This report presents findings from an evidence review conducted by RAND Europe on behalf of the Department of Work and Pensions (DWP) about effective contracting of employment services and health assessments.

1.1. Contracting out public services

DWP is responsible for welfare, pensions and child maintenance policy. It is the biggest public service department in the United Kingdom (UK). It administers the State Pension and a range of working age, disability and ill health benefits to around 20 million claimants and customers.⁴ DWP provides publicly financed health assessments to assess eligibility for health-related benefits (which are the statutory services), and employment support (which are the discretionary services). Some employment services in the UK are contracted out by DWP to providers, including private and not-for-profit organisations, with an intention to manage by objectives,⁵ and improve effectiveness (extent to which results are achieved)⁶ and efficiency (extent to which results are delivered in an economic and timely way⁷).^{8,9}

The evidence behind this model of service provision is growing, especially for commissioning employment services,¹⁰ but it is showing mixed results (discussed in section 2.1.2).^{11,12,13} There is still little research and evidence-based guidance for commissioners on how to contract out employment and health contracts¹⁴ effectively and how to manage such markets over time.

⁴ DWP (2002). As of 12/05/2022: <https://www.gov.uk/government/organisations/department-for-work-pensions/about>

⁵ Management by objectives is a management system that directs activities toward the effective and efficient achievement of organisational objectives through participation in decision making, goal setting, and objective feedback. See: Rodgers, R., & Hunter, J. E. (1992). A foundation of good management practice in government: Management by objectives. *Public Administration Review*, 27-39.

⁶ OECD (2021). Applying Evaluation Criteria Thoughtfully.

⁷ OECD (2021).

⁸ Finn, D. (2011). Sub-contracting in Public Employment Services: Review of research findings and literature on recent trends and business models. The European Commission Mutual Learning Programme for Public Employment Services. DG Employment, Social Affairs and Inclusion

⁹ OECD (1997). Contracting out government services: Best Practice Guidelines and Case Studies. Occasional papers No. 20.

¹⁰ For contracting out health services (but not specifically health assessments) see: Cole, A., Cubi-Molla, P., Elliott, R., Feast, A., Hocking, L., Lorgelly, P., Payne, K., Peek, N., Sim, D., Sussex, J., Zhang, K and Steuten, L. (2021). Making Outcome-Based Payment a Reality in the NHS. Phase 2: Practical Considerations; van Stolk, C., Bjornsson, G., and Goshev, S. (2010). Provider Incentives in Social Protection and Health: A selection of case studies from OECD countries. RAND Corporation

¹¹ Gash, T., Pancharia, N., Sims, S., and Hotson, L. (2012). Making public service markets work: Professionalising government's approach to commissioning and market stewardship. Institute for Government

¹² Langenbucher, K. & Vodopivec, M. (2022). Paying for results: Contracting out employment services through outcome based payment schemes in OECD countries. OECD Social, Employment and Migration Working Papers No. 267

¹³ Stephan, G. (2016). Public or private job placement services: Are private ones more effective?. *IZA World of Labor* 2016: 285 doi: 10.15185/izawol.285

¹⁴ See, for example: Gash, T. and Pancharia, N. (2012). When to contract: Which service features affect the ease of government contracting?

1.2. Research objectives and questions

There is little quantitative evidence to determine, which model of providing employment services and health assessments is better: contracting out or delivering these in-house.¹⁵ Ideally, one would be able to make a direct – quantitative – comparison of effectiveness between services contracted out and those delivered by Public Employment Services (PES) for the same group of people. However, such comparisons are difficult to conduct and thus rare. In addition, such studies often do not allow to fully understand why contracted services performed well (or not so well), when compared to the business as usual, and how the process could be improved.

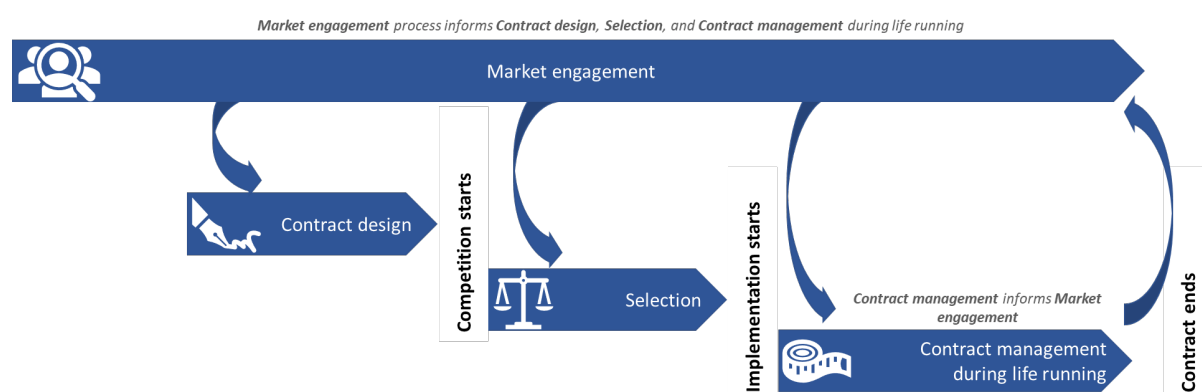
This study, commissioned by DWP and conducted independently by RAND Europe, looks to collate existing evidence from literature on the effectiveness and efficiency of the contracting out model. Furthermore, it employs qualitative interviews to provide additional insights on the matter to address the following main research question:

What are the advantages and disadvantages of contracting out public employment services and health assessments and what is the optimum model for doing so?

This research draws on robust impact evidence (where it exists) and provides qualitative perspectives on the processes of contracting out employment services and health assessments, both in the UK and internationally. Thus, this report helps to improve the understanding and practice of contracting out.

Specific research questions (see Annex 1) are grouped under three broad topics: (i) contract design,¹⁶ (ii) market competition and (iii) live running. These topics cover a comprehensive process of contracting out and include subsequent stages from how contracts are designed, how contractors are selected, and how the contracts are managed (*Figure 1*).

Figure 1. Key stages of the process of contracting out



¹⁵ Other delivery models include partnership and grant funding arrangements.

¹⁶ This area includes benefits and costs of contracting out.

1.3. Methods

This research is based on a targeted review of literature, 46 semi-structured interviews and a synthesis workshop. Annex 1 illustrates how these different methods contributed to answering the specific research questions.

1.3.1. Literature review

A Quick Scoping Review (QSR) was chosen due to pragmatic considerations about the timeline of the research, which took place between November 2021 and March 2022 (the review had to be completed by early February 2022 to inform the programme of interviews – see section 1.3.2).

The QSR focused on sources published in English, related to the UK or other members of the Organisation for Economic Co-operation and Development (OECD) and published between 2011 and 2021. The search strategy is outlined in Annex 2. Additional sources were identified via snowballing, and this included some highly relevant sources published prior to 2011 as well as material shared directly by some interviewees. After being assessed against inclusion/exclusion criteria, data from included sources were inputted in a data extraction template in Excel (see Annex 2), enabling systematic analysis against key themes and specific research questions.

1.3.2. Semi-structured interviews

Interviews were conducted to complement the QSR and to fill in gaps in understanding of the processes of contracting out (employment services and health assessments) and obtaining in-depth views and perspectives from a range of stakeholders around the research questions. A long list of interviewees was agreed with DWP and it included three main groups of stakeholders: (i) **11 DWP staff** (employees with knowledge on the policy and strategy, commercial and contracting issues faced by commissioners), (ii) **20 employment service providers** operating in the UK (primes and small organisations) with experience of delivering contracted services, (iii) **15 experts** (i.e. academics and researchers who investigate the issue of contracting out in the UK and internationally). A total of 46 interviews were conducted between February and March 2022. Interviews were conducted remotely via Microsoft Teams and lasted approximately 60 minutes. Interviews were semi-structured to enable comparability whilst allowing some flexibility to tailor the discussion according to the interests and expertise of the individual. The topic guide for the interviews can be found in Annex 3. Interviews were recorded (with interviewees' permission) and the transcripts were analysed thematically. Interviewee codes used in this report use letters 'C', 'E' or 'P' to indicate interviewees representing commissioners, experts and providers, respectively.

Additional insights were provided by representatives from DWP in form of comments to an early version of this report. When these comments present new or alternative perspectives, these are incorporated with the interview findings and marked as contributions from DWP representatives for clarity and transparency.

1.3.3. Synthesis workshop

The purpose of the workshop was two-fold: (i) to elicit feedback on the preliminary findings from the literature review and interviews (examine if these are plausible, identify inconsistencies and gaps) and (ii) to consider possible solutions to further improve the commissioning processes at DWP.

The online workshop (lasting 2 hours) was held on 24 March 2022 with 18 participants from different stakeholder groups: 11 commissioners of employment support and health assessments,¹⁷ four providers of contracted out employment services and providers' representatives, and three experts: academics and researchers. The list of workshop participants was agreed with DWP and included some stakeholders who participated in the interviews, as well as those who were not interviewed. More information about the workshop is presented in Annex 4.

1.3.4. Strengths and limitations of the methodology

To the knowledge of the authors, this is one of the few studies that explores contracting out in such a comprehensive way, including diverse perspectives from practitioners and academics, commissioners and providers.¹⁸

A number of limitations to this methodology should be noted. Whilst it takes a structured approach, a QSR does not follow the same levels of rigour as a systematic review or rapid evidence assessment.¹⁹ It is possible that some relevant sources were missed, particularly those published prior to 2011. The focus on English-language sources may likewise have resulted in certain findings being excluded. One major limitation to the review is the dearth of literature on contracting out health assessments specifically. Where this evidence exists, it is highlighted as text boxes in this report for greater visibility.

In addition, the study draws on a relatively limited number of in-depth interviews; different viewpoints or additional findings might have emerged if the pool of interviewees were larger. This particularly pertains to the health assessments, where interviews were conducted only with commissioners and the perspectives of providers were not included.²⁰ Interview findings in relation to health assessments – even if representing only the views of commissioners – are included in text boxes for better visibility.

1.4. Structure of this report

The remainder of this report is structured as follows:

¹⁷ This number includes all representatives from the DWP. However, not all of them have been responsible for contracting out services.

¹⁸ Others include Langenbucher & Vodopivec (2022) and Whitworth & Carter (2018).

¹⁹ Collins, A.M., Coughlin, D., Miller, J., Kirk, S. (2015). *The Production of Quick Scoping Reviews and Rapid Evidence Assessments: A How to Guide*.

²⁰ This was due to contract negotiations in health assessments that coincided with the research.

- Section 2 presents key findings in relation to contract design (including cost and benefits, risks, specifications and other aspects associated with Payment by Results (PbR) contracts).
- Section 3 offers insights on market competition (including selection criteria used and types of contracts).
- Section 4 outlines live running (or execution of contracts from the commissioners' perspectives).
- Section 5 offers key conclusions and outlines implications from this research.

2. Contract design

Key findings on contract design

- PbR contracts may bring benefits in terms of programme efficiency and value for money. However, there is mixed evidence on the impact of PbR contracts on programme effectiveness. The main costs associated with PbR contracts are linked with undesirable behaviour from providers such as creaming and parking. Contracting out may also weaken the accountability of public services.
- For commissioners, risks associated with PbR contracts include overpaying, underpaying and market failure. For providers, these are high upfront costs, underperformance and reputational concerns. In addition, contracting out may incur a risk of variable service provision for users/beneficiaries.
- Black box provision allows providers greater flexibility and may facilitate innovation, but evidence is scarce. Flexibility results in greater service variability and may increase the risk of creaming and parking. Minimum service requirements are designed to assure basic service standards.
- The prime model of contracting builds long-term relationships with large contractors who can take on the risk of a PbR contract. However, there is a risk of subcontractors not being able to benefit fully from contracts.
- Contract Package Areas are considered by some interviewees as too large (and not conducive to horizontal integration) and successful contractors may lack local expertise, although there is limited evidence in this area.
- An optimal length of contracts is considered by interviewees between five and seven years, where there is a perceived need for services lasting that long.

This chapter summarises evidence on the benefits, costs and risks associated with the contracting out of employment services (and health assessments – where evidence exists). The chapter considers how these factors are shaped by elements of contract design, such as payment structures (e.g., outcome payments versus service fee), levels of specification (e.g., black box versus more prescriptive approaches), the prime/sub-prime model, contract length and geographical divisions (e.g., Contract Package Areas).

This chapter focuses primarily on the design of PbR contracts, which have dominated the contracting out of employment services in the UK in recent decades (see Box 1). Cost Plus contracts (an example of contracted out provision without PbR) are covered in section 2.4. Some arguments made about the effects of contracting out and the optimum model for doing so are not specific to PbR (even if most contracted out programmes where evidence is available include PbR elements), and where this is the case efforts are made to draw this out in the text.

Before the specific features of PbR contracts are presented, a basic description of the PbR mechanisms are outlined below:

- **Commissioner:** usually a public authority (here: DWP)

- **Service provider(s):** organisations entitled by the commissioner to compete for delivery of services
- **Results:** outcomes of service provision to clients (here: job placements for employment support or established appropriate level of benefit claims for health assessments)
- **PbR contract:** an agreement determining what services are to be provided (when, where, for whom, and how), what the expected targets are, and how are these paid for:
 - **Upfront or service fee** – a fee for services paid to provider per customer for offering some basic service to help providers with an initial cost of setting up and providing the service until results (paid for outcomes) are expected to start to materialise
 - **Outcome payments** – price paid to providers for results achieved; these can vary for client groups (e.g., lower price for clients easier to place into jobs and higher price for clients with complex needs) and may include fees for achieving long-term outcomes (e.g., keeping clients in work over a certain period) or exceeding the targets set for performance. There may be any number of separate outcomes which attract payments.

2.1. Costs and benefits associated with contracting out

The main potential benefits of contracting out employment services identified in the literature and interviews are increased programme efficiency (section 2.1.1) and effectiveness (section 2.1.2). Concepts such as effectiveness, efficiency, cost effectiveness, and value for money are central to understanding the costs and benefits of contracted out provision. Whilst there is no universally agreed definitions of these concepts, this is how they are used in this report:

- **Effectiveness:** the extent to which a programme achieves the desired outcomes.²¹
- **Efficiency:** the degree to which results are achieved in an economic and timely way.²²
- **Cost effectiveness:** the effect of a programme relative to the cost.²³
- **Value for money (VFM):** using public resources in a way that creates and maximises public value.²⁴

A programme may be effective without being efficient, for instance if it achieves the desired outcomes but at a prohibitively high cost. Conversely, programme efficiency may come at the cost of effectiveness if low-cost services are pursued to a degree that compromises quality.

²¹ OECD (2021).

²² OECD (2021).

²³ IDinsight evidence toolkit. As of 16/06/2022: <https://www.evidencetoolkit.org/cost-effectiveness>

²⁴ Department for Transport (2015). Value for Money Framework.

The main possible costs associated with contracting out are ‘gaming’ behaviours on behalf of providers which can undermine programme quality and disadvantage job seekers who face greater challenges finding employment (section 2.1.3), as well as concerns around weakened transparency and accountability (section 2.1.4).

Whether these benefits and costs materialise in reality is another question. On many, if not all, of these points, the existing evidence is limited, with mixed results.^{25,26} It is difficult to reach overarching conclusions about the effect of contracting out (relative to public employment services (PES) provision or no provision) for several reasons:

- Limited empirical evidence, linked to methodological challenges associated with estimating the impact (whether in terms of effectiveness, efficiency, cost effectiveness etc.) of contracting out (see Annex 5 for a more detailed discussion)
- Challenges establishing a counterfactual from which to draw robust comparisons between different forms of provision (e.g., contracted out versus PES)
- Difficulties estimating the effect of contracting out as distinct from effects associated with specific elements of contract design (PbR, black box etc.).
- Empirical studies generally compare contracted out provision (using PbR) to PES provision (or no provision), but where differential outcomes are observed it is difficult to say whether this was due to contracting out per se or to PbR specifically (or indeed other aspects of contract design).

Theoretical arguments made in the literature and empirical evidence about the advantages and disadvantages of contracting out (see Table 1) focus on PbR, so this section focuses on such contracts. However, the available literature does not allow us to compare outcomes for PbR versus other forms of contracted out provision (such as Cost Plus contracts) or to directly compare different PbR designs (e.g., placing lesser or greater emphasis on outcome payments).

Synthesising the evidence in this area to draw overarching conclusions is complicated by the following factors:

- Studies relate to different contexts and use different research methodologies
- Studies differ in terms of the programmes they relate to and the extent to which and how these programmes incorporate different elements of contract design.

2.1.1. PbR contracts may bring benefits in terms of programme efficiency and value for money

Amongst potential benefits of contracting out employment services is the flexibility to scale up or down the provision as needed and the ability to better tailor services to specific customer groups.²⁷ It has been also argued that contracting out employment services can improve their efficiency and cost effectiveness, reducing expenditure on

²⁵ Langenbucher & Vodopivec (2022).

²⁶ Stephan (2016).

²⁷ Langenbucher & Vodopivec (2022).

activities unlikely to produce job outcomes²⁸ and strengthening incentives for delivery organisations to pursue cost savings.²⁹

A desire to increase the efficiency and cost effectiveness of programmes and to achieve VFM has been a key motivating factor for implementing PbR contracts in countries such as the **US** (where they are sometimes described as Pay for Success (PFS) programmes),^{30,31} **Australia**, the **UK**, **Sweden** and the **Netherlands**.³² The empirical evidence described in this chapter is mixed: some studies find PbR contracts for employment programmes are associated with outcomes such as efficiency, cost effectiveness and VFM, whereas this appears not to be the case for other programmes.

In the **UK**, the impact assessment of the Work Programme (WP) assessed cost effectiveness and VFM.³³ Following an Intention to Treat (ITT) approach,³⁴ the study compared outcomes for job seekers claiming out of work benefits before and after March 2017, when referrals to the WP closed. The comparison made in this study was between those eligible for the WP and those who would have been eligible for the WP if it was still in operation. However, the authors recognise limitations such as a potential lack of external validity (generalisability of the findings) associated with focusing on participants at the end of the WP.³⁵ The cost benefit analysis³⁶ found that outcomes achieved by DWP resulted in an estimated £0.83 (with a range of £0.56 to £1.10) in return for each £1 spent by DWP over the observed two-year period, but this rose to £1.76 (with a range of £1.21 to £2.32) when the findings were extrapolated over a four-year period.³⁷ This suggests that over the medium to long-term, the WP was cost effective and achieved VFM for the taxpayer. However, the comparison drawn is with no provision. It is not possible to infer from this analysis whether the WP was more/less cost effective compared PES provision or a different form of contracting out (e.g., non-PbR). Evidence on other UK contracted out programmes is summarised in *Box 1*.

²⁸ Finn, D. (2012). 'Subcontracting in Public Employment Services: The Design and Delivery of "Outcome Based" and "Black Box" Contracts', European Commission

²⁹ Andersson, F., Jordahl, H., and Josephson, J. (2019). Outsourcing Public Services: Contractability, Cost, and Quality. IZA Discussion Paper Series. No 12401.

³⁰ Albertson, K., Fox, C., O'Leary, C., Painter, G., Bailey, K., & Labarbera, J. (2018). Payment by Results and Social Impact Bonds. *University of Chicago Press Economics Books*.

³¹ Child, C., Gibbs, B. G., & Rowley, K. J. (2016). Paying for success: An appraisal of social impact bonds. *Global Economics and Management Review*, 21(1-2), 36-45.

³² Struyven, L. & Steurs, G. (2004). 'Quasi-market reforms in employment and training services: first experiences and evaluation results' in Evaluation of Systems and Programmes: Third report on vocational training research in Europe, Luxembourg: Office for Official Publications of the European Communities, (Cedefop Reference Series 57)

³³ Kay, J & Marlow, S. (2020). 'The Work Programme: A Quantitative Impact Assessment', UK Department for Work and Pensions

³⁴ The Intention to Treat (ITT) approach in randomised experiments includes all participants in the statistical analysis according to the group they were originally assigned regardless of what treatment (if any) they received.

³⁵ The payment model for WP differed across the contract. To begin with, the contract included an upfront service fee, but this was later abolished, moving to 100% outcome payments.

³⁶ Cost benefit analysis for this impact assessment was conducted using the DWP Social Cost-Benefit Analysis Framework. See: Kay & Marlow (2020).

³⁷ Kay & Marlow (2020).

Box 1. Body of evidence on contracting out employment services from the UK

The **New Deals** were a series of workfare programmes³⁸ implemented in the UK between 1998 and 2009 using both contracted out and in-house provision, including targeted programmes for young people (18-24), older people (50+), people with disabilities and lone parents.

- The **New Deal for Young People** (NDYP) (1998-2009) was a mandatory programme for unemployed workers aged 18-24 who had been claiming Job Seeker Allowance (JSA) continuously for six months or more. An evaluation of longer-term outcomes associated with the NDYP found that participants spent on average 64 fewer days claiming JSA and other labour market benefits over a four-year period compared to a comparison group of workers who were slightly too old to participate in the programme.³⁹
- The **New Deal for Disabled People** (NDDP) (1998-2009) was a voluntary employment programme available to people claiming incapacity benefits. Services were provided by Job Brokers comprised of a mix of public, private and third sector organisations.⁴⁰ An evaluation found that more NDDP participants than non-participants left incapacity benefits and entered paid employment.⁴¹ The cost-per-placement for NDDP was estimated to be £2,000-£3,000, rising to £4,000-£5,000 for a sustained job outcome of six months or more.⁴² Cost-benefit analysis found that the Government saved money (£3.41-£4.50 for continuing claimants and between £1.71-£2.26 for new claimants) for each pound spent on the NDDP programme.⁴³

Employment Zones (EZs) (2000-2009) used contracted out employment services (under 'black box' PbR contracts – see section 2.3) to support unemployed workers in deprived areas into employment. EZ provision was originally targeted at long-term unemployed workers aged 25 and older but was later expanded to include young people and lone parents.⁴⁴ A synthesis of the evidence on EZs⁴⁵ concludes that they were more effective than comparative New Deal programmes in terms of job outcomes achieved. However, EZs were found to offer less VFM than comparable New Deal programmes, particularly NDYP.⁴⁶

The **Flexible New Deal** (FND) (2009-2010) was an employment programme for long-term unemployed workers.⁴⁷ In addition to Jobcentre Plus (JCP), services were provided by prime providers commissioned on a PbR basis⁴⁸ who were given considerable flexibility in service provision (the 'black box' approach).⁴⁹ An evaluation based on survey and interview data found that the proportion of young people (aged 18-24) participating in the programme who had secured paid work was higher in Phase 2 areas (where the FND was introduced in 2010) compared to Phase 1 areas (where the FND was introduced in 2009), providing some indication that the programme's predecessor – NDYP – may have produced better outcomes for this group.⁵⁰ This evaluation found

³⁸ Workfare is a term used to describe social assistance programmes that require recipients to participate in work or other activities (e.g., training). See: Lødemel, I., & Trickey, H. (Eds.). (2001). *'An offer you can't refuse': workfare in international perspective*. Policy Press.

³⁹ Beale, I, Bloss, C and Thomas, A. (2008). The longer-term impact of the New Deal for Young People. Department for Work and Pensions Working Paper No 23.

⁴⁰ Greenberg, D. H., & Davis, A. (2007). Evaluation of the New Deal for Disabled People: The cost and cost-benefit analyses. Department for Work and Pensions.

⁴¹ Stafford, B., (2012). Supporting moves into work: New Deal for Disabled People findings. *Scandinavian Journal of Disability Research*, 14(2), pp.165–176. DOI: <http://doi.org/10.1080/15017419.2011.558235>

⁴² Greenberg and Davis (2007).

⁴³ Greenberg and Davis (2007).

⁴⁴ Griffiths, R and Durkin, S. (2007). Synthesising the evidence on Employment Zones. Department for Work and Pensions Research Report 449

⁴⁵ Griffiths and Durkin (2007).

⁴⁶ Griffiths and Durkin (2007).

⁴⁷ Job seekers who had been claiming Job Seeker's Allowance (JSA) for 12 months or more.

⁴⁸ Mulheim, I and Menne, V. (2008). The Flexible New Deal: Making it work. Social Market Foundation. As of 16.06.22: <https://www.smf.co.uk/wp-content/uploads/2008/09/Publication-The-Flexible-New-Deal-Making-it-work.pdf>

⁴⁹ Vegeris et al (2011). Flexible New Deal evaluation: Customer survey and qualitative research findings. Department for Work and Pensions.

⁵⁰ Adams, L., Oldfield, K., Riley, C., Vegeris, S., Husain, F., Bertram, C., ... & Vowden, K. (2011). *Jobseekers Regime and Flexible New Deal Evaluation: findings from longitudinal customer surveys and qualitative research* (Vol. 767). A report of research carried out by Policy Studies Institute on behalf of the Department for Work and Pensions.

that work outcomes achieved by people participating in the FND were largely sustainable. No impact assessment was published for the FND.

Pathways to Work (2003-2011) was an employment programme for people claiming incapacity benefits. Initially delivered by JCP, a provider-led model was introduced in 2008 where services were delivered by contracted private and third-sector organisations. Providers were paid for job outcomes and sustained job outcomes, and received a service fee for taking people onto their caseloads.⁵¹ An evaluation found the net impact of the provider-led Pathways had no statistically significant effect on employment, although there was a reduction in benefit claims.⁵² Outcomes were similar for the JCP and provider-led Pathways.⁵³

Work Choice (2010-2017) was a tailored employment programme for people with a disability or health condition.⁵⁴ An evaluation of the Work Choice programme found that providers experienced challenges achieving job outcome targets.⁵⁵

Work Programme (WP) (2011-2017) was an employment programme for the unemployed.⁵⁶ Results from the impact assessment are discussed elsewhere in this chapter.

New Enterprise Allowance (NEA) (2011-2022) was a programme to support the unemployed into self-employment.⁵⁷ A qualitative evaluation found that the programme offered effective support to people interested in self-employment.⁵⁸ No impact evaluation is available at the time of writing.

The **support for the very long term unemployed trailblazer** was a randomised controlled trial (RCT) of support for people reaching the end of their two years on the WP. The trial identified positive results, with those receiving support spending significantly less time on benefits and more time in employment compared to the control group.⁵⁹

The Work and Health Programme (WHP) (2017-present) is an employment programme for people with disabilities and those long-term unemployed.⁶⁰ No impact evaluation is available at the time of writing.

Intensive Personalised Employment Support (IPES) offers one-to-one support and training for people with disabilities. No impact evaluation is available at the time of writing.

Restart (2021-present) provides support for Universal Credit and JSA Income Based claimants who have been claiming for at least 9 months (with no upper limit on the duration of claims).⁶¹ The Restart Scheme has only been in operation since 2021 so it is too early for results from an impact assessment to be published.

⁵¹ Knight, G et al. (2013). Provider-led Pathways to Work: Net impacts on employment and benefits Department for Work and Pensions Working Paper No 113

⁵² Knight, G et al. (2013).

⁵³ Knight, G et al. (2013).

⁵⁴ A total of 210,330 individuals were referred to Work Choice between 2010 and 2017 and 158,420 started the programme. Of those who started the Work Choice programme, 35% achieved a short-term job outcome and 23% achieved a sustained job outcome. See: [Work Choice: referrals, starts and job outcomes to October 2019 - GOV.UK \(www.gov.uk\)](https://www.gov.uk/work-choice-referrals-starts-and-job-outcomes-to-october-2019)

⁵⁵ DWP (2013). Evaluation of the Work Choice Specialist Disability Employment Programme: Findings from the 2011 Early Implementation and 2012 Steady State Waves of the research. See: [rrep846.pdf \(publishing.service.gov.uk\)](https://publishing.service.gov.uk/rrep846.pdf)

⁵⁶ Almost 2 million individuals were referred to the WP between 2010 and 2017. Around a third (32%) of claimants referred to the programme achieved job outcomes. 39 out of 40 providers on the WP achieved minimum performance levels as set out in the contracts. See [Work Programme statistical summary: data to June 2020 - GOV.UK \(www.gov.uk\)](https://www.gov.uk/work-programme-statistical-summary-data-to-june-2020)

⁵⁷ Almost 300,000 individuals took part in the New Enterprise Allowance. As of December 2021, 161,000 businesses were set up through the NEA programme by 157,000 individuals. See: [New Enterprise Allowance statistics: April 2011 to December 2021 - GOV.UK \(www.gov.uk\)](https://www.gov.uk/new-enterprise-allowance-statistics-april-2011-to-december-2021)

⁵⁸ DWP (2013). New Enterprise Allowance: Qualitative evaluation. See: [rrep836.pdf \(publishing.service.gov.uk\)](https://publishing.service.gov.uk/rrep836.pdf)

⁵⁹ DWP (2013). Support for the very long term unemployed trailblazer: longer term analysis of benefit impacts

⁶⁰ As of February 2022, 310,000 individuals have been referred to the WHP and 210,000 have started the programme. 78% were from the disability group, 10% from the long-term unemployed group and 12% from the Early Access group. See: [Work and Health Programme statistics to February 2022 - GOV.UK \(www.gov.uk\)](https://www.gov.uk/work-and-health-programme-statistics-to-february-2022)

⁶¹ Prior to January 2022, the Restart Scheme was focused on individuals who had been claiming Universal Credit and part of the Intensive Worksearch Regime for between 12 and 18 months.

Evidence from **Australia**⁶² supports the idea of improved cost effectiveness associated with PbR contracts. Summarising findings from evaluations of the Job Network, two sources^{63,64} highlight the fact that providers achieved lower placement costs compared to previous PES provision (i.e., cost per outcome). The cost per employment outcome achieved by contracted providers was around 50 per cent lower compared to public provision.⁶⁵ Whilst outcomes achieved by providers in the Job Network were similar to previous programmes delivered by the PES, these outcomes were achieved with fewer resources, suggesting improvements in efficiency.⁶⁶ However, cost per job outcome achieved is an imperfect measure because there may be differences in the type of services delivered by the two provider types and/or the profile of job seekers they support (particularly in this case because the provision was at different points in time). Moreover, lower costs per job outcome may be achieved by 'gaming' practices such as creaming and parking (see section 2.1.3) i.e., prioritising job seekers with a stronger employment profile for whom achieving outcomes is less resource intensive.

Other studies find no evidence of improved efficiency, cost effectiveness or VFM associated with PbR contracts. An RCT looking at the effect of contracting out employment services for unemployed university graduates in **Denmark** (relative to PES provision) under PbR contracts found no difference in public expenditure associated with the two programmes (coupled with comparable effectiveness, suggesting similar value for money for the taxpayer).⁶⁷ An RCT comparing the impact of job search activities provided by private providers and the PES in **France** found that only publicly delivered services were associated with a reduction in the amount of time spent unemployed, thereby reducing expenditure on unemployment benefits (contributing to VFM).⁶⁸

A literature review drawing on evidence from **France, Germany, Denmark and Sweden** found that whilst contracting out is associated with increased cost effectiveness for services that are more straightforward to contract out, such as cleaning, waste disposal and transportation, for employment services the effect is neutral or negative (i.e., a decrease in cost effectiveness).⁶⁹ This is consistent with findings from within the **UK** where outsourcing in employment shows mixed evidence i.e. including both increased and decreased cost effectiveness.⁷⁰ However, it is not clear whether all the programmes considered in the literature have a PbR design.

Efficiencies and improvements in cost effectiveness achieved by PbR contracts should be offset against transaction costs (e.g., costs associated with writing

⁶² Dockery, A. M. & Stromback, T. (2001). 'Devolving public employment services: preliminary assessment of the Australian experiment', *International Labour Review*, 140(4), pp.429-451

⁶³ Dockery & Stromback (2001).

⁶⁴ Struyven & Steurs (2004).

⁶⁵ Dockery & Stromback (2001).

⁶⁶ Struyven & Steurs (2004).

⁶⁷ Rehwald, K., Rosholm, M., and Svarer, M. (2015). Are Public or Private Providers of Employment Services More Effective? Evidence from a Randomized Experiment. IZA Discussion Paper Series. No. 9365.

⁶⁸ Behaghel, L. et al. (2012). 'Private and Public Provision of Counselling to Job-Seekers: Evidence from a Large Controlled Experiment', IZA Discussion Paper No.6518

⁶⁹ Andersson, F., Jordahl, H., and Josephson, J. (2019). Outsourcing Public Services: Contractability, Cost, and Quality. IZA Discussion Paper Series. No 12401.

⁷⁰ Sasse, T., Guerin, B., Nickson, S., O'Brien, M., Pope, T. and Davies, N. (2019). Government outsourcing: What has worked and what needs reform? Institute for Government.

specifications, evaluating bids and monitoring performance) in assessing VFM.^{71,72} Transaction costs will depend on the approach taken to contracting out and the systems put in place, but generally amount to only a small proportion of total costs. However, stakeholders in the **Netherlands** cited in one qualitative research study described “*costly, over-complex contractual models*” which detracted from the cost effectiveness of contracted out employment services.⁷³

In sum, evidence on the impact of PbR contracts on the efficiency, cost effectiveness and VFM of employment services is mixed. In countries such as **Australia** and the **UK**, PbR contracts have been associated with cost effectiveness and/or VFM, but in other contexts (**France, Denmark**) they have not. There seems to be consensus in the reviewed literature that:

- (i) PbR is not equally well suited for all public services (it works well in simple, transactional services like waste management but less well in more complex services, like health or education)⁷⁴
- (ii) it is a difficult form of contracting⁷⁵
- (iii) benefits (and costs) of PbR contracts in employment services depend on contract design and monitoring arrangements;⁷⁶ concerns have been also raised in the literature that incentives to cut costs created by PbR contracts may result in negative consequences for staff providing services such as poorer employment conditions and lower wages.^{77,78}

⁷¹ Cockx, B., & Baert, S. (2015). Contracting Out Mandatory Counselling and Training for Long-Term Unemployed. Private For-Profit or Non-Profit, or Keep it Public? IZA DP No. 9459

⁷² Rehwald et al. (2015).

⁷³ Lindsay, C. & McQuaid, R. W. (2009). 'New Governance and the Case of Activation Policies: Comparing Experiences in Denmark and the Netherlands', *Social Policy and Administration* 43(5), pp.445-463. Page 458.

⁷⁴ Gash et al. (2012); NAO (2015). Outcome-based payment schemes: government's use of payment by results.

⁷⁵ Sasse et al. (2019).

⁷⁶ Langenbucher & Vodopivec (2022); Stephan (2016); NAO (2015). Outcome-based payment schemes: government's use of payment by results.

⁷⁷ Bach-Mortensen, A. & Barlow, J. (2021). Outsourced austerity or improved services? A systematic review and thematic synthesis of the experiences of social care providers and commissioners in quasi-markets. *Social Science & Medicine*. 276

⁷⁸ Stephan (2016).

Effective contracting of employment and health services

Table 1. Empirical evidence on the benefits of contracting out employment services using PbR

Study	Country	Services contracted out	Recipient group	Contract design	Research/ evaluation design	Comparator group(s)	Outcome type	Specific outcome(s)	Effect of contracting out/PbR
Kay and Marlow (2020)	UK	Work Programme – employment support with flexible ‘black box’ provision	Economically inactive and unemployed. Impact assessment focused on those aged 25+ in receipt of benefits for 12+ months ⁷⁹	PbR, weighted towards outcome payments differentiated by group	Impact assessment using ITT approach	No provision	Cost effectiveness ⁸⁰	Return on investment ⁸¹	+
							VFM ⁸²	Return on investment	+
							Effectiveness	Additional days in employment and not in receipt of benefits	+
Rehwald et al. (2015)	Denmark	Employment counselling and training	Newly registered unemployed with a university degree	PbR (operation subsidy and performance-based bonus)	RCT	PES provision	VFM	Net expenditure	0
							Effectiveness	Regular employment, employment with wage subsidies, non-benefit receipt, unemployment	0
Dockery and Stromback (2001)	Australia	Job Network – job-matching activities, job-search training and intensive assistance	Unemployed	PbR, upfront fee + payment for job outcomes, additional bonus for placing long-	Descriptive statistics	Mixed provision (PES and private contractors)	Cost effectiveness	Net cost per outcome	+

⁷⁹ Payment group 2 – the largest payment group in the WP

⁸⁰ In the medium-long term

⁸¹ For participants, DWP, the Exchequer and society

⁸² In the medium-long term

Effective contracting of employment and health services

Study	Country	Services contracted out	Recipient group	Contract design	Research/ evaluation design	Comparator group(s)	Outcome type	Specific outcome(s)	Effect of contracting out/PbR
				term unemployed					
Struyven and Steurs (2004)	Australia	Job Network – job-matching activities, job-search training and intensive assistance	Unemployed	PbR, upfront fee + payment for job outcomes, additional bonus for placing long-term unemployed	Descriptive statistics	Mixed provision (PES and private contractors)	Cost effectiveness	Net cost per outcome	+
Behaghel et al (2012)	France	Intensive job search programmes	Newly unemployed in specific geographical areas	PbR – 30% upfront fee, 70% performance-based payment	RCT	PES, no intensive support (control group)	VFM	Time spent claiming unemployment benefits	0 relative to no intensive support; effect is + for PES relative to no intensive support
							Effectiveness	Exit rate to employment	+ relative to no intensive support, but effect weaker than PES compared to no intensive support
Cockx and Baert	Belgium (Flanders)	Counselling and training	Long-term unemployed	PbR – 30% outcome-based payment	Event history analysis	PES	Effectiveness	Exit to employment	+
Pastore (2019)	Italy (Lombardy)	Dote Unica Lavoro - voucher system for employment services	Unemployed	Bonus payments linked to achieving sustained job outcomes	Descriptive statistics	PES	Effectiveness	Sustained employment/ self-employment for six months	-

Effective contracting of employment and health services

Study	Country	Services contracted out	Recipient group	Contract design	Research/ evaluation design	Comparator group(s)	Outcome type	Specific outcome(s)	Effect of contracting out/PbR
Koning and Heinrich (2010)	Netherlands	Employment services	People claiming unemployment/disability benefits	PbR – shift from partial to 100% outcome-based payments	Probit model using administrative data	100% PbR compared to lower outcome-based payments	Effectiveness	Job placement rate	+ for cohort claiming unemployment benefits 0 for cohort claiming disability benefits

Note: + means contracting out/PbR had a positive effect compared to the comparison group, - means negative effect and 0 means neutral/no effect

Findings from the interviews

It is important to note that DWP – similar to other commissioners – uses a **hybrid model** which combines outcome payments with service fees.⁸³ Therefore, investigations of the benefits of PbR in this context involve determining the optimal balance of these two elements within a payment model.⁸⁴

Interviewees across all stakeholder groups expressed the opinion that PbR **drives performance** by focusing the attention and efforts of providers on efficient achievement of the desired outcomes.⁸⁵ For this mechanism to work, clearly defined, relatively short-term and measurable outcomes are required.⁸⁶ Weighting a contract towards outcome payments was also thought to **incentivise quick responses** to problems or unexpected changes, so providers avoid any disruption to their achievement of payable outcomes.⁸⁷ Ultimately, PbR creates a ‘lever’ for the commissioner which they can use in combination with the level of specification to drive efficient and effective behaviours in providers.⁸⁸

In a contract weighted towards outcome payments, any ‘waste’ of DWP expenditure can be understood to be minimised as the commissioner will only pay for successful outcomes.⁸⁹ However, this opinion assumes the outcomes would not have been achieved without the contracted service.⁹⁰

‘A fundamental problem is that the outputs being paid for are rarely if ever the outcomes that we are aiming for. (...) A major concern in [value for money] considerations is deadweight – paying for outcomes which would have been achieved without the intervention – and PbR generally does nothing to reduce this.’ - DWP representative

To a certain extent, PbR presents **financial risk** to providers, alongside the potential of making substantial profit.⁹¹ This is explored in greater detail in section 2.2.

On the one hand, one possible effect of PbR is providers are likely to be operating a low cost, ‘lean’ business which can be viewed as a **maximisation of efficiency** for these organisations.⁹² For service users, providers’ response to PbR can result in a **more focused and time-sensitive** provision as wasting service user’s time or failing to create a sustainable outcome is disincentivised.⁹³ However, the model may also incentivise gaming strategies likely to reinforce inequalities amongst service users (discussed in section 2.1.2).

⁸³ C1, E4, P1, P6, P18, P20.

⁸⁴ E4, P1, P6, P20.

⁸⁵ C1, C2, C4, C7, C8, C9, C10, E2, E9, E11, P4, P8, P10, P11, P12, P14, P16, P17, P18, P20.

⁸⁶ C7, C8, E4, P6.

⁸⁷ E11.

⁸⁸ C3.

⁸⁹ C10, E5, E7, P1, P7, P14, P16, P18.

⁹⁰ C8, E9, E15, P3, P13.

⁹¹ C1, C3, C5, E3, P1, P2, P7, P8, P20.

⁹² E11, P12.

⁹³ C8, C9, E4, P2, P14.

On the other hand, the staggering of income produced by PbR can result in a **cashflow problem** for providers which impacts negatively on service quality, investment and innovation.⁹⁴ Fluctuation in payable outcomes over the course of a contract can result in providers **lowering or cutting investment**, especially towards the end.⁹⁵ Even a temporary dip in income, or a period of making a loss, can cause a 'downward spiral' for the provision and the health of the provider organisation.⁹⁶

Some expert and provider interviewees held the view that PbR **drives innovation**.⁹⁷

'The main driver here would be the value for money idea, this idea that you ideally only pay what you get for, but you also, in that sense, obviously spur innovation, because usually [...] there should be incentives to get outcomes for as many people as possible.' - Expert interviewee

This statement suggests that the incentive to achieve more or better outcomes, and to do this more efficiently, encourages the development of new approaches and problem-solving. However, there were also reservations about the potential for innovation in PbR contracts which are heavily weighted towards outcome payments, with the lack of immediate income making investment in innovation almost impossible.⁹⁸

Interviewees differed on how best to facilitate successful PbR contracts: some commissioner and provider interviewees insisted that PbR cannot work efficiently if the commissioner interferes in delivery,⁹⁹ while other commissioners and providers maintained that specification was needed to maximise benefits and contain risks or costs.¹⁰⁰

2.1.2. Mixed evidence on the impact of PbR contracts on programme effectiveness

The contracting out of employment services is expected to improve service quality by fully utilising the knowledge and expertise of external providers.¹⁰¹ It has been argued in the literature that competition drives improvements in the quality and effectiveness of employment services.¹⁰² PbR contracts specifically are thought to lead to improved performance because they encourage providers to focus on the purpose of the service and to invest in effective strategies.¹⁰³ PbR contracts can be used to direct efforts towards specific desired outcomes, such as higher paying and/or more

⁹⁴ C1, E1, E4, E5, E8, E9, P4, P6, P13, P14, P16.

⁹⁵ C1, P3, P4, P6, P7, P14, P20.

⁹⁶ C3, E4, E7, E11, P6, P13, P17.

⁹⁷ C4, E3, E7, E13, P9, P11, P18.

⁹⁸ C1, P3, P4, P6, P7, P14, P18, P20.

⁹⁹ C2, P2, P7, P11.

¹⁰⁰ C8, P8.

¹⁰¹ HoPES – Working Group (2013). PES Efficiency Working Group. Final report. European Commission.

¹⁰² Finn (2012).

¹⁰³ Finn (2012).

sustainable employment.^{104,105} It matters greatly, therefore, which outcomes are linked to payments under PbR contracts.

The impact assessment of the WP in the **UK** compared outcomes for groups of job seekers eligible for the WP¹⁰⁶ and those who would have been eligible for the programme if it had still been in operation after March 2017. The study estimates the WP was associated with 46 additional days in employment and not in receipt of an Out of Work benefit for participants over a two-year period.¹⁰⁷

Cockx and Baert (2015)¹⁰⁸ compare mandatory counselling and training for the long-term unemployed provided by the public, private and non-profit organisations in **Belgium (Flanders)** using event history analysis. Their analysis estimates the likelihood of leaving unemployment for employment (relative to remaining unemployed or leaving the labour market) under different conditions. The PbR element in this contract was relatively small (30%), with a bonus of €500 paid per placement if the provider managed to attain a placement rate at least three percentage points above a target set at the regional level. The study finds that private providers were slightly more effective than PES in achieving job outcomes for the unemployed. One year after assignment, exit rates to employment were 2.8 percentage points higher for private providers compared to the PES, and the return to unemployment was postponed by 1.4 months. Private providers achieved better outcomes on average than not-for-profit organisations, but these differences were small and not statistically significant. The authors hypothesise several reasons for the better job outcomes achieved by private providers. Private providers were new to the market and had a greater incentive to develop a good reputation, particularly because the contracts were likely to be re-tendered. Private providers were larger than non-profits and may have been able to benefit from economies of scale. In addition, even though private providers/non-profits were not able to refuse referrals, there is evidence to suggest that the composition of their clientele differed (with private providers supporting job seekers with better employment prospects). This suggests that some of the differential may be due to selection bias (see Annex 5) rather than provider type.

Literature reviews identify other countries in which PbR contracts have been associated with improvements in the effectiveness of employment services relative to PES provision, including **Ireland** (JobPath programme) and **New Zealand**.¹⁰⁹ However, research from other countries paints a more mixed picture in terms of the impact of PbR contracts on programme effectiveness, as described below.

The evidence on the impact of reforms in **Germany** (2003-2005) which introduced quasi-markets in employment services and a voucher system, are mixed: one

¹⁰⁴ Finn (2012).

¹⁰⁵ Shutes, I. and Taylor, R. (2014). Conditionality and the financing of Employment Services - Implications for the Social Divisions of Work and Welfare. *Social Policy & Administration*. 48 (2), p204-220.

¹⁰⁶ The impact assessment of the Work Programme focused one of the payment groups (Payment Group 2), comprised of individuals aged 25 years or over and in receipt of Jobseeker's Allowance or Universal Credit who had been in the Searching for Work conditionality group for 12+ months.

¹⁰⁷ Kay & Marlow (2020).

¹⁰⁸ Cockx, B., & Baert, S. (2015). Contracting Out Mandatory Counselling and Training for Long-Term Unemployed. Private For-Profit or Non-Profit, or Keep it Public? IZA DP No. 9459

¹⁰⁹ Langenbucher & Vodopivec (2022).

evidence review finds that private providers tend to secure better outcomes in terms of job quality (wages, job satisfaction, fixed-term employment and employment duration) compared to the PES¹¹⁰ but a literature review drawing on multiple studies concludes there is no evidence that private (relative to public) provision, or the introduction of a voucher system increased the likelihood of participants entering employment.¹¹¹ The voucher system was intended to drive improvements in performance through competition and consumer choice.^{112,113} The limited effectiveness of the voucher system in Germany (in terms of driving improvements relative to public provision) may be partly attributable to the fact that many voucher recipients did not use them.¹¹⁴ Job seekers – particularly those with a lower level of education – may not have the necessary information to make a meaningful choice under the voucher system, and in certain regions there may not be a sufficient number of providers to enable true competition.¹¹⁵ Market information was imperfect, since participants often did not have information about the performance of different providers.¹¹⁶ One study comparing job outcomes (defined as at least six months continued employment or self-employment) for private and non-profit providers and the PES under a voucher system in **Italy** identified better outcomes for the PES (and to a lesser extent non-profits) compared to private agencies.¹¹⁷

Koning and Heinrich (2010)¹¹⁸ estimate the effect of private provision of employment services under PbR in **the Netherlands**. Over the period 2002-2005, the PbR element of welfare-to-work contracts was strengthened, moving to a model where 100% of payment was outcome-based. Using administrative panel data, this study finds that shifting to a full PbR model was associated with an increase in job placements observed for unemployed workers, but not for those claiming disability benefits. There was no increase in average duration of employment associated with the shift to full PbR; additional job placements were largely confirmed to temporary positions. The authors conclude:

‘These benefits of performance-based contracting appear to be limited to short-term outcomes, and thus, their effectiveness in increasing the long-run impact of public programs such as welfare-to-work services remains questionable.’¹¹⁹

An RCT evaluating the impact of the contracting out of employment services via PbR contracts for unemployed university graduates in four municipalities in **Denmark**¹²⁰

¹¹⁰ Ayaita, A. et al, (2021). 'Job Placement via Private vs. Public Employment Agencies: Investigating Selection Effects and Job Match Quality in Germany', IZA Discussion Paper No.14024

¹¹¹ Jacobi, L. & Kluve, J., (2006). 'Before and After the Hartz Reforms: The Performance of Active Labour Market Policy in Germany', IZA Discussion Paper No.2100

¹¹² Jantz, B. et al., (2018). 'Marketization and Varieties of Accountability Relationships in Employment Services: Comparing Denmark, Germany, and Great Britain', *Administration and Society* 50(3), pp.321-345

¹¹³ Bruttel, O., (2005). 'Delivering active labour market policy through vouchers: experiences with training vouchers in Germany', *International Review of Administrative Sciences* 71(3), pp.391-404

¹¹⁴ Jacobi & Kluve (2006)

¹¹⁵ Bruttel (2005)

¹¹⁶ Jantz et al. (2018)

¹¹⁷ Pastore, F. (2019). The quasi-market of employment services in Italy. IZA DP No. 12662

¹¹⁸ Koning, P. & Heinrich, C. J., (2010). "Cream-Skimming, Parking and Other Intended and Unintended Effects of Performance-Based Contracting in Social Welfare Services", IZA DP No. 4801

¹¹⁹ Koning & Heinrich (2010).

¹²⁰ Rehwald et al. (2015).

found no evidence of differences in the effectiveness of public and private programmes. However, there was some evidence of better job outcomes secured by private providers for specific groups, namely men and those with prior work experience. This could indicate private providers' greater investment in groups deemed easier to place (creaming and parking, see section 2.1.3) although this is not addressed in the source.

Behaghel et al. (2012)¹²¹ compare outcomes for intensive job search programmes delivered by private providers and the PES in **France** (comparing them to a control group who did not receive intensive support) using an RCT design. Private providers operated under a PbR framework, with a 30% upfront payment and 70% conditional on securing job outcomes (35% linked to starting a job within six months; the remaining 35% paid if the worker was still employed after six months). Due to differences in types of job seekers supported by the two programmes, comparisons could only be drawn in relation to the newly unemployed in certain geographical areas. The study finds that both programmes were effective, increasing exit rates to employment (relative to the control group who did not receive intensive job counselling). However, the impact was greater in the programme delivered by the PES compared to private contractors. After six months, the exit rate was 4.2% higher for the private programme relative to the control group and 9.1% higher for the public programme. The authors argue this relates to the large upfront payment for private providers, encouraging them to enrol as many job seekers as possible. In contrast, the public employment agency had to manage scarce resources and had no incentive to maximise enrolment. In recent years, the involvement of private contractors in providing employment services in France has declined and in particular services targeted at the most vulnerable job seekers have been re-integrated into the PES.¹²²

Findings from the interviews

Any disruption caused to outcomes due to circumstantial or environmental change could be financially debilitating, leaving providers **vulnerable to economic shifts** or other uncontrollable factors.¹²³

'I think the Work Programme was almost all payment by results, which did cause issues with cash flow, did then lead to maybe only [having] companies that could [...] borrow easily, because you end up paying for that borrowing [...] also, if you go too far on the PbR route [...] it leaves providers very vulnerable to fluctuations in the economic context that they can't control, and volume contexts that they can't control.' - Commissioner interviewee

These levels of uncertainty and financial precarity are understood to have the additional effect of **forcing small and medium-sized providers out** of the

¹²¹ Behaghel, L. et al. (2012). 'Private and Public Provision of Counselling to Job-Seekers: Evidence from a Large Controlled Experiment', IZA Discussion Paper No.6518

¹²² Manoudi, A. et al. (2014). 'Small Scale Study on PES Business Models', European Commission

¹²³ C1, C2, C6, C10, E8, P7, P13.

market.¹²⁴ These effects will likely have a negative impact on programme provision, with providers focusing on risk management over delivery and specialist providers not being present to provide tailored support.

Nonetheless, a **narrow focus on outcomes** was seen as detrimental to the quality of the service and the experience of users by interviewees across commissioners, providers and experts alike.¹²⁵ This has been seen to manifest negatively in the following ways:

- an incentive for rapid outcomes, rather than sustainable ones (depending on how outcomes are defined)¹²⁶
- a loss of staff or a less skilled workforce as staff feel over-pressured to reach outcome targets¹²⁷
- the absence of incentives, resources and will to innovate¹²⁸
- a lack of resources allocated to partnership or 'joined up' provision¹²⁹
- the side-lining of service users who may need more costly or time-consuming support before an outcome can be claimed.¹³⁰

Some interviewees, across the three stakeholder groups, suggested that poorly designed PbR can 'hurt' a programme.¹³¹ For example, 'aggressive' PbR was thought to hinder the effectiveness of the Work Programme,¹³² while the cost of financing programmes¹³³ with low service fees was said to divert funds from service users.¹³⁴ Meanwhile, some alternative payment mechanisms were seen, by commissioner and provider interviewees, to be better suited to certain programmes' design,¹³⁵ enabling greater success in programmes such as the Job Entry Target Support programme (JETS).¹³⁶

2.1.3. The main costs associated with PbR contracts are linked with undesirable behaviour from providers such as creaming and parking

One potential cost associated with contracting out employment services is an increased risk of practices such as parking and creaming (or cherry-picking), which have a detrimental effect on the prospects of jobseekers who are harder-to-place.

¹²⁴ E5, E8, P2, P4, P7, P8, P14, P16.

¹²⁵ C8, C9, E1, E2, E4, E5, E9, E11, E15, P3, P5, P8, P13, P17, P20.

¹²⁶ E4.

¹²⁷ C9.

¹²⁸ E1.

¹²⁹ E2.

¹³⁰ P3, P5, P8.

¹³¹ C10, E4, P2, P6.

¹³² E4.

¹³³ This refers to the 'cost of money' in regard to loans and interest incurred by providers.

¹³⁴ P2.

¹³⁵ C3, C10, P15.

¹³⁶ C10.

Evidence from countries including the **UK**,^{137,138} the **USA**,¹³⁹ **Australia**,^{140,141} **Germany**,¹⁴² **Denmark**,¹⁴³ **Italy**¹⁴⁴ and the **Netherlands**^{145,146} supports the idea that contracting out is associated with gaming practices that draw resources and support away from job seekers who are most in need.

Gaming practices are concerning from an equity perspective because the most vulnerable will be left without sufficient support. These practices could also undermine VFM (since some of these outcomes would have been achieved anyway).^{147,148} The term 'gaming' implies deliberate strategies on behalf of providers to maximise profit, but these behaviours may be inadvertent or deemed necessary to avoid making a loss. We differentiate between gaming (which has negative consequences for some customers) and rational economic practices among providers to deliver services in the most economic ways (see section 4.5).

Countries where provision is contracted out have developed and adapted their approach over time, with a particular focus on limiting gaming practices.^{149,150} Evidence suggests the effect of contracting out on gaming practices can be mitigated by 'accountability levers'¹⁵¹ built into contract design. Contract design can be used to weaken or counteract incentives to direct support away from jobseekers who are hard-to-place.

Accountability levers to address cherry picking relate to eligibility and the degree to which providers can choose to accept or reject job seekers.^{152,153} Objective eligibility criteria¹⁵⁴ and an independent referral process,¹⁵⁵ as well as making it mandatory for providers to accept referrals are suggested as tools to address cherry picking. In **Australia**, private providers cannot refuse referrals, and evidence suggests this has been effective in limiting cherry picking.¹⁵⁶ However, there is a danger that if private providers are obliged to accept all referrals, then this will result in parking.¹⁵⁷

A common approach to address creaming and parking is to differentiate payments for different target groups, with higher payments for placing jobseekers with poorer

¹³⁷ Rees, J. et al. (2014). 'Support for All in the UK Work Programme? Differential Payments, Same Old Problem', *Social Policy & Administration* 48(2), pp.221-239

¹³⁸ Carter (2018a).

¹³⁹ Finn, D. (2010). *Outsourcing Employment Programmes: Contract Design and Differential Prices*

¹⁴⁰ Dockery & Stromback (2001).

¹⁴¹ Finn (2012).

¹⁴² Ayaita, A. et al. (2021). 'Job Placement via Private vs. Public Employment Agencies: Investigating Selection Effects and Job Match Quality in Germany', IZA Discussion Paper No. 14024

¹⁴³ Lindsay, & McQuaid (2009).

¹⁴⁴ Pastore, F. (2019). *The quasi-market of employment services in Italy*. IZA DP No. 12662

¹⁴⁵ Finn (2010).

¹⁴⁶ Koning, P. & Heinrich, C. J. (2010). "Cream-Skimming, Parking and Other Intended and Unintended Effects of Performance-Based Contracting in Social Welfare Services", IZA DP No. 4801

¹⁴⁷ Carter, E. (2018a). 'Making Markets in Employment Support: Promises and Pitfalls in the Work Programme's Private Power Market'. PhD Thesis

¹⁴⁸ Dockery & Stromback (2001).

¹⁴⁹ Finn (2010).

¹⁵⁰ Finn (2011).

¹⁵¹ Carter, E. (2021).

¹⁵² Carter (2021).

¹⁵³ Finn (2011).

¹⁵⁴ FitzGerald, C. et al., (2019). 'Walking the contractual tightrope: a transaction cost economics perspective on social impact bonds', *Public Money and Management* 39:7, pp.458-467

¹⁵⁵ FitzGerald et al. (2019).

¹⁵⁶ Dockery & Stromback (2001).

¹⁵⁷ Ayaita et al. (2021).

employment prospects.^{158,159} In **Ireland** job seekers are assigned to different payment groups according to unemployment duration.¹⁶⁰ In **Australia** outcome-based payments increase with the duration of unemployment.¹⁶¹ In **Italy** the payment amount linked to a voucher for achieving a successful job outcome depends on the profile of the job seeker.^{162,163}

Under the WP in the **UK**, participants were placed into one of 10 claimant groups as a proxy for their perceived support needs.¹⁶⁴ Higher fees were paid in response to job outcomes for claimants in groups deemed harder-to-place.¹⁶⁵ The pricing model was influenced by the 'Non-Intervention Rate' (NIR) – the percentage of participants it is estimated would have achieved a successful outcome without assistance.¹⁶⁶ Drawing on WP data and qualitative research,¹⁶⁷ Creaming and parking practices were widespread in the early stages of the WP, something blamed on the programme's weak accountability framework.¹⁶⁸ The authors suggest the design of the payment structure may not have been sufficiently well aligned with the level of difficulty associated with placing job seekers (although it would be difficult, if not impossible, to create perfect alignment), weakening the effect on creaming and parking.¹⁶⁹ This source describes how, rather than using the claimant groups as defined under the DWP payment structure, many providers used their own system of 'triaging' participants, for instance the red, amber, green (RAG) rating system used by some providers. Some providers confirmed at that time the RAG approach was used to determine which participants to direct the most resources to, including within payment groups. Imbalances in job outcomes achieved by the different payment groups (as evident in programme statistics) may indicate some degree of creaming and parking as well as variation in the challenges faced by different groups in finding work.¹⁷⁰ The authors observe:

*'Claimant groups [payment groups in the Work Programme] were a blunt instrument orientated primarily around the prior benefit received and not necessarily coterminous with a customer's distance from the labour market.'*¹⁷¹

Another source focusing on WP notes that for some groups, payments may not be high enough to cover the cost of supporting them into work.¹⁷² The considerable heterogeneity within payment groups (modelling suggests greater variation within

¹⁵⁸ Finn (2011).

¹⁵⁹ Langenbucher & Vodopivec (2022).

¹⁶⁰ Langenbucher & Vodopivec (2022).

¹⁶¹ Langenbucher & Vodopivec (2022).

¹⁶² Langenbucher & Vodopivec (2022).

¹⁶³ Pastore, F. (2019). The quasi-market of employment services in Italy. IZA DP No. 12662

¹⁶⁴ Kay & Marlow (2020).

¹⁶⁵ Carter (2021).

¹⁶⁶ Finn (2012).

¹⁶⁷ Rees et al. (2014).

¹⁶⁸ Whitworth, A., & Carter, E. (2018). 'Rescaling employment support accountability: From negative national neoliberalism to positively integrated city-region ecosystems', *Environment and Planning C: Politics and Space* 36(2), pp.274-289

¹⁶⁹ The authors note that their findings are consistent with the presence of creaming and parking in the Work Programme, but do not prove that these took place.

¹⁷⁰ Rees et al. (2014).

¹⁷¹ Rees et al. (2014).

¹⁷² Carter E. and Whitworth A. (2015). Creaming and parking in quasi-marketised welfare-to-work schemes: Designed out of or designed into the UK Work Programme? *Journal of Social Policy* 44(2): 277–296.

payment groups than between them) may have led providers to focus their efforts on those with the best employment prospects in each group.¹⁷³ Qualitative research found that some WP providers admitted to seeing job-ready participants more frequently. However, most programme participants thought the support they received was adequate and well matched to their needs.¹⁷⁴

As a means of addressing this issue in future programmes, some commentators have argued for the use of ‘accelerator’ or ‘escalator’ models, where providers are paid according to the percentage of participants who achieve job outcomes in each payment group,¹⁷⁵ see Table 3 (section 3.3).

Ayaita et al. (2021)¹⁷⁶ find the introduction of a voucher system in **Germany** in 2002 reduced gaming practices and allowed unemployed workers to choose support from a range of private providers. Agencies could redeem the voucher from the public employment agency if they secured a positive job outcome.¹⁷⁷ The voucher amount received by the provider was higher for the long-term unemployed. Private providers also received remuneration from the new employer of the placed worker. The study finds evidence of gaming practices between 1995 and 2002. In this period, long-term unemployed workers were more likely to be placed by the public employment agency than private providers, suggesting private providers were focusing their efforts on workers with higher qualifications and better employment prospects. The same trend is evident in the period 2003-2018. However, the effect is substantially larger in the earlier time period where the voucher system was not in place, suggesting that vouchers were effective at reducing creaming and parking. Moreover, there is evidence that individuals placed privately with vouchers were more likely than those placed privately without vouchers to have poorer employment prospects (e.g., lower educational qualifications, experience of unemployment).¹⁷⁸

Findings from the interviews

The largest concern around PbR was indeed the possibility of **incentivising gaming** behaviours, including parking and creaming.¹⁷⁹

‘The ones that are most easy to place into a job [...the providers] cream those ones off and work hardest on getting them into a job, but the hardest ones to place, [they] park and do minimal stuff with them, and they won’t get jobs. So, from the provider’s point of view, that’s good in terms of return on their investment. From the government’s point of view, it’s probably not what you want, because the people you want getting into work are those hardest to help ones. The ones that are closer to the labour market may well have found those jobs themselves anyway, and so you’re paying for dead weight.’ - Expert interviewee

¹⁷³ Carter & Whitworth (2015).

¹⁷⁴ Newton, B., Meager, N, and Sainsbury, R. (2014). Work Programme evaluation: the participant experience report Department for Work and Pensions research report 892.

¹⁷⁵ Langenbucher & Vodopivec (2022).

¹⁷⁶ Ayaita et al. (2021).

¹⁷⁷ This was defined as a socially insured job of at least 15 hours per week with a duration of at least three months.

¹⁷⁸ Ayaita et al. (2021).

¹⁷⁹ C1, C7, C8, C9, E1, E4, E7, E9, E11, E13, E13, P1, P2, P5, P8, P11, P17.

However, views were mixed on which types of gaming were most harmful or, indeed, if some instances could be acceptable:

‘The problem with some payment by results systems is the parking aspect, so, to me, creaming and parking are often thrown together [but] they’re very different things: creaming can be “let’s get people who are ready for work in work quickly and then we’ve got more resources”. Parking is [...] the opposite of that: it’s “let’s ignore those people who could be getting more time and more resources because we can focus on this low hanging fruit”.’ - Provider interviewee

Some expert and provider interviewees noted that mitigating strategies can be used to counter this tendency towards gaming in PbR contracts.¹⁸⁰ Possible strategies that were described include increased prescription and auditing,¹⁸¹ the use of complementary policies to incentivise focusing on ‘harder to help’ users,¹⁸² differential or accelerator payment mechanisms,¹⁸³ a reduction of the emphasis on price competition¹⁸⁴ and the use of earning thresholds to calculate payable outcomes.¹⁸⁵

Alternative perspectives on PbR costs and gaming were also offered by representatives from DWP. This new material is incorporated below:

‘The main cost of a PbR contract is the same as with any other type of contract – paying for deadweight or not paying enough to make a difference and the main risk is setting the thing up with performance expectations that aren’t aligned with what’s actually possible.’ - DWP representative

‘There is some (admittedly limited) evidence that net impact of provision can be greater for those who are closer to the labour market. So it is possible in principle that a focus on those who are easier to help might improve overall programme effectiveness in terms of total net impact – but at the same time (...) [it may have] negative consequences for equity.’ - DWP representative

2.1.4. Contracting out may weaken the accountability of public services

The degree to which contracting out services weakens accountability is contested in the literature.¹⁸⁶ Some argue that contracting out weakens accountability, whereas others argue to the contrary that contractual agreements and minimum standards strengthen it. Commissioners remain accountable for the services they procure from

¹⁸⁰ E8, E12, P6.

¹⁸¹ E8.

¹⁸² E8.

¹⁸³ E12.

¹⁸⁴ E12.

¹⁸⁵ P6.

¹⁸⁶ Mulgan, R. (2006). Government Accountability for Outsourced Services

external agencies.¹⁸⁷ However, the details of these arrangements are generally commercially confidential and protected from public scrutiny.¹⁸⁸

Evidence related to the costs and benefits of contracting out health assessments is very limited (see *Box 2*).

Box 2. Health assessments: suitability of the PbR model

The reviewed literature focused primarily on contracting out employment services or took a wider perspective of commissioning public services more broadly. None of the reviewed sources discussed contracting out health assessments.

Some limited evidence emerged from the interviews. It was noted that PbR **may not be the right model for health assessments**.¹⁸⁹ While this payment mechanism can encourage efficiency, the combination of a limited market and the importance of correct procedure during assessments may counteract this.¹⁹⁰

2.2. Risks associated with contracting out

2.2.1. For commissioners, risks include overpaying, underpaying and market failure

When employment services are delivered by PES, the success (or failure) of supporting people into work falls primarily on the public sector. This situation becomes more complex when public authorities commission the services externally: the risk – or a possibility of failing to support people back to work – could arguably rest with providers (their knowledge, abilities, good will, etc. or lack of these), or with commissioners (the way the service contracts were designed and managed), or with both parties. PbR contracts are often described as transferring risk from commissioners to providers.^{191,192,193} However, PbR contracts (and indeed, other contracting out without PbR) are not without risk to commissioners and to an extent these risks can be mitigated by contract design. The government Sourcing Playbook notes that outsourcing complex services carries more risk for commissioners.¹⁹⁴

A key risk for commissioners is overpayment i.e., paying providers for outcomes that would have been achieved even without the services delivered.^{195,196} Provider behaviours such as cherry picking, creaming and parking (which may be exacerbated by PbR – see section 2.1.3) contribute to overpayment since they direct resources away from job seekers for whom programmes might have the greatest impact.¹⁹⁷

¹⁸⁷ Mulgan (2006).

¹⁸⁸ Mulgan (2006).

¹⁸⁹ C3, C8.

¹⁹⁰ C3.

¹⁹¹ Considine, M., O'Sullivan, S., & Nguyen, P. (2018). The policymaker's dilemma: The risks and benefits of a 'black box' approach to commissioning active labour market programmes. *Social Policy & Administration*, 52(1), 229-251.

¹⁹² Finn (2012).

¹⁹³ HoPES – Working Group (2013). PES Efficiency Working Group. Final report. European Commission.

¹⁹⁴ HM Government (2021). The Sourcing Playbook Government guidance on service delivery, including outsourcing, insourcing, mixed economy sourcing and contracting', Open Government licence v3.0.

¹⁹⁵ FitzGerald et al. (2019).

¹⁹⁶ FitzGerald et al. (2019).

¹⁹⁷ Carter (2018a).

Differential payments (i.e., higher payments for securing job outcomes for workers who are harder-to-place) are designed in part to address the risk of overpayment, since providers are incentivised to work with groups for whom job outcomes are less likely to be achieved without support. However, even the differential payments create a rather simplistic framework, given the diversity of the client group. In addition, linking outcome payments for placing the most disadvantaged groups in sustainable work, requires long follow up periods, and late outcome payments – which diminish their appeal to providers. There may also be a risk of fraudulent claims of outcomes by providers (e.g. claiming outcomes which have not occurred or claiming that the individual who achieved outcomes was in a higher payment group – in either case leading to overpayment), although little evidence was identified to substantiate this.¹⁹⁸

On the flip side, underpayment also presents a risk for providers under PbR i.e., not investing enough in the programme to secure the desired outcomes. Underpayment can result in poor provider performance associated with cost-cutting measures and market failure (see section 2.2.2).

From the commissioners' perspective, a key risk is poor performance from providers¹⁹⁹ and, in the extreme, market failure, where the commissioning body needs to 'bail out' the provider or find an alternative provider,^{200,201} incurring high transaction costs.²⁰² Strategies to manage the risk of poor provider performance (other than PbR to strengthen incentives and improve performance) include setting performance targets and minimum service requirements (see section 2.3.4),²⁰³ measuring service quality through monitoring and benchmarking (see section 4.1) and contract provisions that penalise providers for poor performance. The risk of market failure can be mitigated by having a larger number of providers rather than a few 'primes' (see section 2.6).²⁰⁴ However, a larger number of providers will require more effort in terms of managing and monitoring them, increasing the costs to commissioners.

Findings from the interviews

It appears that risk should not be understood as fully transferred from commissioners to providers in PbR contracts, or indeed in any other type of contract.²⁰⁵ This can largely be attributed to the political and social responsibility held by DWP which means contracts should not be allowed to fail without contingency plans and services must be maintained unless low performance or quality demands change or termination.²⁰⁶

¹⁹⁸ Finn (2012).

¹⁹⁹ Langenbucher & Vodopivec (2022).

²⁰⁰ Finn (2011).

²⁰¹ HM Government (2021). 'The Sourcing Playbook: Government guidance on service delivery, including outsourcing, insourcing, mixed economy sourcing and contracting', Open Government licence v3.0

²⁰² Poutvaara, P. and Jordahl, H. (2020). Public sector outsourcing. The desirability of outsourcing the provision of public services depends on their characteristics and market conditions. IZA World of Labor. 65 (2).p1-11.

²⁰³ Langenbucher & Vodopivec (2022).

²⁰⁴ Langenbucher & Vodopivec (2022).

²⁰⁵ C1, C3, C4, C6, C7, E4, P2, P15, P20.

²⁰⁶ C1, C3, E11, E13, P15.

Risks for the commissioner can be understood as a matter of trade-offs in which an assessment is made of which factors can or cannot be controlled.²⁰⁷

In the discussions with interviewees about risk, considerations of financial risk were most prominent for all stakeholder groups. **Overspending** is one such risk; without a mechanism to set a total contract value, exceeding a programme budget may be financially damaging for the commissioner.²⁰⁸

'And the risk is [...] that actually it's too easy or the provider's too successful and we can't afford it, they've performed so highly we run out of money paying them, and that becomes a problem for the Department and for Treasury who are then overspending what they plan to spend. Arguably you could say "well, if the provider's continuing to delivery additionality, it's well worth spending more" and that case can certainly be made, but you still need to find the money from somewhere.' - Commissioner interviewee

This is distinct to **overpaying** in which the commissioner bears the risk that they may not have secured the best price for the service²⁰⁹ or may be paying for outcomes that would have occurred without the provision.²¹⁰ This would be a scenario in which VFM is not achieved. In another, policy or circumstantial changes could mean the commissioner paying for the duration of a contract, knowing its effectiveness, and therefore its VFM, is now reduced.²¹¹

Indeed, **underspending** is another risk for the commissioner, with both financial and social implications.²¹² With 'money left on the table', the commissioner might miss an opportunity to use their full financial capacity for maximum impact on their goal, namely moving jobseekers into employment. In some scenarios, the programme or its replacement could have a **future budget reduced** on the basis of a gap between actual and planned expenditure.²¹³

DWP has already taken the following actions to assess and limit risk:

- using risk profiles to restrict bidders on the basis of financial assets and stability (Financial Viability Risk Assessment)²¹⁴
- utilising insights from modelling and large data sets to forecast risk²¹⁵
- running pilots and other trials²¹⁶
- planning for legal risks or litigation.²¹⁷

The commissioner also, to some extent, bears risk on behalf of the service users, given its political and social responsibilities. This means the commissioner will aim to avoid the social risk of **any reduction in service quality or any direct or indirect**

²⁰⁷ C3, C5.

²⁰⁸ C5, C6, C8, C10, C11, E4, P18.

²⁰⁹ C8, P2, P6.

²¹⁰ C5, C8, E13.

²¹¹ C4, C8, C11, E4, E11, E13, P9.

²¹² C6, C8.

²¹³ C8.

²¹⁴ C4.

²¹⁵ C4.

²¹⁶ C5.

²¹⁷ E4, P9.

harm to service users.²¹⁸ Thus, contracting out, while passing on operational responsibility, is not thought to transfer this risk and the commissioner must, instead, seek to limit the potential for providers to, intentionally²¹⁹ or otherwise,²²⁰ under-serve jobseekers.

The commissioner also considers how the landscape of the provider market itself may pose risks to the maintenance and quality of service provision. Indeed, disproportionate financial risk transfer to providers could increase the **risk of gaming and other negative behaviours** as providers try to secure income.²²¹

Reputational risks must also be considered.²²² These risks include negative public opinion and media coverage,²²³ the morale of DWP staff²²⁴ and the political scrutiny on public expenditure.²²⁵ Damage to the commissioner reputation could impact on the success of the programme (and indeed could spill over to other programmes).²²⁶ If service users have a negative view of providers or DWP, their engagement and efforts within a provision may suffer.²²⁷

Overall, some commissioner and expert interviewees suggested **lessons have been learned** by DWP on risk assessment and risk management.²²⁸ Some expert interviewees noted it is not always possible to price risks²²⁹ while provider and commissioner interviewees emphasised that a mature and experienced market lowers the risk involved in PbR contracts.²³⁰ Further suggested strategies to minimise or account for risk were as follows:

- negotiating more openly with providers on risk to avoid uncertainty driving up price²³¹
- carrying out risk profiles on all providers throughout supply chains²³²
- using a Target Cost Incentive Fee²³³ to set pricing but allow for agreed-upon adjustments as costs change²³⁴
- increasing marketing and communications efforts to manage reputational risk²³⁵
- strengthening minimum service guarantees and customer service standards to protect the service user²³⁶

²¹⁸ C3, C7, E4, P6.

²¹⁹ C3, C7, E4.

²²⁰ C10.

²²¹ E2.

²²² C8, E3, E4, E5, E14, P15, P16, P17, P20.

²²³ E3.

²²⁴ C8.

²²⁵ P3.

²²⁶ DWP representative.

²²⁷ C8, P8, P15, P20.

²²⁸ C2, C3, C4, E13.

²²⁹ E11, E13.

²³⁰ C4, C7, P11.

²³¹ C7, C10, P15.

²³² C4.

²³³ When a Target Cost Incentive Fee is used, allowable costs estimated at the time the price is agreed are used to determine a target cost that is multiplied by the contract profit rate to calculate the contract price. This is then subject to change, based on the difference between actual and estimated allowable costs, according to terms set out in the contract. (see SSRO (2014) Glossary of terms used in report templates: contract reports (including contract pricing statement)).

²³⁴ C4.

²³⁵ P15.

²³⁶ E4.

- setting appropriate contingency plans for low-performing providers.²³⁷

2.2.2. For providers, risks include high upfront costs, underperformance and reputational concerns

For providers, financial risk is a key consideration in taking on a PbR contract. The government Sourcing Playbook notes providers taking greater risk is usually associated with higher profit margins.²³⁸

PbR contracts, particularly those with low upfront payments, may put some, particularly smaller organisations, at risk of cash flow issues, since payments may not be made until some time after costs are incurred.²³⁹ This risk for providers may be managed by shifting the balance from conditional, outcome-based payments to fixed upfront payments.²⁴⁰ However, too great an emphasis on fixed payments may increase the risk of parking²⁴¹ and weaken the advantages of PbR contracts.

In a programme governed by a PbR framework, a key risk for providers is underperformance. Underperformance carries a financial risk since in PbR contracts payment is conditional on achieving a successful job outcome.²⁴² Research from **Australia** has suggested providers in the Job Network invested fewer resources in activities such as training that were with a decline in certain activities or services, for instance training, which were considered riskier investments because they may not result in a job for the job seeker.²⁴³

The risk of provider underperformance is shaped to a large extent by economic conditions, namely the number of job seekers and vacancies in the local labour market.^{244,245} Economic conditions can change after the terms of the contract are agreed and uncertainty in this regard is a key risk to providers, but also to commissioners, for whom provider under- and over-performance is also undesirable. The risk to both parties may be mitigated by building provisions into contracts to allow for changing circumstances and unforeseen events²⁴⁶ such as the possibility of contract renegotiation.²⁴⁷

Participating in PbR contracts can incur some reputational risk for providers. Voluntary and/or community-based organisations may feel that PbR contracts are in tension with their organisational mission to serve certain communities in a broader sense than securing specific job outcomes, and therefore may be less likely to participate in programmes.²⁴⁸ Qualitative research in the **UK** has found that some non-profit organisations were concerned about the reputational risk of participating in

²³⁷ E14.

²³⁸ HM Government (2021) The Sourcing Playbook.

²³⁹ Finn (2011).

²⁴⁰ Behaghel, L. et al., (2012). 'Private and Public Provision of Counselling to Job-Seekers: Evidence from a Large Controlled Experiment', IZA Discussion Paper No.6518

²⁴¹ Behaghel (2012).

²⁴² Shutes, I. and Taylor, R. (2014). Conditionality and the financing of Employment Services - Implications for the Social Divisions of Work and Welfare. *Social Policy & Administration*. 48 (2), p204-220.

²⁴³ Struyven & Steurs (2004).

²⁴⁴ Finn (2011).

²⁴⁵ Finn (2011).

²⁴⁶ FitzGerald et al. (2019).

²⁴⁷ Finn (2011).

²⁴⁸ Shutes and Taylor (2014).

the WP given the emphasis on conditionality and mandatory participation for unemployed workers.²⁴⁹ One may argue providers can choose contracts that suit them best – as other actors (e.g. local authorities, NHS) also contract out service provision. However, DWP is one of the main players in the area of employment.

Findings from the interviews

It is safe to assume providers understand that any economic incentive, including profit, involves risk.²⁵⁰ Viewing **financial risk as an inevitability** implies a provider should make their own judgements on whether their cashflow or reserves allow them to afford any risk attached to a contract.²⁵¹

Nonetheless, financial risk will still be the major focus for most providers.²⁵² The following concerns were raised by interviewees:

- imbalances in risk-sharing between providers and commissioner²⁵³
- financial risks involved in securing premises and hiring staff which might not be required²⁵⁴
- providers' reliance on DWP for referrals²⁵⁵
- impact of shareholders on providers' risk tolerance²⁵⁶
- uncertainty around income due to extrapolation²⁵⁷
- the difficulty of innovating when financial risk is high²⁵⁸
- possibility of going out of business²⁵⁹
- challenge of meeting high targets set by the commissioner to minimise overspend.²⁶⁰

Providers take on **risk around delivery** and the implications, financial and otherwise, of delivery failure.²⁶¹ This is especially the case in a contract with a payment mechanism closer to 'pure' PbR (which is not used, as DWP opts for a hybrid model in which outcome payments are combined with service fees).²⁶² While there are risks associated with mistakes or low performance, providers' main fears will be around disruption, delays or failure caused by factors beyond their control.²⁶³ Indeed, providers may bear the risk that the policy and its programme does not deliver the results envisaged.²⁶⁴

These factors could include **issues with other actors involved**, such as employers²⁶⁵ or even competition with other providers for engagement with these

²⁴⁹ Shutes and Taylor (2014).

²⁵⁰ C10, E3, E4, P6.

²⁵¹ E3, P6.

²⁵² C7, E4, E11, E13, P12, P15, P16, P17.

²⁵³ C1, E1, E3, E4, P2, P11, P15, P20.

²⁵⁴ P2, P5, P6, P16, P19.

²⁵⁵ P10, P13, P15, P19.

²⁵⁶ E2, P6.

²⁵⁷ P16, P17.

²⁵⁸ C4, E1.

²⁵⁹ P12, P13.

²⁶⁰ E1.

²⁶¹ C5, C7, E11, E13, P10, P17.

²⁶² E7, E13.

²⁶³ E3.

²⁶⁴ P10.

²⁶⁵ E3.

employers,²⁶⁶ but most often these factors would involve labour market changes.²⁶⁷ Some provider interviewees observed that even differences in the expected demographics of the service users can impact on delivery and performance, as a programme would have been planned for users with different needs.²⁶⁸

Interviewees from all stakeholder groups identified risks around **varying volumes of referred service users** as the most disruptive to financial stability.²⁶⁹ This variation has most often been caused by changes in the labour market or inaccurate forecasts. The knock-on effect of changing volumes for providers is outlined below:

'A significant issue for us providers [...] is that we will have assumed that, round numbers, 10,000 people are going to flow through our books, and we're going to put [a certain proportion] of those into work, and therefore the revenue, or the profit from the contract will be X and Y, and we will have told banks or, in our case, shareholders, or whoever, that's what's going to happen, and if the volumes are half that, all those assumptions are out, and therefore all of the cost assumptions are out. I may have taken a lease on an office in London assuming that I would need X, Y and Z square footage.' - Provider interviewee

But some provider interviewees also reported that **competing with other parallel programmes** was sometimes the cause of low referrals.²⁷⁰ Interviewees suggested **change clauses** for unprecedented events, such as the COVID-19 pandemic, should be included in contracts so renegotiations or a move to Cost Plus model (see section 2.4) can be triggered to reduce risk for providers.²⁷¹

Additionally, for subcontractors (especially small and medium providers), **reliance on primes for cashflow** can mean an unreasonably high level of financial risk.²⁷² Where large prime providers can rely on their resources to flexibly respond to changes in volume or performance, this is not the case for smaller organisations or even providers with fewer resources and a large supply chain.²⁷³ However, some providers may be less vulnerable to cashflow reliance – for example, the different streams of income received by local authority subcontractors mitigate financial risk, especially since these subcontractors are not required to generate profit.²⁷⁴

But other risks remain. **Reputational risk**,²⁷⁵ for example, may be particularly important to charity or local authority subcontractors who may wish to differentiate themselves from private providers.²⁷⁶

'There is reputational risk [because...] as a charity we also have to be really careful of the customer groups that we are targeting to support. We have a

²⁶⁶ E3.

²⁶⁷ C7, E1, E14, P6, P8, P9, P14, P15.

²⁶⁸ P18.

²⁶⁹ C1, C3, C4, C7, C8, C10, C11, E1, E3, E4, E7, E9, E12, E13, E14, P2, P3, P6, P8, P9, P10, P11, P13, P14, P15, P17, P18, P19.

²⁷⁰ P13, P15.

²⁷¹ C7, E12, P9, P16, P18.

²⁷² P10.

²⁷³ C5, E4, P5.

²⁷⁵ C4, C8, E11.

²⁷⁵ C4, C8, E11.

²⁷⁶ P16, P17, P20.

broad remit of people at disadvantage, so for most of the people that we'd ever work with within DWP [they are] going to meet that [...] benchmark, that's not a problem. However, we did have discussions about the JETS variation within Work and Health Programme, [thinking] "are they really the hardest to help that we really have set up as a charity [to serve]?" - Provider interviewee

Past instances of **contracting failures**²⁷⁷ in both employment services and health assessments²⁷⁸ have made the question of reputational risk particularly important and new entrants to the provider market may be wary of this association.²⁷⁹

Providers will also consider the impact upon their organisation if a policy, programme or the contracting out process **receives public critique or negative publicity**.²⁸⁰ Risks to an organisation's reputation within the provider market may also be assessed. For instance, providers aspiring to prime contracts will want to make positive impressions on the commissioner²⁸¹ and new entrants will wish to be well-regarded not only by DWP, but possibly other government departments.²⁸² It is likely these providers would then avoid contracts which could possibly damage industry reputations by risking low performance or delivery failure.²⁸³

Some provider interviewees also mentioned **extrapolation** as a destabilising practice in which funds are reclaimed by the commissioner after incorrectly paying for an outcome.²⁸⁴ This is one example of the type of uncertainty providers might face which creates risk around PbR contracts.²⁸⁵

Interviewees, including commissioners, experts and providers, expressed the opinion that PbR places **disproportionate risk onto providers**, noting this will also mean increased costs for commissioners as the providers start to price that risk into bids.²⁸⁶ The pressure on providers to maintain results may also **limit provider flexibility**, as financial rigidity makes organisations risk-averse and unable to respond to varying needs or scenarios.²⁸⁷ This could result in a negative impact on service users depending on the costs of tailoring services to their location or specific needs.

2.2.3. Contracting out may incur a risk of variable service provision for users/beneficiaries

Contracting out employment services (not specific to PbR contracts), particularly in combination with a flexible (black box) model of service provision, might incur a risk of variable service provision onto service users.²⁸⁸ On a quasi-market, inconsistencies may develop not just in the form and content of services but,

²⁷⁷ See for instance Sasse et al. (2019).

²⁷⁸ C4, C8.

²⁷⁹ E14, P8.

²⁸⁰ E3, E5.

²⁸¹ P16.

²⁸² P8.

²⁸³ P8, P17.

²⁸⁴ P16, P17.

²⁸⁵ C8, E7, E11, P7, P13.

²⁸⁶ C7, E4, E12, P5, P8, P11, P17.

²⁸⁷ C6.

²⁸⁸ Lindsay & McQuaid (2009).

crucially, in service quality.²⁸⁹ The risk of variable service quality may be contained to some degree by introducing minimum service requirements to ensure some consistency in delivery.²⁹⁰

Findings from the interviews

In the words of one interviewee:

'One of the shortcomings, of course, has been [...] not being able to actually adequately service people who are further from the labour market. That is a major shortcoming that's not been fixed in the UK or Australian model, so therefore it's a risk to the jobseeker of course.' - Expert interviewee

As noted in section 2.1.3, interviewees across the three stakeholder groups warned of the **impact of parking** upon service users.²⁹¹ The deliberate avoidance of serving users with greater or more complex needs is considered to be neglectful, leaving users without the support to which they are entitled. The long-term impact of such neglect could include compounding issues keeping users from the labour market, losing the potential for earnings and other benefits of employment. The combination of different aspects of contract design, such as payment mechanism or specification, could increase or decrease the likelihood of users being parked.²⁹²

There is also a risk that parking is made more likely by the very notion of some service users being **'hard to help'**. It may be that service users are framed as having complex needs simply because the support they need is costly²⁹³ or because progress is inevitably slower, due, for example, to language barriers or digital illiteracy.²⁹⁴ Furthermore, this phrasing may categorise some service users (e.g. people with disabilities) as a homogenous group,²⁹⁵ rather than recognise the diversity of different disabilities and associated support needs.²⁹⁶ Creating diverse supply chains with specialist subcontractors may counter-act these risks, dependent on the dynamics within the prime-sub model (see section 2.5).²⁹⁷

There is also a danger that **service users feel their voices are not heard** during their time on a programme²⁹⁸ and afterwards, during assessment or later policymaking.²⁹⁹ While some expert and provider interviewees emphasised the utility of **surveys** and similar feedback mechanisms,³⁰⁰ others had reservations.³⁰¹ The impact of factors beyond a provider's control, namely government policies, on service user views and levels of satisfaction were thought, by expert and provider

²⁸⁹ Lindsay & McQuaid (2009).

²⁹⁰ Mosley, H., G., (2011). 'Decentralisation of Public Employment Services', European Commission Mutual Learning Programme for Public Employment Services

²⁹¹ C6, C7, C8, E3, E4, E8, E13, P11.

²⁹² C3, E13.

²⁹³ C7.

²⁹⁴ E7.

²⁹⁵ E9, P1.

²⁹⁶ E15.

²⁹⁷ C7, E2, E15, P2, P7, P10.

²⁹⁸ C9, E14.

²⁹⁹ P1,

³⁰⁰ E4, E9, E12, P11.

³⁰¹ C9, E5, E8, E9, P1, P11, P14.

interviewees, to affect customer feedback.³⁰² Conversely, there were concerns among provider and expert interviewees as to whether there was enough available information to allow **public scrutiny** of provisions and their treatment of service users.³⁰³

There is also a risk of missing the balance between maintaining **consistency** across a provision³⁰⁴ and **adapting appropriately** to service user needs.³⁰⁵ Many expert and provider interviewees suggested a lack of locally grounded and ‘joined-up’ services could result in lower quality provision for some users, as the service is less tailored to their needs and location.³⁰⁶ Mitigating these risks could involve greater incorporation of a ‘**customer focus**’ into contract design, whether through greater choice for users between services³⁰⁷ or using customer satisfaction in PbR or other financial incentives.³⁰⁸

Evidence related to contracting our health assessments is scarce (see Box 3).

Box 3. Health assessments: risks in contracting out

The reviewed literature noted some risks in relation to contracting out health assessments. The example mentioned relates to the provider Atos³⁰⁹ withdrawing from conducting Work Capability Assessments (WCA) for Employment and Support Allowance (ESA) in the **UK** due to concerns about reputational risk.³¹⁰ Atos was the sole provider for carrying out work capability assessments in 2008. After DWP identified some quality concerns and following negotiations between DWP and Atos, a settlement was reached for Atos to exit the contract in early 2014 (that is before it was due to end in August 2015).³¹¹

Additional insights were gained from the interviews.

Negative effects on service quality from pressure on front-line staff is a particular concern for health assessment provision.³¹² Increasing demand for related benefits and subsequent growing caseloads could overburden staff, risking mistakes or lower quality assessments. Reputational risks for providers were also mentioned, given previous providers terminating contracts early.³¹³

In health assessments, the demand produced by the **number of benefit claimants** can be unpredictable, especially if policy on specific benefits changes during the contract.³¹⁴

In health assessments, if decision-making is not based on a high-quality assessment, this can have a negative impact on service user’s livelihood and wellbeing (if the individual’s health restrictions are under-estimated)^{315 316} or result in benefit over-payment (if they are over-estimated).

Commissioners emphasise consistency across delivery to minimise such occurrences.³¹⁷ However, interviewees suggested that such quality assurance will not protect service users if policies prove to

³⁰² E5, E9, P1, P11.

³⁰³ P1, P11, E4.

³⁰⁴ C9, E8, P14.

³⁰⁵ P11, P14.

³⁰⁶ E8, E13, P1, P5, P10.

³⁰⁷ E3, E4, E7, E12, P7.

³⁰⁸ E2, P8.

³⁰⁹ <https://atos.net/en-gb/united-kingdom>

³¹⁰ Warren, J., Garthwaite, K., & Bamba, C., (2014). 'After Atos Healthcare: is the Employment and Support Allowance fit for purpose and does the Work Capability Assessment have a future?', *Disability and Society*, 29:8, pp.1319-1323

³¹¹ <https://publications.parliament.uk/pa/cm201314/cmhansrd/cm140327/wmstext/140327m0002.htm#14032769000011>

³¹² C9.

³¹³ C4, C8.

³¹⁴ C3.

³¹⁵ Adams, L., Oldfield, K., Riley, C. and Nightingale, M. (2011). Trial incapacity benefits reassessment: customer and staff views and experiences. DWP.

³¹⁶ E14.

³¹⁷ C9, E2, E14

be unhelpful or even harmful.³¹⁸ Other suggestions included emphasising communication around an assessment as part of the service user experience as well as providing feedback to staff on the outcomes and customer view of the assessment.³¹⁹

2.3. The level of specification in contracted services

2.3.1. Black box provision allows providers greater flexibility and may facilitate innovation, but evidence is scarce

The process of contracting out is determined by specifications issued by the commissioners. This documentation formulates expectations of the purchaser towards the service provider, and it defines, and governs responsibilities and rights among the parties. Such documentation can attempt to regulate every aspect of services to be provided or leave some (or most) aspects in the provider's discretion (black box). Black box refers to:

*'A policy model where the service providers or agents are free to define their own actions including how to personalize support for unemployed people, engage them in unique activities or provide them with training and mentoring.'*³²⁰

Under the black box approach, providers are awarded discretion to determine the nature of provision other than the requirement to meet minimum service requirements (see section 2.3.2). Although there had been previous moves towards a less prescriptive approach to contracting out in the **UK**, the black box approach became more pronounced under the WP.^{321,322} The flexibility afforded by the black box approach was felt to be suitable for the WP because of the need to cater to all long-term unemployed claimants – some of whom will have distinctive or complex needs – within a single employment scheme.³²³

The flexibility afforded by black box contracts may drive improvements in performance because of the potential for providers to deliver individualised, tailored support. For instance, for job seekers with more complex needs and barriers to employment such as the long-term unemployed, those with disabilities or health-related barriers to employment and ex-offenders.³²⁴ Black box provision is also thought to provide greater scope for innovation in programme delivery. Summarising

³¹⁸ E5, E14, P1.

³¹⁹ E14.

³²⁰ Considine et al. (2018).

³²¹ Considine et al. (2018).

³²² Kay & Marlow (2020).

³²³ Rees et al. (2014).

³²⁴ Tubb, H., (2012). 'Activation and Integration: Working with Individual Action Plans', European Commission Mutual Learning Programme for Public Employment Services

the evidence from **Australia**, Finn (2012)³²⁵ notes that flexible delivery models (combination with outcome-based payments) enabled providers to tailor services to different participants, as well as to innovate, developing new practices to tackle specific barriers to employment. Under the black box approach, providers have more flexibility to adapt their services in response to the latest evidence on effectiveness and best practice.³²⁶ Another advantage of the black box approach is that it reduces administrative complexity and running costs for commissioners.³²⁷

Under the right conditions, private providers contracted under black box conditions might generate greater innovation in service delivery compared to the PES, who are more likely be constrained by factors such as capacity, resource and institutional constraints.^{328,329} Empirical evidence in this area is scarce, perhaps reflecting the fact that innovation is a challenging concept to define and measure (see section 4.5).

An evaluation of the Flexible New Deal in the **UK** noted that the black box approach is not necessarily associated with greater innovation in service provision.³³⁰ Stakeholders from the **Netherlands** interviewed for one research study³³¹ expressed concerns that contracting out had reduced innovation and resulted in greater uniformity in service provision. Stakeholders explained this lack of innovation as a result of the drive towards cost saving, benefitting from economies of scale.³³² Under a PbR system, providers may be cautious about innovating and deviating from 'tried and tested' approaches.³³³

Similarly, qualitative evidence has emerged from the **UK** in relation to DWP that resulted only in modest (rather than radical) service innovation: providers were reluctant to invest in expensive, specialist support services for clients and could not access (other) funding pots or co-ordinate with parallel employment support initiatives at the local level.³³⁴

2.3.2. Flexibility in service provision results in greater service variability and may increase the risk of creaming and parking

Black box provision may create opportunities for providers to pursue cost savings at the expense of programme quality^{335,336} unless checks and balances are put in place to prevent this. Black box PbR contracts may increase the risk of provider behaviours such as creaming, parking and cherry-picking.³³⁷ In **Australia**, policy has shifted away from black box provision due to concerns about creaming and parking resulting in low levels of service for job seekers who face greater challenges finding

³²⁵ Finn (2012).

³²⁶ Finn (2012).

³²⁷ Considine et al. (2018).

³²⁸ Finn (2011).

³²⁹ Langenbucher & Vodopivec (2022).

³³⁰ Vegeris et al (2011).

³³¹ Lindsay. & McQuaid (2009).

³³² Lindsay. & McQuaid (2009).

³³³ Lindsay. & McQuaid (2009).

³³⁴ Gash et al. (2012).

³³⁵ Rehwald et al. (2015).

³³⁶ Struyven & Steurs (2004).

³³⁷ Langenbucher & Vodopivec (2022).

employment.³³⁸ The black box approach in the **UK WP** was associated with modest changes in the self-reported behaviour of frontline employment services staff.³³⁹ Compared to before the WP, staff reported greater autonomy in their role, less reliance on standardised assessment tools and greater emphasis on tailoring and client choice. However, this study also found evidence of an increase in creaming and parking (see section 2.1.3) associated with black box provision, such a growth in the proportion of job seekers not participating in any activity and an increase in the proportion of staff saying their agency's practice was to focus on the most capable job seekers. However, the authors recognise changes between the two time periods may also be due to other factors such as economic conditions.

Black box provision may result in variable service provision, which is concerning from the perspective of equity. Ceolta-Smith et al.³⁴⁰ argue black box provision under the **UK WP** resulted in variable provision for job seekers with disabilities and health-related barriers to employment. Whilst policy documents make clear that providers should address health-related barriers and tailor provision to meet the needs of individual customers, including those with disabilities or health conditions, in practice only a minority of primes implemented health-related provision such as Personal Advisors (PAs) with specific health-related expertise, employing healthcare professionals or forming NHS partnerships. The authors note:

*'Individuals facing similar health-related obstacles to employment can expect to receive very different levels and types of support depending on which Primes' programme they are assigned to join.'*³⁴¹

In addition to variability in service provision, black box contracts make it more difficult to gather information about activities and services delivered as part of a programme, and to classify them (e.g., job search assistance, training, subsidised employment).³⁴²

Findings from the interviews

Views diverged on the functionality of black box commissioning. Some provider interviewees advocated for this model,³⁴³ while others across the three stakeholder groups suggested a '**grey box**' was a better solution.³⁴⁴ These contrasting views are explained further below.

The lack of restrictions involved in the black box model may allow providers necessary **flexibility to tailor services** to the local area³⁴⁵ or to labour market changes.³⁴⁶ Expert and provider interviewees expressed the view that, to a certain extent, black box commissioning **passes risk to providers** as they alone are

³³⁸ Finn Finn (2011).

³³⁹ Considine et al. (2018).

³⁴⁰ Ceolta-Smith, J., Salway, S., & Tod, A. M. (2015). A Review of Health-related Support Provision within the UK Work Programme—What's on the Menu?. *Social Policy & Administration*, 49(2), 254-276.

³⁴¹ Ceolta-Smith et al. (2015).

³⁴² Kay & Marlow (2020).

³⁴³ P6, P11, P16.

³⁴⁴ C6, E4, E13, P14, P18.

³⁴⁵ P10, P16.

³⁴⁶ P2.

accountable for their processes, judged only on outcomes.³⁴⁷ Commissioner and provider interviewees argued that the success of supposedly black box programmes has been **undermined by intervention** and additional specification.³⁴⁸ Indeed, in another study the WP was described as “a ‘black box’ with 300 pages of *guidance*”.³⁴⁹ Some providers claimed black box commissioning would work well if the commissioner’s goals were clear, and providers were able to work independently to deliver results.³⁵⁰

More sceptical views of the black box model suggested the lack of restrictions of providers **allows for gaming**.³⁵¹ One interviewee reported that concerns about gaming were the cause of Australian commissioners moving back towards a ‘grey box’ or higher specification.³⁵²

‘In Australia we did have a black box system at the start but it was quickly established that the providers were basically parking and creaming [...] and so there was an inquiry into this in Australia and [...] so the government of the day’s response was to introduce a whole swathe of requirements, mutual obligation requirements, that were intended to ensure that all citizens receive some level of service.’ - Expert interviewee

However, other provider interviewees suggested gaming can be avoided through other aspects of design, such as differentiated payments,³⁵³ and undesirable behaviours might be less prevalent than popularly believed:³⁵⁴

‘What you then saw, in the next set of programmes, is [...] the kind of notion that there was creaming and parking, which I think, [...] without ever actually being really properly tested and proven, has been received as being [...] folklore [that] there was creaming and parking, [when...] I personally didn’t see examples of that.’ - Provider interviewee

Additionally, commissioner and provider interviewees suggested the black box model **drives down prices**³⁵⁵ and **brings down the quality of service** to the ‘lowest common denominator’.³⁵⁶

Providers are then understood to compete for contracts based on price alone and are **incentivised to deliver as cheaply** as possible to maximise their profits on payable outcomes. These concerns about gaming and low service quality suggest **greater specification might mitigate social and reputational risks** for the commissioner.³⁵⁷ Indeed, the social risk might be of such great concern that a

³⁴⁷ E5, P6.

³⁴⁸ C2, P6, P9, P15, P20,

³⁴⁹ Gash et al. (2012).

³⁵⁰ P13, P14, P16.

³⁵¹ E4, E8, P2, P9, P11.

³⁵² E3.

³⁵³ P11.

³⁵⁴ P3.

³⁵⁵ C8.

³⁵⁶ P1.

³⁵⁷ E4.

‘completely black box’ cannot be used at all, so as to protect service users and ensure equal treatment nationally.³⁵⁸

Providers, however, contend that similarity in services delivered, even under the black box model, can be attributed to **existing industry knowledge on ‘what works’** in employment services.³⁵⁹ While this does mean that black box commissioning may not result in a variety of methods or even in significant innovation,³⁶⁰ perverse behaviours or cheap delivery are not inevitable.

The **combination of two elements: payment structure and service requirements (in the PbR contract)** is an important factor in a contract’s success. When designing a contract commissioners face trade offs between these two elements.³⁶¹ For example, some interviewees suggested that a combination of low outcome payments with high level of service requirements would undermine benefits associated with PbR, as providers’ efforts would focus on ensuring compliance with detailed specifications, rather than on finding better or more efficient ways of helping people into work.³⁶² However, some expert interviewees warned that a payment mechanism weighted heavily towards outcome payments (rather than upfront or service fees) can also limit the benefits gained from a black box model, as a scarcity of cashflow and resources may **lower the likelihood of innovation** as providers will be reluctant to take risks.³⁶³

It may be that **different programmes require different levels of specification**, depending on their goals and target service users.³⁶⁴ For example, some provider interviewees reported that Work Choice was a much more prescriptive contract than most but thought to be among the most successful.³⁶⁵

2.3.3. Minimum service requirements are designed to assure basic service standards

There is a trade-off between enabling flexibility and promoting innovation and the need to monitor programmes and to ensure they are being delivered as intended.³⁶⁶ Minimum service requirements or standards specify the type, quantity or quality of services to be delivered as part of contracted out employment programmes. They promote a certain level of quality and consistency in service provision and make it easier to monitor programmes and compare services delivered by different providers.³⁶⁷ Minimum service standards are used to guard against behaviours such as creaming, parking and cherry picking.³⁶⁸ Minimum requirements may differ across groups of service users, with additional support required for those facing particular challenges or considered further from the labour market. In **New Zealand**, service

³⁵⁸ E11, P8.

³⁵⁹ P13, P17.

³⁶⁰ E5, E9, P1.

³⁶¹ P2, P6, P20.

³⁶² C2, E13, P2, P9, P20.

³⁶³ E2.

³⁶⁴ E4, P5.

³⁶⁵ P7.

³⁶⁶ Carter (2018a).

³⁶⁷ Langenbucher & Vodopivec (2022).

³⁶⁸ Finn (2012).

users considered harder-to-place must be supported at least once a week, whereas for those deemed to have better employment prospects contact may only be on a monthly basis.³⁶⁹

In some programmes, for instance the WP in the **UK**, minimum service standards were left to be set by the provider and formed part of the contract with the commissioning body. These minimum service levels were made public so clients were able to judge whether providers delivered what they had promised. The government had the right to withdraw contracts if the minimum service requirements were not met.³⁷⁰ There was considerable variation in the commitments given and the level of detail of minimum service standards under the WP,^{371,372} detracting from their ability to promote consistency. There was also variability in the degree to which minimum service standards were communicated to service users,³⁷³ undermining their effect as an ‘accountability lever’.³⁷⁴

Findings from the interviews

A new perspective on the black box provision was explained as follows:

‘Although the combination of services delivered is in the control of the provider, that does not mean that commissioners are not allowed to see what this is. Some have suggested that ‘glass box’ might be a better term – you can see inside, but not change it.’ - DWP representative

Minimum service requirements might be considered the factor which turns a black box into a ‘grey’ box³⁷⁵ (not because of lack of transparency for commissioners but for their ability to affect service provision). Some interviewees, across the three stakeholder groups, emphasised the importance of these requirements.³⁷⁶ Minimum service requirements might be understood as a **necessary guarantee in a national service**.³⁷⁷ These requirements are a matter of social and legal responsibility in many cases, ensuring that **safeguarding and due diligence** are carried out appropriately.³⁷⁸

The contents of minimum service requirements have been subject to **review and development** by the commissioner over time.³⁷⁹ It was thought by provider interviewees that the inclusion of these conditions in a contract drives improvement and an adherence to high standards of service.³⁸⁰ Expert interviewees, however, believed the current minimum service requirements are too weak and vague.³⁸¹

³⁶⁹ Langenbucher & Vodopivec (2022).

³⁷⁰ Kay & Marlow (2020).

³⁷¹ Finn (2012).

³⁷² Langenbucher & Vodopivec (2022).

³⁷³ Langenbucher & Vodopivec (2022).

³⁷⁴ Carter, E., (2021). ‘More than marketised? Exploring the governance and accountability mechanisms at play in Social Impact Bonds’, *Journal of Economic Policy Reform* 24:1, p.88.

³⁷⁵ C10.

³⁷⁶ C10, E3, E4, E11, E13, P19.

³⁷⁷ C3, P3, P17.

³⁷⁸ E11, P15.

³⁷⁹ E4, E11, E12, P20.

³⁸⁰ P19.

³⁸¹ E4.

Ultimately, those in favour of minimum service requirements emphasised these should focus on **the service provided to jobseekers** and the compatibility of activities with the labour market and employer landscape.³⁸² This could also involve including requirements related to **social values and public service** as well as job outcomes.³⁸³

Other critiques were based on the opposite view: provider interviewees felt the commissioners had moved towards **excessive minimum service requirements** in contracts.³⁸⁴ From this perspective, the requirements may even specify activities that providers are already carrying out³⁸⁵ but, most importantly, the rigidity of these conditions places emphasis on meeting requirements rather than job seeker needs.³⁸⁶ Provider interviewees suggested the addition of a new requirement should be conditional on the removal of another to act as a balance against over-specification.³⁸⁷

Requirements to meet service users face-to-face at specific milestones were the most objectionable to provider interviewees.³⁸⁸ They suggested this requirement **lags behind developments in remote communication** and specificity on dates could hinder support work with enthusiastic users who needed more flexibility.³⁸⁹

‘There’s just one quite irritating [...] requirement stating that every month you must see every client face to face, like physically face to face, and update their action plan. [...] they’ve relaxed that requirement when there are Covid restrictions in place – but if you just take Covid out of it – it just means what you’re doing is basically a mechanical every month, you’re seeing people for the sake of seeing people because you have to, not because it’s [...] anything] to do with the client, where they’re at in their job, in the journey [...] and that’s just ridiculous because we are supposed to have a little bit of expertise in this field and that’s just a very blunt instrument.’ - Provider interviewee

The effect of such rigidity may also risk creating a **‘low trust’ relationship** between the commissioner and providers.³⁹⁰ A punitive approach with contracts focused on penalising providers may have a negative effect on service quality, as well as on long-term relations with the provider market.³⁹¹

‘I’m very much of the mindset that you write rules for the people who are looking to break them, not the ones who are compliant, and so I think you have to have a set of minimum standards but in saying that, [...] that’s what people will revert to and so you really need to be creating incentives then, I think, to make people strive for better.’ - Commissioner interviewee

³⁸² C2, C3, C8, C11, E3, E15.

³⁸³ E5.

³⁸⁴ P6, P9, P18, P20.

³⁸⁵ P17.

³⁸⁶ E3, P1, P3, P9, P11, P12, P15, P18, P20.

³⁸⁷ P3.

³⁸⁸ P6, P9, P15, P20.

³⁸⁹ P12, P15.

³⁹⁰ E7, E8, P9, P11, P15.

³⁹¹ P9.

Some provider interviewees described minimum service requirements as the cause of **excessive focus on process and paperwork**, to the detriment of more efficient work on outcomes.³⁹² The resources used to meet and record compliance were thought by provider interviewees to **limit potential innovation**, alongside the obstacles created by the specifications of the requirements themselves.³⁹³ Scepticism was expressed by expert interviewees about the effectiveness of minimum service requirements which encourage providers to innovate.³⁹⁴

It was also suggested by provider interviewees that **subcontractors might suffer disproportionately** with the burden of meeting minimum service requirements, particularly if their prime has agreed to requirements which are unrealistic for a small or specialist service, either due to resources or the service user profile.³⁹⁵ Indeed, mandated referrals can limit a provider's ability to meet requirements on contact or other activities as **jobseekers may be reluctant or ill-equipped** to engage.³⁹⁶ With cases such as these in mind, it may be better if minimum service requirements are written as **adjustable for different types of jobseeker** to allow for tailored services or even innovation.³⁹⁷

If prescription and monitoring of compliance with minimum service standards increases, it may be that any supposed benefits of contracting out are lost, including reduced administrative costs and the utilisation of provider expertise or innovation.³⁹⁸ To counter such a result, **greater collaboration with providers on minimum service requirements** might be considered.³⁹⁹ This could involve consulting with providers when drafting these requirements,⁴⁰⁰ or even allowing earlier sight of the contract so providers can negotiate or at least begin necessary organisational planning.⁴⁰¹

Evidence related to contracting our health assessments is more limited (see Box 4).

Box 4. Health assessments: level of specifications

The reviewed literature did not consider the specific arrangements for contracting out health assessments. However, some insights were provided by interviews.

The question of specification involves different considerations and can be more straightforward in the health assessment context. While one interviewee reported that black box models have been used less and less over time, they also noted that the earlier use still involved intervention and additional specification from the commissioner eventually.⁴⁰² As mentioned above, the **importance of consistency and correct procedures in health assessments** will likely limit any benefits gained from the flexibility of a black box model.

³⁹² P3, P4, P11, P15.

³⁹³ P12, P20.

³⁹⁴ E5, E7.

³⁹⁵ P13, P16.

³⁹⁶ P16.

³⁹⁷ E7, P16.

³⁹⁸ P4, P6, P15.

³⁹⁹ P7, P15.

⁴⁰⁰ E12, P7.

⁴⁰¹ P11.

⁴⁰² C2.

2.4. Cost Plus model

In contrast to PbR contracts, Cost Plus contracts reimburse contractors for agreed expenses and allow additional payments to enable them to make a profit. This approach transfers no or minimal risk to suppliers.⁴⁰³ The reviewed literature did not discuss this model of contracting out, but interviews provide some useful insights (including on health assessments – see Box 5).

Findings from the interviews

Interviewees across all three stakeholder groups reported that moving contracts onto Cost Plus during the pandemic was greatly appreciated by providers and praised **for safeguarding the long-term capacity of the market and the maintenance of service**.⁴⁰⁴ Expert interviewees also emphasised that during this period all parties demonstrated good will and commitment to the overall goals as providers continued to invest in high performance without the financial incentive of PbR.⁴⁰⁵

For instance, it was noted by provider interviewees that, in retrospect, they would have made higher profits if contracts remained on PbR during this time, or at least once the peak of uncertainty had passed.⁴⁰⁶ Despite this, provider interviewees still **expressed gratitude for DWP's rapid response** to protect and stabilise the provider market.⁴⁰⁷

Regarding future uses of Cost Plus, some interviewees, across stakeholder groups, suggested DWP could **standardise the use of Cost Plus as an emergency measure**, even for scenarios which are more feasible to forecast, such as large economic downturns.⁴⁰⁸

An alternative strategy could plan for a transfer to Cost Plus in the event of **disruptive non-emergency changes**. Some interviewees suggested that an **introductory or concluding period** on contracts which uses Cost Plus could address existing problems with cashflow or reduced investment over the lifetime of a PbR contract.⁴⁰⁹ Placing outcomes for certain client groups on a Cost Plus mechanism while others remain on PbR could also counter parking or assist providers in moving disadvantaged clients or those with complex needs closer to the labour market.⁴¹⁰

Reservations around future uses of Cost Plus were also expressed. One provider interviewee would only choose a Cost Plus contract if they had identified risks around failures to receive prompt and full payments from a client.⁴¹¹ From a commissioner perspective, Cost Plus could be seen to **transfer financial risk back to DWP**.⁴¹² Both commissioner and provider interviewees observed that Cost Plus models **do**

⁴⁰³ HM Government (2021). 'The Sourcing Playbook: Government guidance on service delivery, including outsourcing, insourcing, mixed economy sourcing and contracting', Open Government licence v3.0

⁴⁰⁴ C1, C3, C4, C5, C8, C10, E12, E13, P3, P4, P5, P10, P11, P12, P13, P14, P16, P17, P18, P20.

⁴⁰⁵ E13.

⁴⁰⁶ P2, P6, P12, P17, P20.

⁴⁰⁷ P4, P10, P11, P12, P13, P14, P16, P17, P18, P20.

⁴⁰⁸ C5, E12, P2, P5, P18, P20.

⁴⁰⁹ C3, P2, P14.

⁴¹⁰ P9, P13.

⁴¹¹ P8.

⁴¹² C1, P6.

not incentivise efficiency,⁴¹³ although it may be possible to design a contract with additional efficiency or innovation incentives such as bonuses.⁴¹⁴

One requirement of Cost Plus models is the operation of **open book accounting and cost oversight**.⁴¹⁵ This process can be burdensome to both commissioner and provider, perhaps diverting resources from work on job outcomes.⁴¹⁶

'I think it's extremely difficult to run cost models without having a large amount of cost insight. So, we were only able to really run cost models in the employment space because we'd done the work to establish open book costing management, we'd done the work to understand how money flowed through those contracts, and therefore we were in a position where that insight, aligned to the structures we'd set up, meant there was very little incentive for gaming, and our view would be, very little gaming took place.' - Commissioner interviewee

Debate was also seen to arise around which costs were counted; this process could cause **combative relations** not only between commissioners and providers, but also between primes and subcontractors.⁴¹⁷ These are similar objections to those described above around the **administrative burden** of tracking compliance with minimum service requirements.

Box 5. Health assessments: Cost Plus model

For health assessments, some contracts had used a Cost or Cost Plus model prior to the pandemic.⁴¹⁸ It may be that for this type of service, where **oversight and commissioner involvement in delivery is preferred**, a Cost Plus payment mechanism could be a better choice.⁴¹⁹ It may be **particularly suitable for services for which demand is volatile**, as can be the case for Personal Independence Payment (PIP)⁴²⁰ claims.⁴²¹ In this case, the stability of income for the provider also establishes certainty and continuity of service and investment in quality.⁴²² It was also suggested that the upcoming introduction of new IT systems to health assessments might be paired with a period on Cost Plus as providers adjust and re-calculate their costs and procedures.⁴²³

⁴¹³ C1, C8, P2, P6, P8, P9, P17.

⁴¹⁴ C3.

⁴¹⁵ C1, C2, E12, P5, P10, P12.

⁴¹⁶ P3, P8, P9, P12, P14, P20.

⁴¹⁷ C5, P6, P8, P9, P11.

⁴¹⁸ C1, C2.

⁴¹⁹ C2.

⁴²⁰ More information on PIP is available: [here](#).

⁴²¹ C3, C8.

⁴²² C8.

⁴²³ P18.

2.5. Prime model

2.5.1. The prime model of contracting builds long-term relationships with large contractors who can take on the risk of a PbR contract

The prime model refers to:

*'A system where contracts are awarded to 'large, professional placement agencies at regional level', which in turn 'awards sub-contracts to smaller, specialized, local service providers.'*⁴²⁴

Introduced in relation to the New Deals for the unemployed in 2006, the prime model featured prominently in DWP's 2008 commissioning strategy:

*'The commercial opportunities we shall offer will be arranged into larger, longer lasting (subject to performance achievements) contractual packages which we expect will be delivered by top-tier providers leading and managing diverse supply chains.'*⁴²⁵

The prime model was strengthened under the WP⁴²⁶ and subsequent DWP programmes. Under the WP, prime providers were required to demonstrate an annual turnover of at least £20 million as an indication of the ability to finance upfront investment and take on the risk of a PbR contract.⁴²⁷ The emergence of the prime model has developed in tandem with greater flexibility given to providers under the black box approach (see section 2.3.1).⁴²⁸ Prime providers may or may not provide direct delivery as well as work with subcontractors.⁴²⁹

The rationale behind the prime approach is to develop long-term relationships based on trust and cooperation⁴³⁰ with a small number of large providers who are well placed to take on the financial risk of a PbR contracts.⁴³¹ The prime/sub-prime model was designed to enable DWP and providers to work together on a longer-term basis, taking a more strategic approach and building on previous delivery experience.⁴³² Prime providers work with a network of subcontractors, enabling them to draw on the skills and capacities of smaller, specialist, community-based providers who would themselves not be able to take on the financial risk of a PbR contract. Primes may themselves use competition in their supply chains to improve performance.⁴³³

⁴²⁴ Ceolta-Smith et al. (2015). citing Weishaupt, J. T. (2011). Social Partners and the Governance of Public Employment Services: Trends and Experiences from Western Europe, Geneva: International Labour Organization Page 11.

⁴²⁵ DWP (2008).

⁴²⁶ Ceolta-Smith et al. (2015).

⁴²⁷ Carter, E., (2018a). 'Making Markets in Employment Support: Promises and Pitfalls in the Work Programme's Private Power Market'. PhD Thesis

⁴²⁸ Finn (2012).

⁴²⁹ Finn (2012).

⁴³⁰ Finn (2012).

⁴³¹ HM Government (2021). The Sourcing Playbook.

⁴³² DWP (2008).

⁴³³ Finn (2012).

Involving fewer contracts, the prime model reduces transaction costs (set up and management costs) for the commissioning body.⁴³⁴

The prime model results in contracts being concentrated in the hands of a few, large, predominantly for-profit providers. The approach makes it more difficult for new organisations to enter the market⁴³⁵ and therefore the degree of competitive pressure on providers. Awarding contracts to a smaller number of providers may stifle competition and increase the risks/implications of provider failure.⁴³⁶ Awarding contracts to a larger number of providers also has downsides – larger transaction costs and harder to monitor and enforce minimum service standards.⁴³⁷ In the **UK**, the prime model has been associated with a decline in the number of non-profit organisations involved in service provision.⁴³⁸ Similar trends towards the concentration of contracts in the hands of fewer large providers have also been observed in the **Netherlands, Australia and Denmark**.⁴³⁹

The majority of contracts in the **UK WP** (35 out of 40) went to private sector primes (of the remaining five, two were public sector organisations and three were third sector organisations).⁴⁴⁰ Similarly, the majority of primes operating as part of the Restart Scheme (7 out of 8) are private sector organisations. Although large, for-profit providers may be well placed to take on financial risk of a PbR contract, there is some evidence to suggest that compared to socially driven or not-for-profit providers, private companies may be more inclined towards ‘gaming’ practices (see section 2.1.3), requiring more robust accountability levers to be built into contract design.⁴⁴¹

2.5.2. There is a risk of subcontractors not being able to benefit fully from contracts

The prime/sub-prime model may result in a lack of oversight and control for commissioners,⁴⁴² who have less insight into the content and quality of service delivery, as well as how primes work with subcontractors.^{443,444,445} There is a risk of primes shifting the risks associated with PbR contracts onto their subcontractors⁴⁴⁶ and questions have been raised about whether subcontractors, particularly smaller voluntary or non-for-profit organisations are able to benefit fully from contracts under the prime model.⁴⁴⁷ There is a risk of subcontractors being used as ‘bid candy’ to secure contracts and then not receiving the expected number of referrals.⁴⁴⁸

⁴³⁴ Finn (2011).

⁴³⁵ Carey, G. et al. (2020). 'Quasi-market shaping, stewarding and steering in personalization: the need for practice-orientated empirical evidence', *Policy Design and Practice* 3(1), pp.30-44

⁴³⁶ Langenbucher & Vodopivec (2022).

⁴³⁷ Langenbucher & Vodopivec (2022).

⁴³⁸ Finn (2012).

⁴³⁹ Egdell, V. et al. (2016). 'Third Sector Experiences of Work Programme Delivery', *Journal of Social Policy* 45(4), pp.729-746

⁴⁴⁰ Rees et al. (2014).

⁴⁴¹ Carter (2021).

⁴⁴² Finn (2011).

⁴⁴³ Ceolta-Smith et al. (2015).

⁴⁴⁴ Considine et al. (2018).

⁴⁴⁵ Finn (2011).

⁴⁴⁶ Finn (2011).

⁴⁴⁷ Considine et al. (2018).

⁴⁴⁸ Carter (2018a).

In the **UK** WP there was a requirement to have subcontractors in the bid, but not to work with them after commissioning.⁴⁴⁹ Summarising the evidence, Carter observes that under DWP smaller subcontractors received a low number of referrals.⁴⁵⁰ Qualitative research with 13 third sector organisations in **Scotland** acting as subcontractors in DWP found that many had received fewer referrals than expected and some had not managed to offset the costs of preparing their bids.⁴⁵¹ Some third sector organisations interviewed as part of that study suggested this might be due to primes parking more difficult or complex cases who might otherwise have been referred to specialist organisations such as theirs.⁴⁵² Organisations also reported receiving referrals for different types of clients than they had outlined at the bidding stages, creating additional work for their organisation.⁴⁵³

To counteract supply chain issues, the Merlin Standard⁴⁵⁴ was introduced in 2013, as an assessment and enforcement tool to ensure compliance with DWP's code of conduct in relation to subcontracting provision.⁴⁵⁵ Under the WP, all providers were required to achieve and maintain the Merlin Standard.⁴⁵⁶ However, there is some evidence that, despite the Merlin Standard, primes in the WP passed the risks associated with a PbR contract onto subcontractors,⁴⁵⁷ as well as, in some cases, starving subcontractors of referrals.^{458,459,460} The Merlin Standard was withdrawn in April 2022⁴⁶¹ and replaced with guidance for providers on programmes such as the Work and Health Programme⁴⁶² and Restart.⁴⁶³

Findings from the interviews

In providing **contract management** services, prime providers reduce costs and increase efficiency for the commissioner.⁴⁶⁴ While some commissioner and provider interviewees valued this service as one lowering cost and administrative burdens for DWP,⁴⁶⁵ others in the same two stakeholder groups questioned this.⁴⁶⁶ The prime-sub model is attractive due to the belief that the resources of larger providers ensure **consistency and reliability** in service provision.⁴⁶⁷ This also presumes an ability to bear risk on behalf of both subcontractors and the commissioner.⁴⁶⁸

⁴⁴⁹ Carter (2018a).

⁴⁵⁰ Carter, E. (2021). 'More than marketised? Exploring the governance and accountability mechanisms at play in Social Impact Bonds', *Journal of Economic Policy Reform* 24:1, pp.78-94

⁴⁵¹ Egdell, V. et al. (2016). 'Third Sector Experiences of Work Programme Delivery', *Journal of Social Policy* 45(4), pp.729-746

⁴⁵² Egdell et al. (2016).

⁴⁵³ Egdell et al. (2016).

⁴⁵⁴ <https://www.gov.uk/government/publications/the-merlin-standard-guide-for-dwp-providers/the-merlin-standard-a-guide-for-dwp-providers>

⁴⁵⁵ Finn (2012).

⁴⁵⁶ Finn (2012).

⁴⁵⁷ Whitworth & Carter (2018).

⁴⁵⁸ Whitworth & Carter (2018).

⁴⁵⁹ Carter (2021).

⁴⁶⁰ Egdell et al. (2016).

⁴⁶¹ More information on the Merlin Standard is available: [here](#). Last accessed 17.05.22.

⁴⁶² More information on the Work and Health Programme last accessed 17.05.22: [Work and Health Programme including JETS provider guidance - GOV.UK \(www.gov.uk\)](#)

⁴⁶³ Last accessed 17.05.22: [Restart Scheme provider guidance - GOV.UK \(www.gov.uk\)](#)

⁴⁶⁴ C1, C2, C4, C5, C6, C7, C8, C10, E2, E4, E8, P3, P6, P7, P9, P12, P17.

⁴⁶⁵ C1, C4, P17.

⁴⁶⁶ C8, P6, P7.

⁴⁶⁷ C2, C3, C4, C6, C7, E9, E15, P3, P8, P9, P15.

⁴⁶⁸ C7, C6.

'So overall [...] it's a lot more manageable, standardised and then [still] linking in more of a local flavour [...] and we felt that rather than us having to manage all those individual relationships, if we manage one and allow them [to be] experts of the rest of the market [...] that would give us a more cost effective and value for money model.' - Commissioner interviewee

In the view of commissioner and expert interviewees, the prime model offers **flexibility**, as multiple subcontractors can be deployed for different aspects of a provision and to meet different service user needs.⁴⁶⁹ From another, as expressed by a mixture of interviewees across the three stakeholder groups, the complexity of supply chains might be seen to cause delays and inflexibility as **change takes time to trickle down** to all of a prime's suppliers.⁴⁷⁰ For this reason, some subcontractors would prefer to have direct communication with DWP so as to be better informed of expectations or any upcoming changes.⁴⁷¹ The large number of separate actors may also mean that stimulating innovation requires more effort and resources for effective coordination.⁴⁷²

In the view of interviewees across all three stakeholder groups, a prime-sub model facilitates **variety in the provision of services** and allows for small and medium organisations to use their niche or specialist skills as part of supply chains.⁴⁷³ In other views, also from interviewees from all three groups, the establishment of the prime-sub model **risks losing specialist knowledge** among these smaller providers.⁴⁷⁴ Not all of these interviewees believed this to be inevitable, but it was suggested by both provider and commissioner interviewees that primes have not been sufficiently incentivised to prioritise subcontractors with local and specialist knowledge in their supply chain.⁴⁷⁵

Some interviewees across the commissioners, providers and experts emphasised the importance of **due diligence on prime provider's claims** about their supply chain.⁴⁷⁶ It was also suggested DWP could design selection stages to reward bidders with a history of maintaining satisfied members of a supply chain.⁴⁷⁷ While the frequency of prime providers using specialist providers as 'bid candy' was contested, concerns were expressed around the **risks of sharing information with prime providers** at the bidding stage.⁴⁷⁸ Anecdotal accounts described staff of small and medium providers feeling they were misled about their role in a bid so a prime provider could demonstrate their specialist expertise or research without hiring them.⁴⁷⁹

⁴⁶⁹ C3, E12.

⁴⁷⁰ C3, C5, E2, P3, P13.

⁴⁷¹ P3, P13.

⁴⁷² C4, E2.

⁴⁷³ C3, C9, C10, E9, P4, P6, P11, P14, P15, P16, P19.

⁴⁷⁴ C3, C8, E1, E5, E12, P7, P7, P10, P11.

⁴⁷⁵ C3, C7, P11.

⁴⁷⁶ C4, E4, E9, E13, P3, P7, P17, P18.

⁴⁷⁷ P18.

⁴⁷⁸ E13, P1, P11, P16

⁴⁷⁹ P16, P18.

Overall, the power dynamic between prime and subcontractors appears to vary depending on the **relationships** built. Some provider interviewees reported that they were content in the supply chain of 'their' prime.⁴⁸⁰ Others described **unfair treatment** or maintained that this was a consistent concern for subcontractors.⁴⁸¹ The **Merlin Standard was not felt to be effective** by some expert and provider interviewees,⁴⁸² leaving a gap into which greater DWP intervention would be welcome to offer more systematic protection for subcontractors.⁴⁸³

'I haven't seen any of the analysis of supply chains from the Restart providers so far [...] I think it was seen as being pretty problematic under some of the former schemes and [...] there were real questions as to whether [the] Merlin [Standard] is fit for purpose. I just have not seen enough evidence that it can work, in order to give the Department or other interested parties confidence that a supplier who was being [mistreated] in the supply chain [...] really [has] reasonable recourse.' - Expert interviewee

Other problems experienced by subcontractors were not attributed to prime providers but rather to the model itself. While some primes were thought to **withhold funding from subcontractors**⁴⁸⁴ to protect themselves from financial risk,⁴⁸⁵ others might simply be unable to pass down referrals due to unexpected changes to the volume or demographic of service users.⁴⁸⁶

However, provider interviewees suggested the prime provider had a **responsibility to subsidise** specialist work when the prime could afford to,⁴⁸⁷ as well as to protect small subcontractors from retrospective extrapolation which could be crippling.⁴⁸⁸ Additionally, some small and medium providers in supply chains may be nervous about primes who also work as subcontractors in other areas⁴⁸⁹ or who are capable of taking work over from their supply chain if they wish to lower costs.⁴⁹⁰

Expert interviewees suggested the solidification of a prime-sub model has left DWP beholden to prime providers because continuation of employment services is a priority.⁴⁹¹ While this may be an exaggerated fear, provider interviewees worried the **lack of commissioning on the other tiers**, namely those for smaller organisations, forecast a semi-monopoly for the dominant prime providers.⁴⁹²

Interviewees suggested the following alternatives which could provide similar benefits to the prime-sub model:

⁴⁸⁰ P3, P4, P8, P10.

⁴⁸¹ C3, E13, P1, P6, P8, P9, P10, P11, P15, P16, P17, P18.

⁴⁸² E13, P1, P9

⁴⁸³ P1, P6.

⁴⁸⁴ C6, C8, E1, E4, P3, P4, P16, P20.

⁴⁸⁵ P20.

⁴⁸⁶ P16.

⁴⁸⁷ P3.

⁴⁸⁸ P20.

⁴⁸⁹ E1, P16, P17, P19.

⁴⁹⁰ P16, P17.

⁴⁹¹ E6.

⁴⁹² P16, P20.

- establishing a DWP case management IT system for all providers to use during contracts (allowing for shared benefits from economies of scale usually gained from large Contract Package Areas or CPAs) – this is in addition to the Provider Referrals and Payments (PraP) system for referrals and claiming payments for outcomes that is already in place: the new system was suggested for case management during the duration of the contract so small providers were not disadvantaged with IT costs, and primes did not have to rely on for economy of scale⁴⁹³
- moving to a licensing model in which providers are accredited and then free to move into and out of the provider market at any time, subject to performance standards.⁴⁹⁴

2.6. Contract Package Areas may be too large

Contract Package Areas (CPAs) relate to the geographical division of service provision in specific contracts. The **UK** WP had 18 CPAs, each consisting of two or three separate contracts, with a total number of 40 contracts awarded.⁴⁹⁵ Claimants were randomly allocated to one of the primes in a CPA if they had not found work within a specified period. Competition between primes in the Work Programme operated at the CPA level. Minimum Performance Levels were set at the level of the CPA, requiring primes to deliver results at 10% higher than the estimated (regional) non-intervention rate.⁴⁹⁶ Primes who did not achieve the Minimum Performance Level for each CPA could lose their contracts. A second competitive mechanism at the CPA level was the ‘market share shift’, where higher performing primes received a greater proportion of referrals.⁴⁹⁷ While we cannot know the net impact of any one provider, we can know how they are performing relative to their competitors within the CPA. An evaluation found the market share shift was perceived by DWP staff to have a limited impact due to the relatively small proportion of referrals (5%) moved to higher performing providers in the CPA.⁴⁹⁸ Other programmes differ in the design of CPAs. The Work and Health Programme had 6 CPAs and the Restart Scheme has 12.

CPAs are intended to reflect and respond to spatial variation across the **UK** in the functioning of local labour markets.⁴⁹⁹ Providers bidding for contracts as part of the Work Programme were expected to adapt pricing to reflect the ease of achieving job outcomes within the region.⁵⁰⁰ There was no variation in pricing below the CPA level of geographical division.

⁴⁹³ C8.

⁴⁹⁴ E3. Also: Australia Productivity Commission, 2002, “Independent Review of the Job Network”, Report No.21, AusInfo, Canberra, p.xx., p.11.1, p.11.4-11.6, p.11.7. Please also see Figure 2 for further explanation of a licensing model.

⁴⁹⁵ Kay & Marlow (2020).

⁴⁹⁶ Carter (2018a).

⁴⁹⁷ Kay & Marlow (2020).

⁴⁹⁸ Foster et al (2014). Work Programme Evaluation: Operation of the commissioning model, finance and programme delivery.

See: [Work-programme-evaluation-operation-of-the-commissioning-model-finance-and-programme-delivery-summary.pdf](#)

⁴⁹⁹ Carter (2018a).

⁵⁰⁰ Carter (2018a).

The rationale behind introducing large CPAs in the WP was to enable large, private sector organisations to participate in the quasi-market.⁵⁰¹ Having a smaller number of large CPAs reduces complexity and lowers transaction costs for DWP. However, some have argued that CPAs are too large and fail to capture localised differences in labour market conditions and provider performance.⁵⁰² Carter⁵⁰³ finds little variation in the performance of primes operating in the same CPA, masking considerable variation at the level of local authority areas. This results in a lack of transparency and accountability for localised differences in provider performance:

*'Because the contractual agreements are sited at the level of CPAs (...) local authorities [areas] with unacceptably low job outcome performance are contractually invisible since they are diluted, concealed and deemed irrelevant within the CPA aggregate performance figures.'*⁵⁰⁴

Setting competition at the CPA level might encourage providers to focus their efforts on better performing local areas, neglecting areas where the local labour market is more challenging (a regional version of creaming and parking).⁵⁰⁵

Questions have been raised about the degree of variability in service provision across CPAs, in light of the black box approach to contract design in the WP (see section 2.3). Focusing on health-related provision, Ceolta-Smith et al.⁵⁰⁶ highlight considerable variation across CPAs and across providers, raising concerns about inequitable service provision in terms of quality and content.

Findings from the interviews

Interviewees, across commissioners, experts and providers, expressed the view that CPAs **should be smaller** than they are now and have been in the recent past.⁵⁰⁷ An alternative strategy could see larger CPAs for mainstream programmes and smaller for specialist programmes which require greater integration with local economic zones and specialist support services.⁵⁰⁸

In terms of effect on the provider market, creating **fewer and larger CPAs could limit competition** and exacerbate the semi-monopoly of prime providers mentioned above.⁵⁰⁹ Previous programmes in the UK and elsewhere have attempted to address this issue by issuing contracts to more than one provider in each CPA, however service user needs and lack of information limited any competitive drivers.⁵¹⁰ It may be that the concentration of contracts among fewer providers can be remedied by **promoting diverse and locally integrated supply chains.**⁵¹¹

⁵⁰¹ Carter (2018a).

⁵⁰² Carter (2018a).

⁵⁰³ Carter (2018a).

⁵⁰⁴ Carter (2018a).

⁵⁰⁵ Carter (2018a).

⁵⁰⁶ Ceolta-Smith et al. (2015).

⁵⁰⁷ C8, C10, E2, E4, E13, P1, P2, P9, P12, P13, P14.

⁵⁰⁸ E4, P1.

⁵⁰⁹ C1, P8, P10, P12, P13, P18.

⁵¹⁰ C5, E9, E13, P10, P18.

⁵¹¹ P18.

'A criticism which was made of [...] fairly large contract areas is that the prime contractors are insufficiently attuned to local conditions. So, they don't have the relationships with local organisations that might be important players locally, and there's no reason why they couldn't be. So, a prime contractor can work with the range of subcontractors they feel appropriate, and that [...] could be somebody's a specialist, because they deal with a particular type of disability or other need, or they could be a specialist, because they happen to know everything which goes on within a particular town.' - Commissioner interviewee

Some commissioner and provider interviewees suggested that fewer CPAs were desirable to promote **simplicity and consistency** across geographies.⁵¹² However, the drawing of boundaries, and subsequent size of CPAs, also need to consider the four nations' borders⁵¹³ and the complexity of devolved budgets.⁵¹⁴ Some provider interviewees even argued that consistency and efficacy in services was hampered by the division of a national provision among different providers while still acknowledging the dangers of monopolies.⁵¹⁵

It was also noted that fewer and larger CPAs also result in **high value contracts**.⁵¹⁶ But contracts must still be well-funded, avoiding heavy discounting, to avoid resources being spread thinly over a large geographical area.⁵¹⁷ These high value CPAs will be attractive to bidders and so likely increase competition, driving a healthy market of providers.⁵¹⁸ However, the maintenance of large-scale, high value contracts **excludes small and mid-size providers from bidding**, keeping contracts among a small number of prime providers.⁵¹⁹ But this effect may not be inevitable; consortiums of small providers could combine their local expertise to bid for large CPAs.⁵²⁰

Ultimately, the design of large CPAs depends on the logic of **economics of scale**. While some provider interviewees believed this was an accurate assessment,⁵²¹ commissioner and expert interviewees disagreed that economies of scale were at play to a significant degree.⁵²² Interviewees across the three stakeholder groups agreed that the existence of fewer CPAs **lowers administrative commercial costs and management burdens** for the commissioner.⁵²³ However, if a programme is less effective due to the design of geographical divisions, these savings may be undermined.

For example, some interviewees, across the three stakeholder groups, contended that the geographical boundaries of CPAs have been and continue to be arbitrary

⁵¹² C1, C9, P4, P6, P11.

⁵¹³ C1, E1, P14, P16.

⁵¹⁴ E12, P13.

⁵¹⁵ P6.

⁵¹⁶ C1, C2, C11, P7, P12, P13, P14, P20.

⁵¹⁷ P7, P20.

⁵¹⁸ C11, P18.

⁵¹⁹ P8, P10, P13.

⁵²⁰ P12.

⁵²¹ P8, P9, P11, P20.

⁵²² C8, E1.

⁵²³ C8, E1, E13, P7, P13, P14, P18.

and **not always well aligned with functional economic zones** or regional labour markets.⁵²⁴ Advocates for smaller CPAs suggested these were better suited to tailoring a programme to local resources and restrictions.⁵²⁵ This includes **variation between urban and rural areas** or between wealthy and less prosperous areas, all contained within large CPAs.⁵²⁶

While large CPAs involve transferring a great deal of administrative and operational responsibility, it is also possible that less risk is held by the commissioner if smaller CPAs are created which **lowers the impact of scenarios such as delivery failure**.⁵²⁷

Very little evidence was found in relation to health assessments (see Box 6).

Box 6. Health assessments: CPAs

In health assessment contracts lots replace the notion of CPAs. It seems monopolies on delivery have been less of a concern, with contracting out to **two suppliers being considered 'diverse' for this market**.⁵²⁸ Additionally, as the devolved Scottish government will be taking over PIP in Scotland, the commissioner's hands may be tied on some aspects of geographical division.⁵²⁹ While a smaller number of suppliers might be more appropriate in health assessments, compared to employment services, there have been issues in health due to the small number of primes.⁵³⁰

2.7. Length of contracts ideally falls between five to seven years

Several advantages of longer contracts have been noted in the literature. Longer contracts give greater scope for providers to invest in services, resources and human capital,⁵³¹ to develop their capacity and expertise⁵³² and to invest in innovation.⁵³³ For providers, longer-term contracts are necessary to recoup their investments under PbR contracts.⁵³⁴ Longer contracts may reduce transaction costs for commissioners since the tendering process is less frequent. A longer-term perspective is also important in relation to the provision of employment services, where gains may take some time to emerge, for instance after poor performers have been excluded from the market.⁵³⁵ However, there is a balancing act since in other respects shorter contracts offer advantages. Shorter contracts may improve competition/performance

⁵²⁴ C8, E1, E5, E13, P1, P12, P16, P17, P19.

⁵²⁵ C7, C10, E1, E2, E4, E5, E9, E13, E15, P1, P3, P4, P5, P7, P8, P10, P13, P14, P17, P19.

⁵²⁶ E2, P17.

⁵²⁷ C11, P11, P14.

⁵²⁸ C3.

⁵²⁹ C3.

⁵³⁰ DWP representative.

⁵³¹ Finn, D. (2012). 'Subcontracting in Public Employment Services: The Design and Delivery of "Outcome Based" and "Black Box" Contracts', European Commission

⁵³² Kay, J & Marlow, S., (2020). 'The Work Programme: A Quantitative Impact Assessment', UK Department for Work and Pensions

⁵³³ Kay & Marlow (2020).

⁵³⁴ Struyven, L. & Steurs, G., (2004). 'Quasi-market reforms in employment and training services: first experiences and evaluation results' in Evaluation of Systems and Programmes: Third report on vocational training research in Europe, Luxembourg: Office for Official Publications of the European Communities, (Cedefop Reference Series 57)

⁵³⁵ Stephan, G. (2016). Public or private job placement services - Are private ones more effective? IZA World of Labour. 285, p1-10.

because it is easier for new entrants to join the market⁵³⁶ and for poor performers to be excluded.⁵³⁷ Shorter contracts enable the commissioners to review and adapt contract terms.⁵³⁸ However, provisions can be built into longer-term contracts allowing commissioners and providers to renegotiate terms.⁵³⁹ There is also greater danger of designing contracts that are not well matched to the realities of the labour market with longer contract durations.⁵⁴⁰

Findings from the interviews

Interviewees, across all three groups of stakeholders, stated that **five years** was the optimal length of a contract for employment services.⁵⁴¹ There was variation, with some suggesting three years⁵⁴² or ranges of three to five years⁵⁴³ or five to seven years.⁵⁴⁴ Others suggested seven years as ideal, if an extension on five years.⁵⁴⁵ Some expert interviewees suggested one-year contracts were optimal for employment services.⁵⁴⁶ Others thought contracts should be shorter than they have been,⁵⁴⁷ although **many more believed increasing the typical length of contracts would be beneficial.**⁵⁴⁸

The time taken in **start-up and wind down** was commonly referenced as the reason for designing longer contracts, as these periods reduce the amount of time in which a contract is producing maximum outcomes.⁵⁴⁹ Furthermore, too short a contract **discourages investment** from the provider and can make securing premises or staff very challenging.⁵⁵⁰ Additionally, short contracts result in more frequent tender, bidding and negotiation stages, which **drives up commercial transaction costs** for DWP.⁵⁵¹

However, if contracts are too long this will ultimately **limit market entry** and reduce the benefit of competition by locking the commissioner in to contracts with a few providers for an extended period of time.⁵⁵² This could be problematic if DWP, or indeed providers, are **restricted in responding to political and economic change** for which the contract or programme does not account.⁵⁵³

Longer contracts also increase the importance of **robust mechanisms for monitoring** and managing poor performance.⁵⁵⁴ The impact of low service quality or

⁵³⁶ Struyven & Steurs (2004).

⁵³⁷ Finn, D. (2011). "Subcontracting in public employment services: review of research findings and literature on recent trends and business models", European Commission

⁵³⁸ Finn (2011).

⁵³⁹ Finn (2011).

⁵⁴⁰ Langenbacher & Vodopivec (2022).

⁵⁴¹ C1, C4, C6, C8, C10, E4, E11, E12, P2, P3, P6, P8, P9, P10, P11, P13, P14, P19.

⁵⁴² C7, C11, E1, E7, P1, P12, P16, P17, P18, P19.

⁵⁴³ P4, P5, P12.

⁵⁴⁴ P20.

⁵⁴⁵ C1, C4, C10, E6, P7, P12.

⁵⁴⁶ E7.

⁵⁴⁷ C4, C6, E7, P9, P16.

⁵⁴⁸ C4, C8, C9, C10, E2, E3, E5, E8, E9, E12, E13, P5, P12, P13, P17, P18, P19.

⁵⁴⁹ C4, C7, C8, C9, C10, C11, E2, E3, P1, P2, P4, P10, P11, P14, P17.

⁵⁵⁰ C4, C6, E8, E9, E12, P2, P6, P8, P7, P12, P17, P18, P19.

⁵⁵¹ E8, E12.

⁵⁵² C4, C7, E2, E13, P11, P12, P18.

⁵⁵³ C2, C3, C8, E11, P12, P13, P16.

⁵⁵⁴ C10, E5, E7, P18.

poor performance over a long period of time could have more severe long-term implications than issues which occurred over a one-to-three-year period.

Indeed, programmes of different lengths, which may have been chosen for a specialist type of intervention, will require different types of support from the commissioner.⁵⁵⁵ These types of calculations will likely be part of the assessment of trade-offs when designing a contract and should be considered in combination with the other features discussed above, such as payment mechanisms and CPAs.

Evidence in relation to health assessments was very limited (see Box 7).

Box 7. Health assessments: length of contracts

For health assessments, longer contracts were much more typical, with some interviewees even suggesting ten years as an optimal length, citing a need for stability and to ensure the retention of skilled health professionals carrying out the assessments.⁵⁵⁶

⁵⁵⁵ C3, E1.

⁵⁵⁶ C1, C2.

3. Market competition

Key findings on market competition

- Selection criteria often combined quality and price. The examples identified show the weight of the price ranges between 20% and 40%. The quality aspects address issues usually aiming to capture relevant experience and expertise.
- There is little evidence of how different types of contracts impact market competition in the short- & long-term. However, there is some evidence on the effect of contracting out on market competition: this evidence points to some providers being ‘pushed out’ or forced to adapt to secure contracts and deliver services. The dominance of prime contractors in the market has now been well established.
- Trade-offs between optimising short-term and long-term outcomes has been a process of ‘trial and error’. One of the challenges is finding an optimal balance between upfront fees and outcome payments to create incentives for providers and savings for commissioners.

3.1. Selection criteria often combined quality and price

Overall, there was some evidence found on selection criteria and scoring mechanisms used in commissioning employment services within the literature reviewed. Country-specific examples are listed in Table 2.

Table 2. Selection criteria

Country	Criterion (quality or experience)	Weight	Price
Australia	Past performance	30%	No information
	Achieving outcomes	30%	
	Meeting employers’ needs	30%	
	Governance	10%	
Belgium	Description of the implementation methods	50%	20%
	Expertise: past experience and competence of case workers	20%	
	Location of service provision (accessibility by public transport)	10%	
Canada	(Unspecified) qualitative criteria	70%	22%
	Final stage interview	8%	
Ireland	Implementation	18%	40%
	Delivery	30%	
	Contract management and governance	12%	
United States	Proposed services and client journeys	30%	No information
	Providers experience and strength of partnerships	20%	
	Other (unspecified) qualitative aspects	50%	

Source: Cockx & Baert (2015),⁵⁵⁷ Langenbucher & Vodopivec (2022).

In **Denmark** commissioning was decentralised, and regional PES were given freedom to decide the types of service, target groups, contract forms, and payment models, as long as at least 10% of the unemployed users were in services provided by non-public providers.⁵⁵⁸

Germany included quality (in addition to price) in their decision to select contractors: private providers, and most of their services, had to be approved according to criteria such as capacity and personnel; certification companies themselves also had to be approved by the national accreditation body, Deutsche Akkreditierungsstelle.^{559,560}

The **Netherlands** varied the selection criteria according to the needs of particular groups, placing greater emphasis on the professional competence of contractors, rather than price, for those working with more disadvantaged groups.⁵⁶¹

As evident from Table 2, **price** is a common criterion used. However, it has both advantages and disadvantages. On the one hand, it helps to reduce cost. Proponents of contracting out argue that introducing competition for providers incentivises them to better use resources and thus increases efficiency in service delivery: competition among providers confines the provision of a service to those best suited for it and transfers some cost savings to the buyers.⁵⁶² However, incomplete (or less detailed) contracts give providers scope for cutting costs and quality.⁵⁶³ On the other hand, price may have negative impact on quality, as providers remove expensive elements to become more competitive.^{564,565} Empirical evidence from the Netherlands helps to illustrate this point. The fall in price and focus on short term results for outcome payments led providers to remove long-term training for the unemployed from their offer.⁵⁶⁶

By the same token, **relevant, proven experience and expertise** are frequently used as criteria but while their intent is to ensure service quality, these criteria are far from perfect. They can also inhibit access for new providers; they may be unreliable, as outcomes achieved may depend on a multitude of factors; they are also hard to use to objectively compare between providers with different capacities to generate outcomes.^{567,568} Another example from the Netherlands illustrates how the quality requirements can be a barrier to gain entry to the market. In a Dutch tendering process, new providers had much lower chances: the average success rate to secure the contract was 30% but this varied between new entrants and more established

⁵⁵⁷ Cockx, B., & Baert, S. (2015). Contracting Out Mandatory Counselling and Training for Long-Term Unemployed. Private For-Profit or Non-Profit, or Keep it Public? IZA DP No. 9459

⁵⁵⁸ Jantz, B. et al. (2018). 'Marketization and Varieties of Accountability Relationships in Employment Services: Comparing Denmark, Germany, and Great Britain', *Administration and Society* 50(3), pp.321-345

⁵⁵⁹ <https://www.dakks.de/en/home-en.html>

⁵⁶⁰ Jantz et al. (2018).

⁵⁶¹ Finn (2012).

⁵⁶² Andersson et al. (2019).

⁵⁶³ Stephan (2016).

⁵⁶⁴ Jantz et al. (2018)

⁵⁶⁵ OECD (2006); in: Lindsay & McQuaid (2009).

⁵⁶⁶ Finn (2011).

⁵⁶⁷ Struyven & Steurs (2004).

⁵⁶⁸ Stephan (2016).

providers (with 14% and 43% success rates respectively).⁵⁶⁹ Barriers to market entry (such as red tape or high costs) need to be low because risks associated with PbR contracts may discourage some providers from the tendering process and limit competition.⁵⁷⁰

The selection process is different in voucher systems, where jobseekers obtain a voucher which entitles them to receive employment support and it is jobseekers (and not the commissioners) who select their providers.^{571,572} In Germany, providers can enter and exit without going through a lengthy tendering process or entering into long-term commitments on set conditions. This creates easier access for small providers (including small- and medium-sized enterprises, SMEs) but poses challenges for brand recognition.⁵⁷³ However, users often lack information to make informed choices about selecting the right provider for them. As such, implementing schemes that include a degree of user choice is challenging.⁵⁷⁴ While different forms of quality recognition were identified in the reviewed literature, these were contested:

- Star Ratings in Australia⁵⁷⁵ – (see section 4.1) these Star Ratings were designed to aid service users in choosing between providers based on performance and satisfaction matrices.⁵⁷⁶
- Certification in Germany – where PES lose direct influence over providers: providers are accountable to (private) certification companies.⁵⁷⁷

There is little evidence to support practitioners to determine what market-shaping activities they should undertake and under which circumstances.⁵⁷⁸

Findings from the interviews

The major topic of concern regarding the selection process was **an excessive emphasis on price** in previous rounds of bidding.⁵⁷⁹ Some interviewees did suggest this has improved⁵⁸⁰ but, generally, many interviewees were critical of the short- and long-term impact of damaging practices around discounting.⁵⁸¹

'I think there was a concern, at one point, around the price scoring, around the amount of discount which [...] if they hadn't [...] changed that [for Restart], it probably would have meant [...] kind of a race to the bottom, so we were really pleased when that was changed.' - Provider interviewee

⁵⁶⁹ Struyven & Steurs (2004).

⁵⁷⁰ Stephan (2016).

⁵⁷¹ Bruttel, O., (2005). 'Delivering active labour market policy through vouchers: experiences with training vouchers in Germany', *International Review of Administrative Sciences* 71(3), pp.391-404

⁵⁷² Jantz et al. (2018)

⁵⁷³ Bruttel (2005).

⁵⁷⁴ Conolly et al., (2010).

⁵⁷⁵ Struyven & Steurs (2004).

⁵⁷⁶ O'Halloran, D. et al, (2022). 'The development of the Australian Unemployed Workers Union Rating Scale (AUWURS) of employment service providers', *Australian Journal of Public Administration*, (Early View), p.3.

⁵⁷⁷ Jantz et al. (2018).

⁵⁷⁸ Carey, G. et al. (2020). 'Quasi-market shaping, stewarding and steering in personalization: the need for practice-orientated empirical evidence', *Policy Design and Practice* 3(1), pp.30-44

⁵⁷⁹ C1, C6, C8, E4, E12, E15, P2, P18, P19, P20.

⁵⁸⁰ P9, P14.

⁵⁸¹ E4, E13, P2, P15, P18, P19, P20.

'It's just really important that we don't [buy] this stuff too cheaply, because the people who are going to suffer are the people that we're trying to support. If the service levels and the quality of the service ends up getting cut in order to achieve that cheaper price, it doesn't serve anybody well.' - Commissioner interviewee

Additional insights were offered by representatives from DWP who commented on the price criterion and noted that:

'For FAS [Financial Assistance Scheme] we are using the UK Government "Bid Evaluation guidance note"⁵⁸² which advocates using a Price per Quality Point [PQP] metric.' - DWP representative

Indeed, PQP is a novel approach and it has not been discussed in the reviewed literature. A PQP is calculated by dividing the bid price by the quality score to give an output price per quality point: the lowest PQP is the most economically advantageous tender.⁵⁸³

The selection criteria are known to seek indicators of strong performance.⁵⁸⁴ However, there were calls for a greater emphasis on **assessing track record** – even outside this field (i.e. employment support) – and rewarding this in selection.⁵⁸⁵

Some interviewees across the commissioners, providers and experts alike regretted **the exclusion of smaller providers based on reserve or turnover requirements**⁵⁸⁶ or the resources needed to produce bids.⁵⁸⁷ In addition, some expressed concerns about the treatment of smaller providers by primes during the bid stage (discussed in section 2.5).⁵⁸⁸ One interviewee suggested that contracts with heavy weighting to outcome payments only attracts large providers to bids.⁵⁸⁹

A past tendency for bids to **over-promise on performance** was noted.⁵⁹⁰ One interviewee suggested that DWP management of performance offers during the Restart bidding phase was good example of course-correction.⁵⁹¹ This greater intervention by the commissioner used existing experience, knowledge and research to set parameters for the expected pricing and feasible performance levels for Restart bids. This then allowed both for sharing evidence with bidders on a realistic performance offer and limiting the attraction of low-cost bids.

⁵⁸² As of 15/07/2022:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/987130/Bid_evaluation_guidance_note_May_2021.pdf

⁵⁸³ As of 15/07/2022:

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/987130/Bid_evaluation_guidance_note_May_2021.pdf

⁵⁸⁴ C2, E7, P5, P6, P16, P18, P20.

⁵⁸⁵ P16, P18.

⁵⁸⁶ C8, E4, E5, P1, P4.

⁵⁸⁷ P1, P7, P10, P12.

⁵⁸⁸ P1, P16, P17.

⁵⁸⁹ E4.

⁵⁹⁰ C8, E1, E4, P3, P16.

⁵⁹¹ P9.

There was some reporting of previous **provider involvement with developing the tender**, largely in market engagement.⁵⁹² Nonetheless, more collaboration between DWP and providers on the programme, contract and selection question was recommended.⁵⁹³ In an example of good practice, some interviewees praised DWP's previous use of 'local actors' to formulate qualitative selection questions which would limit surface-level answers from bidders.⁵⁹⁴ This involved consulting smaller organisations and local, supporting, actors in the employment services industry, while formulating questions that should be asked to prime providers in their bid.

While, overall, provider interviewees reported positive relations with DWP, some noted a few points of dissatisfaction. Some provider interviewees felt there was a gap between the commissioning strategy they discussed with DWP staff (mostly in the policy team) in informal conversations and the strategy that was seen in the tender announcement – in terms of the scope for innovation, minimum service requirements and the weight of price criterion in the selection process (attributed by interviewees to the commercial team).⁵⁹⁵ Others reported feeling that DWP due diligence on their bid found inaccuracies which they believed could have been resolved or disproven, if dialogue was facilitated to a greater extent in this stage of the process.⁵⁹⁶

In general, some providers thought the outcomes of selection processes were unpredictable.⁵⁹⁷ This was thought to be a negative characteristic as existing connections or embeddedness in local areas were not thought to be taken into account appropriately, meaning newcomers to an area would win a competition and start from scratch. Some provider interviewees also felt that confident bids with overly optimistic claims were rewarded over those which expressed uncertainty or caution on volumes or performance.⁵⁹⁸ Others felt that local geographical embeddedness not sufficiently rewarded, also resulting in a waste of the provider's local investment from previous programmes.⁵⁹⁹

Suggestions for changes to the selection process included:

- continuing to collect feedback from bidders to further inform future competitions⁶⁰⁰
- choosing providers through a licensing system to satisfy the need for proven experience and expertise⁶⁰¹
- shifting to a social partnership model in which private providers, government and relevant third sector organisations voluntarily work together in the design and implementation of commissioning public services⁶⁰²

⁵⁹² C5, C11, P4, P6,

⁵⁹³ C8, E7, P4, P6, P8, P10, P11, P13.

⁵⁹⁴ E7, P7.

⁵⁹⁵ P6, P7, P8.

⁵⁹⁶ P2.

⁵⁹⁷ P2, P6, P16.

⁵⁹⁸ P2, C8,

⁵⁹⁹ P2, P5, P11, P16, P19

⁶⁰⁰ C1, E9.

⁶⁰¹ E1.

⁶⁰² E3. See also: Welsh Government (2021). 'Draft Social Partnership and Public Procurement (Wales) Bill'; Scottish Government (2018). 'Report – to date – of the Strategic Public Social Partnership (PSP) Model in Scotland.'

- encouraging non-profits and charities to include margins to provide for re-investment or risk-management in their financial bids even when they are not-for-profit organisations.⁶⁰³

It appears that, despite some critique, the methods of selection model are acceptable to some interviewees, with the caveat that the current model was not thought to promote innovation or facilitate a 'market disruptor'.⁶⁰⁴

3.2. There is little evidence of how different types of contracts impact market competition in the short- & long-term

Beyond discussing the selection criteria in contracting out employment services (section 3.1), little empirical evidence was found in the reviewed literature on how the type of contract affects market competition, or even, more broadly, on how the performance could be improved in the short- and long-term.⁶⁰⁵

Experiences documented in the literature include examples from a few countries but these focus on the effect of contracting out on competition (rather than the effect of a specific type of contracts):

- In the **UK**, DWP sought to animate ongoing competition by using 'market share shift' to reallocate a portion of the caseload to better performing providers. In practice, this has been a weak tool.⁶⁰⁶
- Rapid use of contracting out in the **Netherlands** led to the emergence of a 'pluriform market',⁶⁰⁷ referring to the complexity of an intensified tender mechanism with many purchasers and limited co-ordination between them, with no evidence of increased efficiency nor equity.⁶⁰⁸
- **Denmark** first saw a rapid expansion and then contraction of contracted out provision due to poorly designed financial reward systems.⁶⁰⁹

Former state monopoly systems transitioned in some countries to one characterized by the co-existence of the public and private actors: e.g. Germany,⁶¹⁰ France⁶¹¹ and Israel.⁶¹² While the situation in each country has likely changed over time, snapshots were captured in the reviewed literature that provide some insight into the shape of the market at the time the studies were carried out:

- In **France**, in 2007-2008, there were 11 private providers including (i) job agencies supplying temporary workers, (ii) consultancies specialised in the placement of

⁶⁰³ C4.

⁶⁰⁴ E1, P6, P7, P8, P9, P14.

⁶⁰⁵ Carey et al. (2020).

⁶⁰⁶ OECD (2014).

⁶⁰⁷ OECD (2006); in: Lindsay & McQuaid (2009).

⁶⁰⁸ Struyven & Steurs (2004).

⁶⁰⁹ Lindsay & McQuaid (2009).

⁶¹⁰ Ayaita et al. (2021).

⁶¹¹ Behaghel et al. (2012).

⁶¹² Benjamin (2020).

workers after mass layoffs, and (iii) international firms specialised in the placement of foreign jobseekers.⁶¹³

- The number of further education providers in **Germany** in 2002 was estimated between 5,000 and 30,000 providers, but PES still had a strong position and exerted considerable influence over the quality and price on the market.⁶¹⁴
- Looking at the commissioners in 2002, **Australia** and the **Netherlands** were the two extremes: Australia had only one purchaser, while in the Netherlands there were a number of buyers and providers.⁶¹⁵ There were 44 large for-profit providers operating in Australia and thousands of providers in the Netherlands.⁶¹⁶

Some studies focused on the impact of contracting out on specific market players (i.e. not-for-profit organisations). According to an empirical, longitudinal study, the **WP in the UK** changed the landscape contracting the share of the market held by (smaller) third sector organisations and expanding the large private sector presence in the delivery of employment support leading to a loss of diversity at the market.⁶¹⁷ However, the third sector organisations examined in the study were able to adapt and find alternative funding. Those which stayed in the market had to work in a similar way to private sector providers. **In Germany**, some not-for-profit providers were ‘pushed out’ but many others were able to secure contracts, improve performance, develop and deliver services in line with their values and missions.⁶¹⁸ Some authors point to a risk of declining diversity of providers leading to the homogenisation of employment services.^{619,620}

Findings from the interviews

Commissioner, expert and provider interviewees characterised the landscape of the provider market as one of ‘**feast and famine**’.⁶²¹ This refers to the risk that after a long programme, with substantial funding, there will be a ‘cliff edge’⁶²² for providers until the next is commissioned, assuming they are even successful in the next round of bids. For those that are successful, there remains a need to scale up or down unpredictably in response to the programmes commissioned by DWP.⁶²³ This lack of long-term stability restricts planning and worsens financial risk for providers, sometimes to the extent that smaller contractors cannot afford to participate.⁶²⁴

‘You saw these massive changes in the contracting framework [...] like, you had the Flexible New Deal: it was then completely scrapped, and all the supply chains torn up, and you moved onto the Work Programme, and then the Work Programme ends, you do feel, well, “what happens to all of these providers?”. I think I’ve seen surveys where they ask providers, [...] “what’s going to happen if

⁶¹³ Behaghel et al. (2012).

⁶¹⁴ Bruttel (2005).

⁶¹⁵ Struyven & Steurs (2004).

⁶¹⁶ Langenbucher & Vodopivec (2022).

⁶¹⁷ Egdell et al. (2016).

⁶¹⁸ Finn (2011).

⁶¹⁹ Egdell et al. (2016).

⁶²⁰ Fuertes, V., Jantz, B., Klenk, T. and McQuaid, R. (2014). ‘Between cooperation and competition: the organisation of employment service delivery in the UK and Germany’, *International Journal of Social Welfare*, 23: S1, S71–S86.

⁶²¹ C1, C4, C10, E2, E4, E12, P2, P16, P17, P18, P20.

⁶²² P2, P17.

⁶²³ C1, C4, P12, P13.

⁶²⁴ C7, C8, E8, E12, E5, P2, P4, P7, P11, P16.

you don't get Work Programme funding, or what's going to happen when the Work Programme ends?". And obviously, they always say, we might close, but, then, we don't know that that's true, and I think there's a tendency for everyone to always give that pessimistic view.' - Expert interviewee

Furthermore, a combination of large-scale PbR contracts has firmly established the **dominance of prime contractors** in the market.⁶²⁵ These may be corporate entities who also operate beyond this market, allowing overall revenue and reserves to offset the risk or fluctuations of PbR and the 'feast-famine' effect.

These knock-on effects are also thought to **stifle entry to and exit from the market**.⁶²⁶

'I think, potentially, because of the volume of risk and the subject matter expertise that people build up over a period, your market is constrained. There's a high barrier to entry, for new entrants, and the provisions [...] of the contract will, in part, be driving that' - Provider interviewee

Interviewees across all three groups were concerned with the prospect of small providers with specialist skill sets being pushed out or lost.⁶²⁷ Even meeting the requirements for putting bids together can be beyond the resources of smaller providers.⁶²⁸ Some provider interviewees also questioned the efficiency and fairness of the logistics or timing of the tender and bidding stages.⁶²⁹

Other factors in instability for this market include the **risk of variation in volumes** and the production of contracts based on **discounted prices and unrealistic performance**.⁶³⁰ This latter phenomenon was described as creating a 'vicious cycle of poor results' which has a negative impact for commissioner, clients and providers.⁶³¹

Alongside well-designed contracts, some commissioner and provider interviewees emphasised the importance of **maintaining good relationships** between commissioner and providers for the sustainability of the market long-term.⁶³² For example, the commissioner needs to **communicate clearly with the market** about its short- and long-term goals.⁶³³ Indeed, due to uncertainty around commissioner strategy or behaviour, some providers may have decided to price risk into their bids.⁶³⁴ This is a cost that could be avoided.

Suggestions for sustaining a healthy and competitive market included:

⁶²⁵ E4, E13, P5, P10, P12.

⁶²⁶ C4, E4, E7, E8, E13, E15, P6, P7, P8, P11, P15, P18.

⁶²⁷ C7, C8, E8, E12, E5, P2, P4, P7, P11, P16.

⁶²⁸ E8, E13, E15, P7, P11, P15.

⁶²⁹ P1.

⁶³⁰ E1, E4, E8, E14, P8, P13.

⁶³¹ E1.

⁶³² C9, P1, P14.

⁶³³ C1, C9, P4, P5, P13, P17, P18.

⁶³⁴ C10.

- running competition within CPAs⁶³⁵
- outsourcing Jobcentre Plus⁶³⁶
- changing to a licensing model⁶³⁷
- continuing to include local authorities among contractors⁶³⁸
- creating greater continuity of geographies for selected providers⁶³⁹
- building financial insight to understand drivers of market behaviour.⁶⁴⁰

While some provider interviewees suggested market share shift mechanisms⁶⁴¹ for the maintenance of competition within the market, commissioners critiqued prior use.⁶⁴²

3.3. Trade-offs between optimising short-term and long-term outcomes has been a process of ‘trial and error’

As noted above, price is an important and delicate element in the market competition. Commissioners need to set the price in such a way as to:

‘attract providers to participate in tenders, to ensure that payment models create the right incentives to achieve the desired outcomes, as well as to formulate performance expectations to achieve value for money.’⁶⁴³

While the ways in which prices are set (and their values) may vary and are usually not made public, there is some information on the balance between upfront fees and outcome payments which is crucial to ensuring that contracting out employment services brings mutual benefits for commissioners and providers.^{644,645,646} The process of contracting out has been through ‘trial and error’ in most countries⁶⁴⁷ and continuous learning and adjustment constitute the path forward.

The most recent review and analysis of PbR in contracting out employment services⁶⁴⁸ identifies a range of practices from different countries on setting the outcome payments and their overall weighting in the contracts (Table 3).

⁶³⁵ C11.

⁶³⁶ P18.

⁶³⁷ P18.

⁶³⁸ P17.

⁶³⁹ E8.

⁶⁴⁰ C1.

⁶⁴¹ P18.

⁶⁴² C4.

⁶⁴³ Langenbucher & Vodopivec (2022).

⁶⁴⁴ Bach-Mortensen, A. and Barlow, J. (2021). Outsourced austerity or improved services? A systematic review and thematic synthesis of the experiences of social care providers and commissioners in quasi-markets. *Social Science & Medicine*. 276.

⁶⁴⁵ van Gestel, N. et al. (2019). From quasi-markets to public-private networks: Employers' engagement in public employment services. *Policy Administration*. 53, pp.434-448.

⁶⁴⁶ Struyven & Steurs (2004).

⁶⁴⁷ Behaghel, L. et al. (2012). 'Private and Public Provision of Counselling to Job-Seekers: Evidence from a Large Controlled Experiment', IZA Discussion Paper No.6518

⁶⁴⁸ Langenbucher & Vodopivec (2022).

Table 3. Structure of outcomes payments and weights of outcome payments as % of the total

Country	Outcome payments	Weight of outcome payments*
Australia (Jobactive – Stream B)	Payable after 4 weeks, 12 weeks, and 26 weeks	89%
Canada (Employment Services Transformation)	Payable at months 1 and 12	40-60%
France (Private Placement Operators scheme)	Payable upon entering employment (35% of maximum contract value), and at month 6 if client remains employed (further 35%)	70%
Ireland (JobPath)	Payable 13, 26, 39, and 52 weeks	65% or more
Italy (Reintegration vouchers)	Payments vary according to type of employment contract and client characteristics (variation between permanent contracts, fixed-term over 6 months, and fixed-term between 3-6 months)	Over 90%
Korea (National Employment Support Programme)	Outcome payments vary depending on: <ul style="list-style-type: none"> • Employment competency of individual • Income in the job • Speed of finding employment 	58-73%
New Zealand (no name)	Payment triggered after 0, 3, 6, 9, and 12 months of employment	73-78%
Sweden (KROM)	At entry of employment or education/training programme	70-80%
Sweden (STOM)	Employment outcome payment at 4 months (minimum) Education and training outcomes at 20 weeks (minimum)	50-65%
United Kingdom (WP)	Payments after 13 weeks or 26 weeks, and then for every additional 4 weeks in employment (each dependent on payment group)	75-100%†
United Kingdom (Restart)	Payment after individual has earned 6 months of cumulated earnings at 16hrs times National Living Wage within 18 months of starting the programme	70%
United States (CareerCompass and CareerAdvance)	Payment triggered based on sustained employed after 1, 6, and 12 months	30-40%
United States (Wisconsin Works)	Payment at 31 days and 91 days, alongside payment (respectively) for job placements of individuals who have received 2 years minimum of social assistance benefits and placement into high-wage jobs	Approx. 60%

Note: * The remainder of the payment is formed by the upfront service fees. † The delivery fee declined over the course of the contract to 0. Over the whole seven years it was effectively under 80%

Source: Langenbucher & Vodopivec (2022), DWP

There have been calls for a better understanding that sustainable outcomes require long-term investment and long-term monitoring (which also requires stable

contracting out frameworks).^{649,650,651} As indicated in Table 3, several countries do incentivise sustainability of employment in their contracted out programmes.

Carter proposes a more nuanced approach to payment groups and PbR, in which modelled likelihoods of moving *closer* to work can be measured as part of provider performance and outcomes.⁶⁵² Indeed, a similar approach has been used in some of the countries and programmes examined by the OECD (e.g. Canada).⁶⁵³ Rather than provider performance being measured solely in terms of the extent to which the final desired (employment) outcomes are achieved, it is also possible to measure achievement of intermediate outcomes which are known to represent movement towards work; and to add rigour to this by explicitly modelling the extent to which those intermediate outcomes increase the likelihood of finding work.

Findings from the interviews

Interviewees across all three categories emphasised that **lower costs should not be equated with VFM**.⁶⁵⁴ Some commissioners and providers believed a balance of VFM and provider goodwill or capacity has been achieved so far.⁶⁵⁵ Some of them even reported how this has been demonstrated in past impact assessments.⁶⁵⁶

Others reported outstanding issues in the market which were detrimental to sustaining positive short-term and long-term outcomes for all parties. Firstly, according to provider interviewees, the management of risk around volumes may disproportionately protect the commissioner.⁶⁵⁷ Additionally, some commissioner and expert interviewees mentioned the co-existing risks and potential associated with re-negotiation and change schedules.⁶⁵⁸

Both commissioner and provider interviewees suggested the approach to contracting out should direct fewer punitive measures towards providers and, instead, facilitate greater incentives for desired behaviours.⁶⁵⁹

'There's a perception we don't want government to be seen to be giving bonuses, while I do think personally there should be more of an incentive structure [...] If they deliver, why aren't we paying them bonuses, you know what I mean? [...] the opposite of [...] you deliver late, and you get a late charge [...], we don't do the early side of it, and I think that's because we've got limited funding [...] and we don't] want to be seen to be paying all this additional money out from a public perception point of view.' - Commissioner interviewee

⁶⁴⁹ The proposal centres on the multivariate binary logistic regression models and variables used are grouped as follows: (i) benefit type and history, (ii) health and disability characteristics, (iii) demographics, (iv) employment and household characteristics, and (v) local context. Across these variables there are statistically significant and sizeable effects that lead the author to conclude that improved understanding of the likely achievement of employment outcomes could help design an alternate payment system to better calibrate payment levels. See Carter (2018a).

⁶⁵⁰ Egdell et al. (2016).

⁶⁵¹ Koning, P. & Heinrich, C. J., (2010). "Cream-Skimming, Parking and Other Intended and Unintended Effects of Performance-Based Contracting in Social Welfare Services", IZA DP No. 4801

⁶⁵² Carter (2018a).

⁶⁵³ Langenbacher & Vodopivec (2022).

⁶⁵⁴ C6, C7, C8, E2, E4, E13, P12, P13, P17.

⁶⁵⁵ C2, C10, P12.

⁶⁵⁶ C10, P12.

⁶⁵⁷ C7, P6, P7, P8, P9, P11, P12, P14.

⁶⁵⁸ C3, C7, E4, E14, P11.

⁶⁵⁹ C4, C7, P2, P5.

'If the system penalises those who fall short of standards/targets and does not reward those who exceed them, that providers will just build this into their pricing' - DWP representative

As mentioned above, providers' dissatisfaction with the timing of contracts and the 'feast-famine' effect⁶⁶⁰ might be seen as a further hindrance to sustained quality delivery. Additional areas of contention include the following:

- the role of commercial versus policy priorities⁶⁶¹
- opinions on whether the market has matured and stabilised⁶⁶²
- the belief that VFM has previously been achieved (New Deal) but since lost⁶⁶³
- the question of whether market freedom is preferable to strong stewardship.⁶⁶⁴

Despite these observations, many reported good relations between DWP and their providers.⁶⁶⁵ It was felt that DWP has learned lessons while contracting out.⁶⁶⁶ For example, one commissioner interview stated DWP understood that previous strategies had driven providers to offer unrealistic bids and that DWP had realised that course-correction during a programme is sometimes necessary.⁶⁶⁷ Expert interviewees also observed that they have seen DWP experimenting over the years with a commendable intention to learn and improve:⁶⁶⁸

'I think it's important that we find space to acknowledge where there has been learning from experimentation and this kind of incremental improvement [...] and I definitely think you can see some of that learning coming through in Restart in particular.' - Expert interviewee

Provider interviewees suggested DWP management of entry into the market and costing per service user is now 'broadly'⁶⁶⁹ or 'about' right, improving on past approaches.⁶⁷⁰ It was thought that DWP has recognised previous issues and worked hard, in particular, on the Commercial Agreement for Employment and Health Related Services (CAEHRS) framework⁶⁷¹

Of course, some critiques of the approach remain;⁶⁷² namely, the need for DWP to use the knowledge generated so far to better define their overall objectives, plan these with long-term goals and then map their desired outcomes to these.⁶⁷³

⁶⁶⁰ C6, E12, P5, P18, P20.

⁶⁶¹ C1, P1, P2.

⁶⁶² C3.

⁶⁶³ P16.

⁶⁶⁴ C4, E13, P15.

⁶⁶⁵ C9, C10, P2, P5, P6, P7, P8, P14.

⁶⁶⁶ C8, E4, E13, P6, P15, P19.

⁶⁶⁷ C8.

⁶⁶⁸ E4, E13.

⁶⁶⁹ P6.

⁶⁷⁰ P19.

⁶⁷¹ P15.

⁶⁷² P1, P20.

⁶⁷³ C7, E2, E14, P1, P2, P9, P13, P17, P18, P20.

4. Live running

Key findings on management of contracts during life running

- While contract management is critical for ensuring the quality of service provision, there is no agreed definition of service ‘quality’ or how to measure it. Performance measurement and monitoring aim to encourage innovation, improve performance and limit gaming.
- Incentives to encourage collaboration or information sharing among providers are limited. There is also little evidence on sharing contract resources amongst providers and other stakeholders.
- Expertise and specialist knowledge of providers’ staff is key and need for professional development of staff was noted.
- Innovation in employment services may mean a new client group or new skills training for jobseeker or new ways to provide existing training. It may include making organisational processes more efficient. Competition intends to encourage innovation and this needs to be considered early on. However, evidence for innovation taking place in PbR contracts is limited.

4.1. Measuring service quality

4.1.1. There is no agreed definition of service ‘quality’ or how to measure it

Managing contracts, and monitoring the quality of services during the live running (or contract delivery), is important as this forms one of the ways commissioners can mitigate the risk of undesired behaviours (such as creaming or parking) among the providers.⁶⁷⁴

There is no universally applicable definition of quality in employment services.^{675,676,677,678,679} What it means to provide ‘high quality’ services will depend on the aims and objectives of the policy, for instance, whether the service is geared towards enabling sustained and/or higher paying job outcomes and/or job outcomes for specific groups of job seekers.

Another difficulty is setting targets for the said outcomes, as these need to anticipate the situation on the labour market at least 6 months or even for the length of the contract i.e. a couple of years ahead (if the economy is doing better than expected, the targets will have been set too; conversely, if the economy is not doing well, then

⁶⁷⁴ Langenbucher & Vodopivec (2022).

⁶⁷⁵ Bach-Mortensen, A. and Barlow, J. (2021). Outsourced austerity or improved services? A systematic review and thematic synthesis of the experiences of social care providers and commissioners in quasi-markets. *Social Science & Medicine*. 276.

⁶⁷⁶ Considine, M., O’Sullivan, S., & Nguyen, P. (2018). The policymaker’s dilemma: The risks and benefits of a ‘black box’ approach to commissioning active labour market programmes. *Social Policy & Administration*, 52(1), 229-251.

⁶⁷⁷ PES – Working Group (2013). PES Efficiency Working Group. Final report. European Commission.

⁶⁷⁸ Sienkiewicz, L. (in collaboration with ICF) (2016). Measuring Customer Satisfaction with PES. Increasing PES effectiveness by meeting customer needs. *Employment, Social Affairs and Inclusion*.

⁶⁷⁹ Stephan, G. (2016). Public or private job placement services - Are private ones more effective? *IZA World of Labour*. 285, p1-10.

the target will have been set too high). As such, the actual performance is more difficult to interpret, in that providers who exceed their target may in fact achieve little impact (and have high deadweight); by the same token, providers who fail to achieve their targets may actually achieve net impact. This is important as their performance affects the cost effectiveness of the overall programme, even if this is difficult to measure.

Since setting a good job outcomes target is difficult, it is important to define other measures of quality. Contracted out services will be expected to also include specifications for minimum service standards, which represent a baseline level of quality that all providers should achieve (see section 2.3.3). Service quality may also be understood more broadly than the terms of a PbR contract, in terms of wider factors such as customer experience and satisfaction or desirable outcomes such as innovation that do not tend to feature as criteria for outcome-based payments.

Findings from the interviews

Numerous and varied definitions of service quality emerged from the interviews. Some interviewees thought **outcomes** were the only clear demonstration of service quality.⁶⁸⁰ Others distinguished outcomes as 'performance', with 'quality' referring to the **customer experience**.⁶⁸¹

'I suppose the performance is about, [...] just get people into jobs [...] but the quality would be the customer experience so from when they're coming into the centre, how the centre looks, how they're looked after, is it an environment that's going to give them confidence? [Are they]] going to get the support that they need? Are they getting a good quality experience through whether it's CVs, application, job searching etc.? [...] and how they're treated, because you can be very focused on performance but not give that quality of delivery to the customer.' - Provider interviewee

Others tied service quality to the **processes** used to reach positive outcomes.⁶⁸² Combinations of the factors listed above were also used when defining quality. These included: outcomes plus (customer) service,⁶⁸³ outcomes plus processes⁶⁸⁴ and outcomes plus ethics.⁶⁸⁵ The latter refers to balancing outcomes against basic ethical and moral safeguarding of the service users' best interests and wellbeing in any assessment of quality.

For some, '**fidelity**', referring to the alignment of delivery with bid promises and contractual requirements, was the true indicator of quality.⁶⁸⁶ The use of quality

⁶⁸⁰ E9, E11, P2, P5, P8, P11.

⁶⁸¹ P10, P14.

⁶⁸² P7, P17.

⁶⁸³ E5, P18.

⁶⁸⁴ P1.

⁶⁸⁵ P8.

⁶⁸⁶ P13, P5.

assurance **certification**⁶⁸⁷ and the assessment of a service's **social value**⁶⁸⁸ were also suggested.

Concerns were raised about **over-specification or specification of misguided proxies** for quality.⁶⁸⁹ This was suggested to mistake standardised processes with quality service delivery, to the detriment of providers' capacity to maximise value for each client.

However, commissioner, expert and provider interviewees still asserted that aspects of quality assurance or control should be specified in contracts.⁶⁹⁰ Others also emphasised that the selection process could better seek out provider suggestions to improve quality or even use these proposals as the basis of live running quality measurement during contract delivery.⁶⁹¹

Regulations and professional standards were seen to be underutilised, where they exist.⁶⁹² Some interviewees suggested the introduction of 'kite-marks'⁶⁹³ while others proposed that an industry equivalent to Ofsted is established⁶⁹⁴ or a similar model to the Star Ratings, used in Australia, or other metrics which could even be visible to service users are adopted.⁶⁹⁵

At the time of writing this report there was little publicly available evidence on measuring service delivery in relation to health assessments (see Box 8). In late summer 2022 some information became available on the UK Parliament website through written questions, answers and statements.⁶⁹⁶

Box 8. Health assessments: service quality

The quality concerns surrounding the contract with Atos (see Box 3) were raised and the need for improvements was identified following 'a series of reviews'.⁶⁹⁷ According to DWP, lessons have been learnt: the negotiated agreement that covered the remaining term was 'more robust, with an agreed performance regime that gives [DWP] confidence delivery goals will be achieved. It is that same commercial rigour that will underpin the new procurement for these services'.⁶⁹⁸

According to DWP, provider performance on Personal Independence Payment (PIP) assessments is measured against quality, performance delivery targets and customer experience.⁶⁹⁹ The monthly performance measures against contractual targets in February 2022 show among others:

- Quality of assessments derived from the audit of all assessment reports at 94.4% (where the target was defined as 95% or more reports in each month deemed as acceptable or acceptable without significant learning points)

⁶⁸⁷ E7, P1, P2, P17.

⁶⁸⁸ E1, P20.

⁶⁸⁹ C4, C10, E15, P7, P11, P14, P15, P16.

⁶⁹⁰ C4, C6, C8, E1, E15, P19, P20.

⁶⁹¹ C6, P8, P9.

⁶⁹² C2, E1, E7, E15, P1, P7, P17, P19.

⁶⁹³ P1.

⁶⁹⁴ E1, E4, P2, P13.

⁶⁹⁵ C11, E3, E4, E7, P1, P2, P18.

⁶⁹⁶ For example: <https://questions-statements.parliament.uk/written-questions/detail/2022-06-27/26035>; <https://questions-statements.parliament.uk/written-questions/detail/2022-07-14/36914>; <https://questions-statements.parliament.uk/written-questions/detail/2022-07-14/36915>; <https://questions-statements.parliament.uk/written-questions/detail/2022-07-14/36916>

⁶⁹⁷ <https://publications.parliament.uk/pa/cm201314/cmhansrd/cm140327/wmstext/140327m0002.htm#14032769000011>

⁶⁹⁸ <https://publications.parliament.uk/pa/cm201314/cmhansrd/cm140327/wmstext/140327m0002.htm#14032769000011>

⁶⁹⁹ For a complete set of metrics please see: <https://questions-statements.parliament.uk/written-questions/detail/2022-06-27/26035>

- Claimant Satisfaction rate at 97.8% (92% is the annual target).⁷⁰⁰

Additional evidence in relation to health assessments was gained from the interviews. **Clinical governance standards** are a central part of the assessment process, making quality at once simpler and more complex.⁷⁰¹ Compliance with specified processes is of great importance during these assessments, while outcomes are not likely to be subject to targets.⁷⁰² Caseloads⁷⁰³ and number of complaints⁷⁰⁴ are used as indicators in quality assessment. Observations of service delivery have been carried out in some health assessment contracts.⁷⁰⁵

Challenges around representing user voices were also noted given that the decision, which is not taken by the provider, will likely dominate a user's impression of the service.⁷⁰⁶

4.1.2. Performance measurement and monitoring aims to encourage innovation and limit gaming

Contracting out and the introduction of competition requires strong monitoring and measurement of quality and overall performance.⁷⁰⁷ As noted by Langenbucher & Vodopivec, performance management in PbR contracts incorporates two aspects: monitoring quality of services delivered to clients and examining providers' performance against (outcome) targets.⁷⁰⁸ These aspects are usually set out in contract specifications through minimum service requirements (which regulate processes to reduce possibility for gaming) and performance targets (that incentivise innovation). Naturally, there is a tension between these two and balancing these is a challenge for commissioners.⁷⁰⁹

According to the OECD, performance measurement targets in PbR employment contracts include supporting clients (back) into work, reducing the time clients are without work, reducing the risk of long-term or recurring unemployment and securing client satisfaction from services received.⁷¹⁰ These targets can be measured through:

- monitoring providers (e.g. inspections, site visits) including accreditation, certification or licencing systems
- collating client feedback (e.g. conducting surveys or monitoring client complaints)
- measuring quantifiable achievements (e.g. referrals versus placements, sustainable employment outcomes, etc.).⁷¹¹

A number of existing practices illustrate the diversity in approaches taken by commissioners in addressing this:

⁷⁰⁰ <https://qna.files.parliament.uk/qna-attachments/1489210/original/SLAs%202018%202022.pdf>

⁷⁰¹ C2.

⁷⁰² C2, C3, C9.

⁷⁰³ P14.

⁷⁰⁴ C9.

⁷⁰⁵ C3, C9, E4, P5, P14, P17.

⁷⁰⁶ C2, C3, E14, P11.

⁷⁰⁷ Andersson et al. (2019)

⁷⁰⁸ Langenbucher & Vodopivec (2022).

⁷⁰⁹ Langenbucher & Vodopivec (2022).

⁷¹⁰ Langenbucher & Vodopivec (2022).

⁷¹¹ Langenbucher & Vodopivec (2022).

- **Austria** is monitoring contract performance using a Balanced Scorecard which includes levels of customer and staff satisfaction.⁷¹²
- From 2022, **Australia** is moving to a new licensing system but it is not clear if the Star Ratings model, which was widely discussed in the reviewed literature,⁷¹³ continue.⁷¹⁴ The plans for the new system require providers to apply for a licence, which after (successful) assessment according to the Provider Performance Framework would be granted to selected providers for three years. This licence would be extended for high performing providers after an annual review, while licences for low performers would be revoked and made available to others – as the number of licences would be capped. The application process can be seen in Figure 2.⁷¹⁵ Examples of performance measures in the Provider Performance Framework include:
 - achieving employment outcomes for job seekers
 - progressing job seekers towards employment (e.g. successfully completing designated milestones, such as training, to increase employability skills)
 - quality of service given to job seekers and employers
 - job seeker and employer satisfaction
 - achieving results for disadvantaged cohorts.⁷¹⁶

Figure 2. Example of an organisation interacting with the Australian licensing system

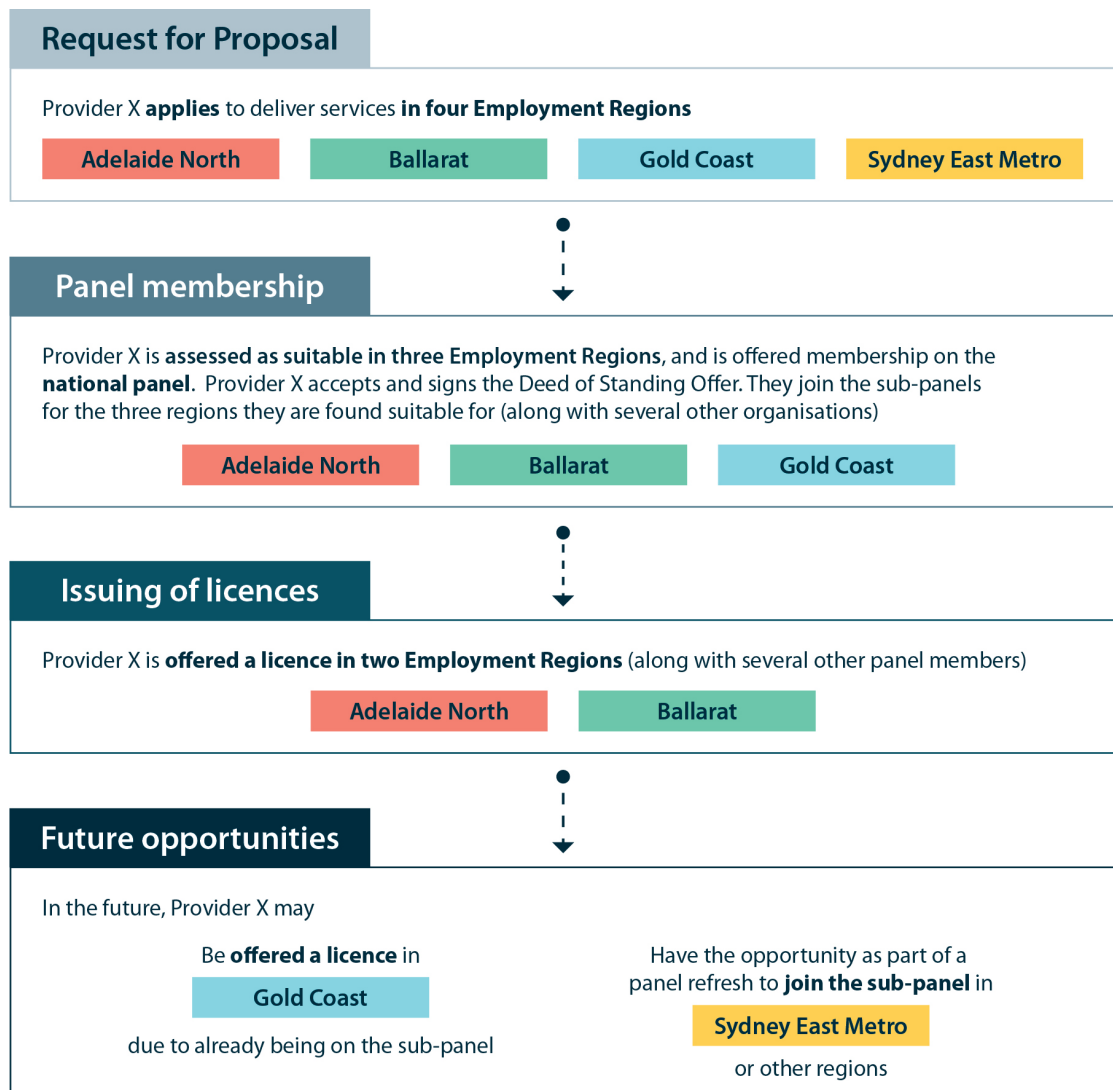
⁷¹² HoPES – Working Group (2013). PES Efficiency Working Group. Final report. European Commission.

⁷¹³ Langenbucher & Vodopivec (2022).

⁷¹⁴ DESE (2021). As of 19/08/2022 : <https://www.dese.gov.au/new-employment-services-model/resources/new-employment-services-update-and-proposed-licensing-system-webinars-frequently-asked-questions>

⁷¹⁵ DESE (n.d.). Proposed licensing system for the New Employment Services Model: Discussion paper. As of 03/05/2022: <https://www.dese.gov.au/new-employment-services-model/proposed-licensing-system-nesm-discussion-paper>

⁷¹⁶ DESE (n.d.).



Source: Department of Education Skills and Employment (DESE, n.d.). Proposed licensing system for the New Employment Services Model

Findings from the interviews

The potential for **user feedback** to better inform the monitoring of performance and service quality was emphasised across all three groups of interviewees.⁷¹⁷

Service user voices were described as ‘chronically underused’,⁷¹⁸ with one interviewee reporting that ‘review systems’ have been seen to facilitate transparency and dialogue between users and providers.⁷¹⁹ Another speculated that service users with disadvantages or complex needs might provide more useful feedback on quality and performance, given the standardisation of provision for ‘mainstream’ users.⁷²⁰

It may be that overarching or political contexts also impact on service user feedback in the employment services setting. For instance, changes in the profile of service users during Restart or the introduction of stricter sanction regimes over time could

⁷¹⁷ C3, C9, C11, E2, E12, E13, E14, P4, P9, P11, P18.

⁷¹⁸ E13.

⁷¹⁹ C11.

⁷²⁰ E2.

shift the focus of user feedback from the provision to other aspects of their experience.⁷²¹

Observations of service delivery have been carried out in some contexts.⁷²² These might be a complementary tool to user feedback for the purposes of establishing multiple monitoring perspectives.

There is evident frustration with quality monitoring that focuses exclusively on compliance. These practices were characterised as **'tick-box' activities or 'paperwork'** which distract from seemingly more meaningful work.⁷²³ Examples of such activities include insistence on 'wet signatures'⁷²⁴ or set numbers of in-person meetings.⁷²⁵

However, some interviewees held the view that the **minimum service requirements** used so far have been neither sufficient nor accompanied by an appropriate level of monitoring.⁷²⁶ In comparison, commissioners in Australia are thought to have sophisticated methods for logging each of a provider's interactions with a job seeker.⁷²⁷ This large data set can then be combined with providers' qualitative reports on quality and progress to assess performance.⁷²⁸ Accreditation and professional standards bodies could also play a role in monitoring quality and driving improvement:

'Making sure that every [...] provider's advice and guidance is matrix accredited would be a good start in showing that people are professionally qualified. Or [if] there was professional systems in place, or ensuring people are all members of ERSA⁷²⁹ or a trade body or something like that where there was good practice sharing [...] I know not all prime contractors are a member of [ERSA], but it just seems that there's opportunities for [...] not self-policing of quality, but [...] the dispersal and spreading of quality. Then also [...making] the measurements [...] more objective, because there is no Ofsted for employment services.' - Provider interviewee

Accountability measures are evidently an important aspect of monitoring quality. Determining which measures should be used is less clear. One possibility involves withholding payment for low quality or poor performance.⁷³⁰ However, this leans hard on the punitive lever, without considering incentives or the role of clear and shared definitions of quality in facilitating improvement.

Beyond specifications and monitoring, it may be that **sustained investment** (mostly in longer contracts) is an effective tool in improving focus on quality and facilitating high performance.⁷³¹

⁷²¹ P9.

⁷²² C3, C9, E4, P5, P14, P17.

⁷²³ E3, E13, P1, P4, P8, P11, P12, P13, P14.

⁷²⁴ P4, P16.

⁷²⁵ P11, P15, P16, P18.

⁷²⁶ E3, E4, E5, E12, E13, E14.

⁷²⁷ E3.

⁷²⁸ E3.

⁷²⁹ Employment Related Services Association. See: [here](#).

⁷³⁰ C2, E4, E7, E15, P18.

⁷³¹ E2, E13.

4.2. Incentives to encourage collaboration or information sharing among providers are limited

Providers operating in a quasi-market may experience both incentives and disincentives to collaborate and share information with other providers. Networking and relationship building with other providers helps to build trust and foster collaboration between providers.⁷³² This is particularly important where customers with complex needs interact with multiple providers who need to coordinate with each other.^{733,734,735,736} Collaboration across providers has the potential to improve the efficiency and effectiveness of employment services if providers are able to learn from one another by sharing examples of good practice and innovation.^{737,738} At the same time, there may be a reluctance to share information and to collaborate with organisations with which one competes for resources (for instance, if the number of referrals is linked to performance) or when bidding for future contracts.⁷³⁹ This reluctance or difficulty is further exacerbated by a lack of transparency, commercial sensitivities and contract confidentiality.⁷⁴⁰

Collaborating (or even merging) with other providers may make the market more stable or predictable and weaken competition,⁷⁴¹ although from the commissioner's perspective, competition is crucial to the effective functioning of quasi-markets.^{742,743} What appears from one perspective to be collaboration designed to improve efficiency for all providers might appear to others as collusion, countering the competitive market forces are fundamental to driving quality under a PbR framework.^{744,745}

Commissioners may take steps to incentivise or facilitate collaboration between providers whilst maintaining healthy competition in a quasi-market. This may occur at the bidding stage, for instance by adapting bid and tender requirements to include SMEs or partnerships, or to establish relationships with other (public service) stakeholders like the NHS.⁷⁴⁶ Commissioners may also remove barriers to information-sharing to incentivise collaboration, for example, through improving

⁷³² Struyven & Steurs (2004).

⁷³³ Carter, E. (2021). 'More than 91arketized? Exploring the governance and accountability mechanisms at play in Social Impact Bonds', *Journal of Economic Policy Reform* 24:1, pp.78-94

⁷³⁴ Carter (2018a).

⁷³⁵ Whitworth & Carter (2018).

⁷³⁶ Van Geste et al. (2019).

⁷³⁷ Ejler, N. and Sidemann, P. (2016). Application of Process Efficiency Techniques in PES. European Commission Analytical Paper.

⁷³⁸ Davern, E., Nunn, A. and Scopetta, A. (2021). Stakeholder conference "The power of PES partnerships" 20-21-22 April 2021 Synthesis paper

⁷³⁹ Langenbacher & Vodopivec (2022).

⁷⁴⁰ Struyven & Steurs (2004).

⁷⁴¹ Struyven & Steurs (2004).

⁷⁴² Bruttel, O., (2005). 'Delivering active labour market policy through vouchers: experiences with training vouchers in Germany', *International Review of Administrative Sciences* 71(3), pp.391-404

⁷⁴³ Timo Weishaupt, J. (2014). Central Steering and Local Autonomy in Public Employment Services. PES to PES Dialogue.

⁷⁴⁴ Bruttel (2005).

⁷⁴⁵ Timo Weishaupt, J. (2014). Central Steering and Local Autonomy in Public Employment Services. PES to PES Dialogue.

⁷⁴⁶ Ceolta-Smith et al. (2015).

systems and efficiency for sharing and changing current legal/commercial restrictions.^{747,748}

Findings from the interviews

Some commissioner, expert and provider interviewees reported they had seen no evidence of collaboration.⁷⁴⁹ These and others also suggested incentives for collaboration were non-existent or insufficient.⁷⁵⁰

It may be that the **narrow focus on claimable outcomes** promoted by PbR discourages collaboration⁷⁵¹ or, indeed, that the **competition of the quasi-market** itself means collaboration is not a commercially viable practice.⁷⁵² Some expert interviewees spoke of instances of staff poaching as an example of how collaboration is disincentivised.⁷⁵³

Other disincentives include barriers to data-sharing, whether due to the requirements of contracts or concerns around intellectual property.⁷⁵⁴ Indeed, for some health assessment contracts, exclusivity or the absence of other providers made collaboration impossible.⁷⁵⁵

Nevertheless, there are examples of limited collaboration between providers.⁷⁵⁶ These largely take the form of **roundtables and forums** for sharing success stories or good practice.⁷⁵⁷ Other, more practical, instances include **sharing points of contact with employers** for efficiency.⁷⁵⁸

Interviewees across all three stakeholder groups emphasised that collaboration is the product of sustained and mutual **relationship building**.⁷⁵⁹ This may be informal and could be the result of commercial staff moving between providers and building networks and goodwill.

The following incentives to promote collaboration were suggested:

- introducing contractual obligations around collaboration, either requiring primes to share best practice with their own supply chain⁷⁶⁰ or upgrading existing 'exhortation' and 'encouragement'⁷⁶¹ for prime-to-prime collaboration to compulsory requirements which would be monitored along with other standards⁷⁶²
- launching a collective target or collaboration bonus⁷⁶³

⁷⁴⁷ Scharle, A. (2011). New Developments, Good Practices and Lessons Learnt. One year of the PES to PES Dialogue: Discussions at the Dissemination Conference. DG Employment, Social Affairs and Inclusion.

⁷⁴⁸ Struyven & Steurs (2004).

⁷⁴⁹ C7, E7, E8, E13, P1, P6, P7, P8, P13.

⁷⁵⁰ C3, C7, C9, E1, E2, E5, E13, P1, P2, P4, P7, P9, P12, P14, P17.

⁷⁵¹ E2, P4.

⁷⁵² C1, C2, C7, C10, E1, E5, E7, E9, E12, E14, P1, P2, P4, P7, P9, P13, P15, P16, P17, P20.

⁷⁵³ E9, E12.

⁷⁵⁴ C3, E13, P6, P10, P11, P12, P15, P16, P18.

⁷⁵⁵ C2.

⁷⁵⁶ C3, C8, C9, C11, E9, E12.

⁷⁵⁷ C9, C11, E2, E3, E7, P1, P8, P9, P10, P11, P14, P15, P16, P18, P19.

⁷⁵⁸ E12, P16.

⁷⁵⁹ C1, C3, C8, E1, E2, E11, E12, P1, P5, P9, P10, P11, P12, P14, P15, P19, P20.

⁷⁶⁰ E4.

⁷⁶¹ E12, P15.

⁷⁶² E12, P9, P11, P15, P17, P18.

⁷⁶³ C1.

- facilitating shared problem-solving (e.g., for a case management system).⁷⁶⁴

While the commissioner's 'encouragement' to collaborate was noted,⁷⁶⁵ collaboration is thought to require resources and perhaps greater facilitation from DWP.⁷⁶⁶

'People can share information and this sort of thing and therefore we can pay for [...providers] sharing bits and pieces. So, I am aware of those items that are in the contract [...] they're in the set up – are they used effectively in life is a different question and I'm not sure they are currently.' - Commissioner interviewee

'I think [...] you also have to create the space for providers to be able to share with each other. They are a fairly – below the kind of very high-level company to company, [...] they are a fairly collaborative bunch, but in the old days there used to be a lot more networking, a lot more events, a lot more activity, and a lot of that has got squeezed out because the contracts have been much tighter, there's been less funding in the system and because the [DWP] doesn't support it.' - Expert interviewee

The ReAct partnership was named as a good example of a collaboration scheme which brings together providers delivering the Restart programme, led and funded by six of the prime providers for the purpose of sharing evidence and good practice.⁷⁶⁷ However, while the potential of the ReAct partnership was noted, it was believed a lack of financial resources was the cause of two providers opting out of this instance of collaboration.⁷⁶⁸

4.3. There is little evidence on sharing contract resources amongst providers and other stakeholders

In addition to sharing information and examples of good practice with other delivery organisations, providers may share contract resources, either with providers operating within the same programme or a wider set of stakeholders. Resource sharing (as opposed to information sharing) was mentioned infrequently in the literature, making it difficult to draw conclusions about the extent to which this occurs and the circumstances in which it may be facilitated. Resource-sharing is referenced more frequently in the literature in relation to the PES compared to contracted provision.⁷⁶⁹

⁷⁶⁴ P2.

⁷⁶⁵ C3, C5, C10, E4, E8, P8, P15, P19.

⁷⁶⁶ E1, E2, P4, P5, P16, P20.

⁷⁶⁷ See the ReAct website [here](#) for more information,

⁷⁶⁸ P16, P20.

⁷⁶⁹ Sultana and Watts. (2005). Career Guidance in Europe's Public Employment Services

Concerns around the commercial sensitivity of contract resources may deter providers from sharing these with other organisations. Sharing contract resources amongst providers and other stakeholders is heavily dependent on trust and likely driven by the desire to help service users.⁷⁷⁰ For instance, Innovation Fund projects in the UK have been described as cooperative efforts between parties to achieve shared social goals based on relationships of trust.⁷⁷¹

Findings from the interviews

Interviewees across all three stakeholder groups suggested deliberate **DWP facilitation** is required to enable any sharing of contract resources beyond their remit.⁷⁷² This may be due to legal restrictions or the logistics of liaising with other public services with whom DWP should already have contacts. In fact, interviewees, from commissioner, expert and provider groups, reported they had never seen this type of resource sharing.⁷⁷³ There seems to be a distinct **lack of incentives** for collaboration of this kind.⁷⁷⁴

A few examples of resource sharing amongst providers and other stakeholders involved **'joined-up' service provision** which aimed to prioritise holistic support for service users. Interviewees reported on examples of this type of collaboration occurring within the cities of Manchester⁷⁷⁵ and Sheffield.⁷⁷⁶ The Talent Match programme was also named as an instance of 'joined-up' operation being facilitated.⁷⁷⁷ Other instances include the development of employer networks, mentioned above, which could be shared by providers and other similar initiatives.⁷⁷⁸

Despite some commissioner and provider interviewees indicating a lack of evidence for this type of collaboration, they claimed they would **expect providers to initiate** this behaviour.⁷⁷⁹ The commissioner interviewees suggested plans for sharing resources with other stakeholders, especially public services, should be requested as part of providers' bids.⁷⁸⁰

Another factor hindering this type of resource-sharing may be a **scarcity of resources** to share:

'Given the relatively low level of funding that the providers had, you know their financial incentive was to implement relatively low-cost measures or to persuade other organisations to fund an intervention where they could. Of course, then that impacted on other organisation's budgets. If you try and do things on the cheap, you can't then really expect too much.' - Expert interviewee

⁷⁷⁰ Carter, E., (2021). 'More than marketised? Exploring the governance and accountability mechanisms at play in Social Impact Bonds', *Journal of Economic Policy Reform* 24:1, pp.78-94

⁷⁷¹ Carter (2021).

⁷⁷² C3, C7, E4, E5, E13, P1, P11, P18, P20.

⁷⁷³ C3, C7, C11, E8, P1, P6, P8, P10, P11, P13, P16, P17, P20.

⁷⁷⁴ C7, C11, P2, P11, P17.

⁷⁷⁵ E4, E13.

⁷⁷⁶ E4.

⁷⁷⁷ E2.

⁷⁷⁸ C11, P2, P17, P19.

⁷⁷⁹ C8, P2, P20.

⁷⁸⁰ C8.

This would come under a broader problem of provider capability in which the investment in outreach and collaboration required exceeds the current financial model and skills basis.⁷⁸¹ Competing for funds or risking confusion as to the 'ownership' of outcomes are further disincentives linked to this issue.⁷⁸²

There was very little evidence on sharing information and collaboration amongst providers in relation to health assessments (see Box 9). This can be partly explained by the fact that the number of providers offering these services is much smaller, compared to employment support.⁷⁸³

Box 9. Health assessments: collaboration and sharing resources

In the health assessment context, interviewees believed resource sharing could and did occur but had to be requested.⁷⁸⁴ However, given the sensitivities and different expectations of the health assessment tasks, this was considered appropriate and understandable.⁷⁸⁵

4.4. Expertise and specialist knowledge of providers' staff is key and need for professional development of staff was noted

Outsourcing employment services can enable access to specific skills, knowledge and expertise held by providers, potentially reducing costs and improving service quality.^{786,787} Compared to the PES, providers may have more in-depth knowledge of how to support specific groups of service users, particularly those with complex needs (also discussed in relation to the prime/sub model in section 2.5), and/or a better understanding of local labour markets.⁷⁸⁸ This specialist expertise may enable providers to deliver more efficient or higher quality services compared to the PES.⁷⁸⁹ Providers – particularly private sector organisations – may be better placed than the PES to investment into new and innovative technologies,⁷⁹⁰ which could lead to gains in efficiency and/or effectiveness. Private sector providers may also have strengths in managerial efficiency and entrepreneurship, which have benefits for the quality of employment services.⁷⁹¹

A lack of capacity and knowledge can impede on the delivery of services. Rees et al. (2014) describe an example of how job advisors in one prime provider lacked the knowledge of where to refer individuals with complex needs (for specialist support)

⁷⁸¹ C11, E2, E14.

⁷⁸² E5, E8, E15, P13.

⁷⁸³ C3.

⁷⁸⁴ C3.

⁷⁸⁵ C3.

⁷⁸⁶ Sasse et al. (2019).

⁷⁸⁷ Ejler & Sidelmann (2016).

⁷⁸⁸ Dockery & Stromback (2001).

⁷⁸⁹ Sasse et al. (2019).

⁷⁹⁰ Sasse et al. (2019).

⁷⁹¹ van Gestel et al. (2019).

effectively delaying needed support (and parking such clients).⁷⁹² There is also a possibility that where providers do have relevant knowledge and expertise, these may be used to cream more effectively in PbR models.⁷⁹³ Freedom to use own skills and capabilities encourages flexibility, but risks providers working in their own best interests.⁷⁹⁴

Findings from the interviews

Some interviewees across all three stakeholder groups emphasised the benefits of provider **agility and efficiency** and, in the case of prime contractors, a unique capacity to adapt and react in **operational management**.⁷⁹⁵

In regard to expertise in the delivery of employment (support) services, some expert and provider interviewees reported there was **no ‘secret knowledge’ on ‘what works’** being withheld by providers from the commissioner.⁷⁹⁶ With most methods of engagement with jobseekers being established and shared across providers, this view again suggests that operational expertise is the true ‘added value’ brought by providers to the delivery of this service.⁷⁹⁷ Provider interviewees expressed pride in their support and management of specialist subcontractors as a crucial facilitation of connecting expertise to the delivery of employment services.⁷⁹⁸

Discussion of provider expertise mostly emphasised the **specific and local expertise of small, charitable and specialist providers** who make up the supply chain managed by a prime contractor.⁷⁹⁹ However, this also raised potential difficulties associated with operating as a smaller or charitable provider in a prime-sub and PbR model.⁸⁰⁰ Concern was expressed by expert and provider interviewees around the survival of organisations in this environment and the **subsequent loss of expertise** as a result.⁸⁰¹ With this expertise most likely related to more specific categories of ‘harder-to-help’ or disadvantaged jobseekers, this would also negatively impact service users.⁸⁰² Some expert and provider interviewees also believed existing specialist organisation’s full potential was not being met, largely due to the financial constraints of PbR and a lack of connected services.⁸⁰³

Indeed, it appears contracts have **lacked financial incentives** for providers to fully utilise their expertise or that of their supply chains. This was attributed by expert and provider interviewees both to a tendency for PbR to narrow focus onto faster outcomes⁸⁰⁴ and the constraints of specification around customer journeys, or even who could be referred to the programme.⁸⁰⁵

⁷⁹² Rees et al. (2014).

⁷⁹³ Egdell et al. (2016).

⁷⁹⁴ Rehwald et al. (2015).

⁷⁹⁵ C5, C9, E5, E13, P2, P16, P18.

⁷⁹⁶ E1, E2, E5, P1.

⁷⁹⁷ P7.

⁷⁹⁸ P19.

⁷⁹⁹ C8, E2, E12, E13, E15, P5, P7, P10.

⁸⁰⁰ E12.

⁸⁰¹ E4, E12, P10.

⁸⁰² E12, P12.

⁸⁰³ E2, E13, P7, P13.

⁸⁰⁴ E2, P4, P7, P13.

⁸⁰⁵ P7, P11, P17.

'We tend to talk about programmes as being open programmes where anybody can refer onto it [or] closed programmes [which] we deliver for DWP – but I think [...] one of our specialisms is that over the years we've created really good partnerships with other organisations so they may very well be supporting unemployed people and they very well be a better referral pathway for that person than DWP, but that's not allowed. So I think they don't allow us to use that specialism of "we've got really good relations with other organisations" and I think that's a really missed opportunity' - Provider interviewee

The expertise of staff who are interacting with service users was also emphasised as crucial.⁸⁰⁶ Providers need the motivation and resources to **recruit, train and retain skilled staff**.⁸⁰⁷ One interviewee also emphasised that greater involvement of ground-level staff in bids and perhaps even contract design could better utilise this expertise in 'hands-on' service delivery.⁸⁰⁸

Several interviewees across commissioners, providers and experts suggested **DWP could learn more from provider expertise**.⁸⁰⁹ From a commissioner perspective, such knowledge could be utilised in future policymaking, or even to choose to take certain services back in-house.⁸¹⁰ In the view of some providers and experts, better utilisation of provider expertise could shape improved selection processes and contract design, arguably benefitting all parties.⁸¹¹

It is worth noting that the historical and continued expertise of DWP was also raised. This referred to the Department's **unique access to data and analytical insights** as the commissioner, as well as institutional learning over many decades.⁸¹²

Some evidence specific to health assessments can be found in Box 10.

Box 10. Health assessments: expertise and specialist knowledge

The question of provider expertise at a staff level is **more complicated in the health assessment context**.⁸¹³

'We also have to make sure that they don't become too specialist, because specifically in our contract, [...] they bring specialist knowledge in terms of how you might deal with different disabilities or environments, so there's a really strong clinical and continuous learning presence [...] But again [...] I'm not buying assessors who deliver assessments specifically for certain types of people, as it were. So, therefore, it's really important that they use that expertise in a broad sense.' - Commissioner interviewee

There are concerns around maintaining a generalist approach to health assessments, with fears that a little knowledge could be more dangerous than none.⁸¹⁴ This could manifest in staff holding assumptions around health conditions or disabilities which impact their assessment of the service user's claims and capacities. It could also lead to confusion around the role of the assessor for the

⁸⁰⁶ E3, E4, E8, E9, P1, P6, P14.

⁸⁰⁷ E3, E9, P6.

⁸⁰⁸ P10.

⁸⁰⁹ C5, E2, P2, P6, P10, P13, P16, P19.

⁸¹⁰ C5.

⁸¹¹ E2, P2, P6, P10, P13, P16, P19.

⁸¹² C8, E1.

⁸¹³ C2, E14.

⁸¹⁴ C2, E14.

service user, risking a belief they should have received a 'specialist' assessor to match their health condition or disability.⁸¹⁵

4.5. Encouraging innovation through PbR

4.5.1. Innovation in employment support services can mean different things

Before findings on if and how PbR contracts encourage innovation are presented, this section outlines what innovation in employment services may entail. Thijs and Staes (2012) detail numerous types of service innovations adapted from Hartley (2005).⁸¹⁶

- product innovation, including specific skills training for the long-term unemployed
- service innovation, developing new ways for how services can be provided and includes tailored coaching for jobseekers
- process innovation, developing new ways in which organisational processes are designed; examples of this type include 'lean approaches' looking into processes to make them as efficient as possible
- position innovation, looking for new contexts or 'customers' with examples of adapting the PES to changing contexts of individuals needing services
- strategic innovation through new goals/purposes for the organisation; this includes privatisation and outsourcing additional services (e.g., counselling)
- governance innovation looking at new forms of citizen engagement, including further decentralisation of job centre functions
- rhetorical innovation of new language and concepts.

4.5.2. Competition intends to encourage innovation

Competition is one of the key reasons to outsource services in the first place, as there is belief it can act as a spur for innovation.⁸¹⁷ Competition in service provision can stimulate innovation, for example, in the education and health care sectors.⁸¹⁸ Evidence suggests the payoffs to innovation are small when competition is between public entities.⁸¹⁹

Decentralisation, in relation to innovation, allows for local flexibility, thus promoting upwards competition.⁸²⁰ Competition can result in more testing of innovations, which in turn improves service quality and cost effectiveness.⁸²¹

⁸¹⁵ C2, E14.

⁸¹⁶ Hartley (2005) in: Thijs, N. and Staes, P. (2012). Organisational development, improvement and innovation management in Public Employment Services. European Commission.

⁸¹⁷ Sasse et al. (2019).

⁸¹⁸ Poutvaara & Jordahl (2020).

⁸¹⁹ Poutvaara & Jordahl (2020).

⁸²⁰ Timo Weishaupt, J. (2014). Central Steering and Local Autonomy in Public Employment Services. PES to PES Dialogue.

⁸²¹ Dockery & Stromback (2001).

4.5.3. Encouraging innovation needs to be considered early on

For the outsourcing of contacts, engagement in innovative thinking needs to occur from the start.⁸²² This can be through early dialogue with potential suppliers and increased understanding of new technologies.⁸²³ There is also encouragement to consider innovation-based procedures, and inviting the market to suggest novel and innovative solutions to problems.⁸²⁴ Partnerships and collaboration can also enable experimentation and testing of new approaches.⁸²⁵

The literature suggests there is a need within contracts for flexibility to encourage and allow for innovation.^{826 827} Delivery bodies need flexibility within their contacts to innovate.⁸²⁸ Therefore, the challenge is to design pricing, processes and outcome measures in ways that on the one hand providers have the flexibility (which enables innovation) whilst on the other hand the service quality for the users is protected.⁸²⁹ Flexibility in programme design can be achieved, for example, through design waivers, and allowance for special budget allocations for innovative local programs.⁸³⁰ An example in the Netherlands highlighted that a too strict monitoring regime will impose a high administrative burden on both the department and providers, with the potential to reduce flexibility and deter innovation.⁸³¹

There is, however, a challenge of encouraging innovation early on and allowing for flexibility. Sultana and Watts (2005) identify a key challenge for the PES to confront. This is seeking to find the right balance between, on the one hand, encouraging innovative, flexible, and context-sensitive responses in service delivery between clients and labour markets demands, and on the other, maintaining standards across all providers.⁸³² This aims to ensure service users have guaranteed access to the same quality service, irrespective of their geographical or social location.⁸³³ Increased flexibility and freedom can act as a 'double edged sword',⁸³⁴ as it creates the opportunity for reduced costs and potential for innovation, but also gives the opportunity for providers to act in their own best interest.⁸³⁵

Notably, research suggests that certain regions and areas are more likely to innovate than others.^{836,837} Carter (2018) commented that the local/city region system is most innovative compared to the central government. Leading city-regions can articulate a

⁸²² HM Government (2021). 'The Sourcing Playbook: Government guidance on service delivery, including outsourcing, insourcing, mixed economy sourcing and contracting', Open Government licence v3.0

⁸²³ HM Government (2021).

⁸²⁴ HM Government (2021).

⁸²⁵ Davern et al. (2021).

⁸²⁶ Carter (2021).

⁸²⁷ Mosley (2011).

⁸²⁸ Carter (2021).

⁸²⁹ Finn. (2010).

⁸³⁰ Mosley (2011).

⁸³¹ Struyven & Steurs (2004).

⁸³² Sultana and Watts. (2005). Career Guidance in Europe's Public Employment Services.

⁸³³ Sultana and Watts. (2005).

⁸³⁴ Rehwald et al. (2015).

⁸³⁵ Rehwald et al. (2015).

⁸³⁶ Carter (2018a).

⁸³⁷ Finn (2011).

progressive vision of locally integrated employment support.^{838,839} Integration is key for providers to be able to respond sensitively to context-specific needs and individual needs, which requires connections across advice, skills, housing, among other support.⁸⁴⁰ Some findings from the UK suggest in certain contexts, and under certain contractual agreements, improve outcomes for particular groups and encourage innovation to service delivery.⁸⁴¹

4.5.4. There is limited evidence for innovation taking place

Opportunities for innovation are often not taken.⁸⁴² Examining prospectuses of the Work Programme, Ceolta-Smith et al. (2015) comment that some prime providers showed promise of designing innovative interventions, but it was not clear if these would have been effective or cost effective.⁸⁴³ Indeed, the authors questioned whether a black box commissioning can effectively stimulate innovation for health-related approaches.⁸⁴⁴ Carter (2018) notes that providers appeared to use increased freedom and flexibility to maximise profits and deprioritise services to those more difficult to move into work.⁸⁴⁵

As noted earlier (section 3.1), price competition and emphasis on reduced costs within programme design can act to stifle innovation, resulting in reduced quality and the provision of a “bare bones” service.⁸⁴⁶ Innovation is also often expensive, and consequently “low hanging fruit” innovations are selected when handling the balance between service quality and cost trade off.⁸⁴⁷

A too strict monitoring regime, or full specified contract, imposes high administrative burden on both the department and providers, and has the potential to reduce flexibility and deter innovation.⁸⁴⁸ Therefore, over-monitoring can quash innovation.

Decentralisation can also weaken the capacity of central management to such an extent that it becomes difficult to manage an innovative project on a large, national scale.⁸⁴⁹ An example provided from Lindsay and McQuaid (2009) highlights that in the Netherlands, government representatives acknowledged disappointment at the lack of innovative practice, despite increases of flexibility in this market.⁸⁵⁰

The reviewed literature also provided several examples of how innovation is encouraged or promoted, primarily focusing on the PES, rather than contracted out services. This may be due to the fact that if innovation occurs through black box, or even grey box commissioning, it is more difficult to capture, document and share

⁸³⁸ Whitworth & Carter (2018).

⁸³⁹ Carter (2018a).

⁸⁴⁰ Carter (2018a).

⁸⁴¹ Finn (2011).

⁸⁴² Carter (2018a).

⁸⁴³ Ceolta-Smith et al. (2015).

⁸⁴⁴ Ceolta-Smith et al. (2015).

⁸⁴⁵ Carter (2018a).

⁸⁴⁶ Finn (2011).

⁸⁴⁷ Considine et al. (2018).

⁸⁴⁸ Struyven & Steurs (2004).

⁸⁴⁹ Sultana and Watts. (2005). Career Guidance in Europe's Public Employment Services

⁸⁵⁰ Lindsay, C. & McQuaid, R. W., (2009). 'New Governance and the Case of Activation Policies: Comparing Experiences in Denmark and the Netherlands', *Social Policy and Administration* 43(5), pp.445-463

more broadly due to the issues around transparency, information sharing and collaboration discussed above.

Findings from the interviews

Overall, **risk aversion** appears to stifle innovation by providers; this means that any rewards of experimenting in service delivery are not often anticipated or, indeed, attempts at innovation would risk breaking with the specifications of the contracts.⁸⁵¹

Specifically, interviewees across all three stakeholder groups suggested innovation was **not incentivised by PbR** in its current form.⁸⁵² The risk of failing to claim payments after investing in innovative methods appears to be too great for providers to take on. Furthermore, a basic lack of cash-flow to spend on innovative schemes is also seen as a consequence of the PbR model, if up-front investment or service fees are low.⁸⁵³ Some also claimed that innovation was **restricted by specifications or compliance**.⁸⁵⁴ This restriction was either a matter of specifications not permitting other activities or of compliance-related activities taking up resources which could then not be used to innovate.

However, these interviewees and others also suggested there is a consensus within the market, and its advisors, on **tried and test methods in employment support services**.⁸⁵⁵ This means that, even when given the opportunity to innovate - via black box specification or other models - providers are not motivated to deviate from established procedures. In fact, some commissioner and provider interviewees believed this to be a wise decision, with some scepticism expressed around a search for 'magic dust' or 'reinventing the wheel'.⁸⁵⁶ Instead, there were suggestions that innovation be focused on service users known to require additional support⁸⁵⁷ or proposals aiming to meaningfully shift long-term (un)employment trends.⁸⁵⁸

A distinction between **incremental and disruptive or radical innovation** was also suggested:⁸⁵⁹

'There are sort of two types of innovation [...] there's radical innovation and then there's incremental innovation and radical innovation is [...] doing something completely new and completely innovative and experimental, and then incremental innovation is tweaking [...] and I think what we have seen is that from payment by results [...] they are good on incremental innovation and adaptation [...] so I think the argument for the incremental bit is pretty good [...] but not the radical innovation, so doing something completely new is just not something that's done in [...] payment by results because you're on the hook with the financial risk so why would you do something completely new?' - Expert interviewee

⁸⁵¹ C1, C4, C5, E2, E8, E9, E12, P13, P16, P20.

⁸⁵² C1, C4, C9, C11, E5, E8, E9, E11, E13, P2, P4, P5, P17, P18, P20.

⁸⁵³ C4, C11, E2, E5, E9, P17.

⁸⁵⁴ E8, E13, E15, P2, P9, P14, P16, P17.

⁸⁵⁵ C4, C7, E1, E4, E9, E13, P5, P7, P12, P17.

⁸⁵⁶ C4, E2, P12.

⁸⁵⁷ E5.

⁸⁵⁸ P8.

⁸⁵⁹ E11.

In relation to employment, it may be that the quasi-market structure **does not generate sufficient competitive drivers** to motivate providers to radical innovation.⁸⁶⁰ Small shifts and improvements within a provider organisation may, however, generate enough incremental innovation to improve their performance, relative to other competitors.⁸⁶¹ In contrast, one provider interviewee suggested that a competitive drive to agree to low price points means shifting provider focus from improvement and innovation towards survival and sustainability.⁸⁶²

Suggestions for **better promotion of innovation in the future** included increasing focus on innovation at the tender and bidding stages⁸⁶³, reducing barriers to data-sharing⁸⁶⁴ and using different payment structures or pilots to better facilitate innovation without the risk or resource-scarcity described here.⁸⁶⁵

Some evidence specific to health assessments can be found in Box 11.

Box 11. Health assessments: innovation

One interviewee observed that the health assessments contracts have seen incremental innovation, with changes proposed and then evaluated.⁸⁶⁶ This is likely to be expected due to the nature of the work commissioned and the different identification of 'outcomes' in comparison to outsourced employment services.

⁸⁶⁰ E1, E11, P18.

⁸⁶¹ E7, E11, P12.

⁸⁶² P18.

⁸⁶³ C7, P7, P8, P11.

⁸⁶⁴ C5, E15, P9, P18.

⁸⁶⁵ C4, C5, C11, E1, E2, E3, E4, P2, P5, P13, P17, P20.

⁸⁶⁶ C2.

5. Conclusions

5.1. The aims of this research

The contracting out approaches promise effectiveness and efficiency gains for the commissioners and society at large, and financial rewards for those involved in service delivery. There is a long tradition of contracting out employment services and health assessments in the UK. Despite this, there are still gaps in understanding of how the processes of commissioning could be further improved. Addressing these gaps would help DWP (and other commissioners), as well as providers and clients to reap fuller benefits from different models of contracting.

This study reviews available evidence from the literature and held by key stakeholders: commissioners, providers and experts in the field. It takes a comprehensive approach to the contracting out process: from the point of designing contracts and its different aspects (such as associated costs, benefits and risks, level of specifications, etc.) through to issues arising during market competitions and live running of the awarded contracts.

The main research question this review responds to is **about the advantages and disadvantages of contracting out public employment services and health assessments and the optimum model(s) for commissioning these.**

5.2. Summary of findings and key conclusions

5.2.1. Contracting models: key considerations and trade-offs

The reviewed evidence focuses primarily on contracts utilising PbR mechanisms, which are common in the commissioning of employment services in the UK and internationally. There were no international examples of commissioning of health assessments in the reviewed literature.

PbR contracts are associated with potential advantages (for instance, a better use of specialist knowledge and cost reductions in service delivery) but also with possible drawbacks (such as gaming practices, disadvantaging small, specialist providers and non-profits). Whether the benefits outweigh the costs largely depends on contracts:

*'Unless contracts are well specified and payments well structured, negative incentive effects will seriously detract from social outcomes.'*⁸⁶⁷

⁸⁶⁷ Dockery & Stromback (2001).

Some of these undesirable costs and risks can be addressed by contract design or specification. However, addressing one characteristic of PbR contracts often has implications for other features, as summarised below. Commissioners may face tensions or trade-offs between:

- supporting the largest number of people into work and effectively reaching job seekers with the most significant barriers to employment
- the number of job outcomes achieved and the degree to which these are high-quality and sustainable
- creating scope for innovation and adaptation (e.g., to the local labour market) whilst ensuring a baseline level of quality is met by all providers
- allowing providers flexibility to tailor services to meet the needs of service users without resulting in highly uneven or variable service provision
- building long-term relationships with providers with the capacity to take on the risk of PbR contracts and creating a diverse marketplace where smaller, more specialist providers can flourish.

Optimum contract design will depend on the aims and objectives of the programme, meaning it is not possible to draw overarching conclusions about the ‘best’ approach. The design of PbR contracts will be informed by the unique circumstances of the country, region, client group, and labour market.^{868 869} While there is no single recipe for a flawless PbR contract model, the evidence base is growing, allowing commissioners to make better informed choices suited for their needs.

Effective contracting out requires successfully establishing a quasi-market.^{870 871} Again, the evidence and guidance available for commissioners in the area of market stewardship is scarce but growing numbers of successful and unsuccessful experiences and examples provide some indication for future actions. These lessons are further emphasised below.

5.2.2. Contract design: optimal contract model to achieve the desired outcomes

Table 4 lists the main features of PbR contracts, outlines levers through which they could be optimised and flags how these levers may lead to adverse effects elsewhere. However, these need to be considered with caution as other factors – also discussed in this report may affect delivery (for example, the way in which the targets are set versus how the economy will do during contract duration which may give an inaccurate impression of actual impact/effectiveness).

Table 4. Optimising PbR contracts to achieve the desired outcomes

⁸⁶⁸ Stephan (2016).

⁸⁶⁹ Struyven (2004).

⁸⁷⁰ Stephan (2016).

⁸⁷¹ Langenbucher & Vodopivec (2022).

Benefits, costs and risks in PbR contracting	Assumption	Lever	Associated negative effects
Benefit: improved effectiveness and cost-efficiency	Focus on outcomes drives better results ('what gets measured, gets done')	Increased weighting for outcome payments	Gaming behaviour by providers Disadvantaging specialist providers / subcontractors by primes
Cost: high transaction costs	Reducing costs of commissioning through PbR increases VFM to commissioners	Less detailed contract specifications (black box) Fewer minimum service requirements Longer, fewer and bigger contracts	Variability in service provision Deteriorating service quality Reducing market competition
Risk: gaming strategies	Reducing incentives to gaming minimises risks and costs to commissioners	Differentiate outcome payments by client group More minimum service requirements More detailed contract specifications (grey box)	Limiting flexibility of providers and potential for innovation
Risk: primes' monopoly	Market competition stimulates greater effectiveness and cost savings for commissioners	Smaller CPAs or two-stage competitions within existing CPAs Shorter contracts (Financial) incentives or contractual requirements for primes to work with smaller providers	Increased costs for commissioners
Risk: quality of services	Safeguards must be in place to ensure service users receive high quality provision	More minimum service requirements Strengthening monitoring arrangements or introduce a licencing / accreditation system Strengthening enforcement or breakout clauses	Increased costs for commissioners Reduced incentives (increased costs) for providers

A key consideration in commissioning employment services is minimising the risk of creaming and parking, enabling job seekers with the greatest barriers to employment to benefit fully from programmes. Creaming and parking undermine the equity of programmes and – when these result in deadweight – weaken value for money for

the commissioner. However, mechanisms to address creaming and parking (minimum service standards, greater specification in contracts and differentiated outcome payments) may have undesirable effects (see Table 4). Greater consistency in provision may help to ensure job seekers with the greatest barriers to employment receive support, but the flip side of this may be less innovative, tailored services. It is also extremely difficult to differentiate outcome payments in a way that accurately reflects job seekers' distance from the labour market. This issue taps into fundamental questions about who employment programmes are for and what they hope to achieve; there can be no 'one size all' solution.

5.2.3. Market stewardship and service quality

Findings in this report show a number of lessons learned by commissioners in contracting out employment services in different contexts. These include:

- Developing a long-term strategy for the market and engagement with providers in the commissioning process so that numerous programmes are informed by the current situation on the market and commissioned continuously, with some overlaps, to avoid 'feast and famine' cycles feared by providers
- Using selection criteria carefully, especially since too much emphasis on price may harm service quality and drive some providers away, thus stifling competition (and innovation)
- Considering implications for the altered composition of the market and a more dominant role of providers that can operate at scale but offer relatively similar services, as opposed to smaller and specialist providers, including not-for-profits and SMEs. Evidence shows that contracting out has so far led to this result but it is unclear if (and how) this affects the service users or the market in the long term.

It is challenging to find consensus on definitions of service quality in the provision of the services considered in this report. While the 'success' of a programme may speak through itself through outcomes such as sustained employment placements, achieving strong impact and VFM, the maintenance of consistent, fair and tailored engagement with users is also of concern. Similarly, strong provider expertise can manifest in high-quality provision and strong outcomes, but commissioners might also fear that this expertise instead translates into more effective gaming behaviours.

Some perspectives would conceptually separate performance from quality, with the latter encompassing indicators such as positive customer experience or compliance with minimum service requirements. It may be that standardisation across programmes, and thus the industry, could help align stakeholders on the question of quality. This standardisation could involve the introduction of regulatory and professional standards with observations by an inspection body similar to Ofsted.

The findings from this report on measuring and monitoring quality identified the following obstacles and opportunities:

- Service user feedback has so far been under-utilised but has certain limitations as a metric of quality measurement, with users' views on an overall policy or other factors beyond provider control potentially shaping their assessment of delivery

- Emphasis on fidelity and compliance is seen as an important safeguard for service users and guard against many risks for the commissioner, although providers can feel limited in their delivery of high-quality and tailored provision by ‘tick-box’ exercises
- Accountability measures such as withholding payment for poor performance can be an important lever in contracts but it may be that a punitive approach reduces opportunities to improve quality through sustained investment and mutual learning
- Cashflow restrictions and a narrow focus on outcomes is seen by some as an obstacle to the use of provider expertise which could improve the quality of service, especially for disadvantaged users or those with complex needs or a greater ‘distance’ to travel to employment.

5.2.4. Collaboration

Providers operating in a quasi-market may experience both incentives and disincentives to collaborate and share information and resources, both with other providers and with external stakeholders. Collaboration may foster mutual learning and the dissemination of effective and innovative practice, enabling providers to learn from the specialist expertise of other organisations. From the providers’ perspective, however, there is a tension between the benefits of collaboration and the potential risks of sharing commercially sensitive information with competitors in a quasi-market. The sharing of information and resources is heavily dependent on trust, and in the context of employment services may be driven by a shared social mission to produce the best possible outcomes for service users. It is therefore plausible that the level of collaboration amongst providers in various contracting models vary significantly: while it seems limited in PbR contracts, it may be more common in other models (e.g. partnership or grant agreements).

Commissioners may seek to drive improvements in performance by encouraging and incentivising collaboration, removing barriers to information-sharing through improving systems and changing current legal/commercial restrictions. However, commissioners will be mindful of the need to maintain healthy competition in a quasi-market and to ensure that collaboration does not enter into the territory of collusion.

5.2.5. Innovation

A key motivation for introducing PbR contracts, particularly in combination with flexible (black box) provision, has been encouraging innovation in service provision. This review finds no evidence of PbR contracts leading to break-through innovations in effectively supporting clients into jobs. To an extent, this may reflect the fact that innovation is hard to define and measure. Whilst a theoretical argument is often made in the literature that PbR contracts encourage innovation, empirical evidence on this point is scarce. However, incremental changes in how, when or where services are provided may make small differences for providers to operate more economically, while maintaining the focus on outcomes. The evidence to support this is found in examples of cost savings achieved in some countries and testimonies from a number of interviewees.

With this in mind, the debate about whether the black box approach is superior to grey box in stimulating (breakthrough) innovation is constrained by a lack of evidence. Given that possibilities to save costs through incremental innovation cannot be exercised endlessly, it is also conceivable that the scope for this may be declining. Moreover, elements of PbR contracts thought to encourage innovation, for instance flexible provision, may have disadvantages such as greater scope for creaming and parking and variable service provision. PbR contracts may not always be the most appropriate vehicle to stimulate innovation in commissioning employment services and other forms of contracting may be considered and explored (such as small-scale grants, competitions).⁸⁷²

5.3. Remaining evidence gaps and future research

One of the most evident gaps relates to contracting out health assessments. The evidence of how these services are organised in other countries, as well as perspectives of the UK current, past and prospective providers (including their healthcare and administrative staff) would help better understand both the state of play in contracting out these services and possible ways forward.

Most of the evidence reviewed and collated in this report presents perceived costs and benefits of PbR contracts and how different levers within this contracting model can mitigate some risks but potentially increase others. Some of those risks are not unique to PbR but may be applicable to contracting out more broadly. Stronger evidence on the actual costs incurred and benefits achieved of different models of provision (in-house, contracted out (including PbR, Cost Plus and payment for service contracts), partnership, grant funding arrangements) is needed to be able to compare these. On that basis it would be possible to draw more informed conclusion on which works better for whom, under what conditions, and why.

⁸⁷² See for example: NESTA (2018). Funding innovation: A practice Guide. Making money work harder. As of 18/05/2022: <https://media.nesta.org.uk/documents/Funding-Innovation-Nov-18.pdf>

References

- Adams, L., Oldfield, K., Riley, C. & Nightingale, M. (2011). Trial incapacity benefits reassessment: customer and staff views and experiences. DWP.
- Adams, L., Oldfield, K., Riley, C., Vegeris, S., Husain, F., Bertram, C., ... & Vowden, K. (2011). Jobseekers Regime and Flexible New Deal Evaluation: findings from longitudinal customer surveys and qualitative research (Vol. 767). A report of research carried out by Policy Studies Institute on behalf of the Department for Work and Pensions.
- Albertson, K., Fox, C., O’Leary, C., Painter, G., Bailey, K., & Labarbera, J. (2018). Payment by Results and Social Impact Bonds. University of Chicago Press Economics Books.
- Andersson, F., Jordahl, H., and Josephson, J. (2019). Outsourcing Public Services: Contractability, Cost, and Quality. IZA Discussion Paper Series. No 12401.
- Australia Productivity Commission (2002). “Independent Review of the Job Network”, Report No.21, AusInfo, Canberra.
- Ayaita, A. et al, (2021). 'Job Placement via Private vs. Public Employment Agencies: Investigating Selection Effects and Job Match Quality in Germany', IZA Discussion Paper No.14024
- Bach-Mortensen, A. & Barlow, J. (2021). Outsourced austerity or improved services? A systematic review and thematic synthesis of the experiences of social care providers and commissioners in quasi-markets. *Social Science & Medicine*. 276
- Beale, I, Bloss, C and Thomas, A. (2008). The longer-term impact of the New Deal for Young People. Department for Work and Pensions Working Paper No 23.
- Behaghel, L. et al. (2012). 'Private and Public Provision of Counselling to Job-Seekers: Evidence from a Large Controlled Experiment', IZA Discussion Paper No.6518
- Bruttel, O., (2005). 'Delivering active labour market policy through vouchers: experiences with training vouchers in Germany', *International Review of Administrative Sciences* 71(3), pp.391-404
- Benjamin, O. (2020). Mothers’ negotiation of welfare support: the emergence of privatized entitlement, *International Journal of Sociology and Social Policy*, 40(11/12), pp.1439-1454.
- Carey, G. et al. (2020). 'Quasi-market shaping, stewarding and steering in personalization: the need for practice-orientated empirical evidence', *Policy Design and Practice* 3(1), pp.30-44
- Carter, E. (2018a). 'Making Markets in Employment Support: Promises and Pitfalls in the Work Programme’s Private Power Market'. PhD Thesis
- Carter, E. (2021). 'More than marketised? Exploring the governance and accountability mechanisms at play in Social Impact Bonds', *Journal of Economic Policy Reform* 24:1, pp.78-94
- Carter E. and Whitworth A. (2015). Creaming and parking in quasi-marketised welfare-to-work schemes: Designed out of or designed into the UK Work Programme? *Journal of Social Policy* 44(2): 277–296.

- Ceolta-Smith et al. (2015). citing Weishaupt, J. T. (2011). Social Partners and the Governance of Public Employment Services: Trends and Experiences from Western Europe, Geneva: International Labour Organization Page 11.
- Ceolta-Smith, J., Salway, S., & Tod, A. M. (2015). A Review of Health-related Support Provision within the UK Work Programme—What's on the Menu?. *Social Policy & Administration*, 49(2), 254-276.
- Child, C., Gibbs, B. G., & Rowley, K. J. (2016). Paying for success: An appraisal of social impact bonds. *Global Economics and Management Review*, 21(1-2), 36-45.
- Cockx, B., & Baert, S. (2015). Contracting Out Mandatory Counselling and Training for Long-Term Unemployed. Private For-Profit or Non-Profit, or Keep it Public? IZA DP No. 9459
- Cole, A., Cubi-Molla, P., Elliott, R., Feast, A., Hocking, L., Lorgelly, P., Payne, K., Peek, N., Sim, D., Sussex, J., Zhang, K and Steuten, L. (2021). Making Outcome-Based Payment a Reality in the NHS. Phase 2: Practical Considerations; van Stolk, C., Bjornsson, G., and Goshev, S. (2010). Provider Incentives in Social Protection and Health: A selection of case studies from OECD countries. RAND Corporation
- Collins, A.M., Coughlin, D., Miller, J., Kirk, S. (2015). The Production of Quick Scoping Reviews and Rapid Evidence Assessments: A How to Guide.
- Conolly, A., Day, N., Hales, J., Kinsella, R. and O'Connor, W. (2010). The operation of choice in Multiple Provider Employment Zones. Department for Work and Pensions research report no. 650.
- Considine, M., O'Sullivan, S., & Nguyen, P. (2018). The policymaker's dilemma: The risks and benefits of a 'black box' approach to commissioning active labour market programmes. *Social Policy & Administration*, 52(1), 229-251.
- Davern, E., Nunn, A. and Scopetta, A. (2021). Stakeholder conference "The power of PES partnerships" 20-21-22 April 2021 Synthesis paper
- Department for Transport (2015). Value for Money Framework.
- Department for Work and Pensions (2020). 'The Work Programme – A quantitative impact assessment.' As of 24.11.2021: <https://www.gov.uk/government/publications/the-work-programme-impact-assessment>
- DESE (2021). As of 19/08/2022 : <https://www.dese.gov.au/new-employment-services-model/resources/new-employment-services-update-and-proposed-licensing-system-webinars-frequently-asked-questions>
- DESE (n.d.). Proposed licensing system for the New Employment Services Model: Discussion paper. As of 03/05/2022: <https://www.dese.gov.au/new-employment-services-model/proposed-licensing-system-nesm-discussion-paper>
- Dockery, A. M. & Stromback, T. (2001). 'Devolving public employment services: preliminary assessment of the Australian experiment', *International Labour Review*, 140(4), pp.429-451
- DWP (2002). As of 12/05/2022: <https://www.gov.uk/government/organisations/department-for-work-pensions/about>

- DWP (2008). Quoted in DWP (2015), Review of the Merlin Standard. Department for Work and Pensions research report no. 907, pp. 15.
- DWP (2013). Evaluation of the Work Choice Specialist Disability Employment Programme: Findings from the 2011 Early Implementation and 2012 Steady State Waves of the research. See: rrep846.pdf (publishing.service.gov.uk)
- DWP (2013). New Enterprise Allowance: Qualitative evaluation. See: rrep836.pdf (publishing.service.gov.uk)
- DWP (2013). Support for the very long term unemployed trailblazer: longer term analysis of benefit impacts
- DWP (2014). "Maximus appointed to carry out Health assessments for the Department of Work and Pensions", As of 17/06/2022: <https://www.gov.uk/government/news/maximus-appointed-to-carry-out-health-assessments-for-the-department-for-work-and-pensions>
- Egdell, V. et al. (2016). 'Third Sector Experiences of Work Programme Delivery ', Journal of Social Policy 45(4), pp.729-746
- Ejler, N. and Sidelmann, P. (2016). Application of Process Efficiency Techniques in PES. European Commission Analytical Paper.
- Finn, D. (2010). Outsourcing Employment Programmes: Contract Design and Differential Prices
- Finn, D. (2011). Sub-contracting in Public Employment Services: Review of research findings and literature on recent trends and business models. The European Commission Mutual Learning Programme for Public Employment Services. DG Employment, Social Affairs and Inclusion
- Finn, D. (2012). Subcontracting in public employment services: the design and delivery of 'outcome based' and 'black box' contracts: analytical paper. Centre for European and International Studies Research.
- FitzGerald, C. et al., (2019). 'Walking the contractual tightrope: a transaction cost economics perspective on social impact bonds', Public Money and Management 39:7, pp.458-467
- Foster et al (2014). Work Programme Evaluation: Operation of the commissioning model, finance and programme delivery. See: Work-programme-evaluation-operation-of-the-commissioning-model-finance-and-programme-delivery-summary.pdf
- Fuertes, V., Jantz, B., Klenk, T. and McQuaid, R. (2014). 'Between cooperation and competition: the organisation of employment service delivery in the UK and Germany', International Journal of Social Welfare, 23: S1, S71–S86.
- Gash, T., Panchamia, N., Sims, S., and Hotson, L. (2012). Making public service markets work: Professionalising government's approach to commissioning and market stewardship. Institute for Government
- Gash, T. and Panchamia, N. (2012). When to contract: Which service features affect the ease of government contracting?
- Greenberg, D. H., & Davis, A. (2007). Evaluation of the New Deal for Disabled People: The cost and cost-benefit analyses. Department for Work and Pensions.

- Griffiths, R and Durkin, S. (2007). Synthesising the evidence on Employment Zones. Department for Work and Pensions Research Report 449
- Hartley (2005) in: Thijs, N. and Staes, P. (2012). Organisational development, improvement and innovation management in Public Employment Services. European Commission.
- HM Government (2021). 'The Sourcing Playbook. Government guidance on service delivery, including outsourcing, insourcing, mixed economy sourcing and contracting.' As of 24.11.2021: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/987353/The_Sourcing_Playbook.pdf
- HoPES – Working Group (2013). PES Efficiency Working Group. Final report. European Commission.
- Jacobi, L. & Kluve, J., (2006). 'Before and After the Hartz Reforms: The Performance of Active Labour Market Policy in Germany', IZA Discussion Paper No.2100
- Jantz, B. et al. (2018). 'Marketization and Varieties of Accountability Relationships in Employment Services: Comparing Denmark, Germany, and Great Britain', *Administration and Society* 50(3), pp.321-345
- Kay, J. & Marlow, S. (2020). 'The Work Programme: A Quantitative Impact Assessment', UK Department for Work and Pensions
- Knight, G et al. (2013). Provider-led Pathways to Work: Net impacts on employment and benefits Department for Work and Pensions Working Paper No 113
- Koning, P. & Heinrich, C. J., (2010). "Cream-Skimming, Parking and Other Intended and Unintended Effects of Performance-Based Contracting in Social Welfare Services", IZA DP No. 4801
- Langenbucher, K. & Vodopivec, M. (2022). Paying for results: Contracting out employment services through outcome based payment schemes in OECD countries. OECD Social, Employment and Migration Working Papers No. 267
- Lindsay, C. & McQuaid, R. W. (2009). 'New Governance and the Case of Activation Policies: Comparing Experiences in Denmark and the Netherlands', *Social Policy and Administration* 43(5), pp.445-463. Page 458.
- Lødemel, I., & Trickey, H. (Eds.). (2001). 'An offer you can't refuse': workfare in international perspective. Policy Press.
- Manoudi, A. et al. (2014). 'Small Scale Study on PES Business Models', European Commission
- Mosley, H., G., (2011). 'Decentralisation of Public Employment Services', European Commission Mutual Learning Programme for Public Employment Services
- Mulgan, R. (2006). Government Accountability for Outsourced Services
- Mulheirn, I and Menne, V. (2008). The Flexible New Deal: Making it work. Social Market Foundation. As of 16.06.22: <https://www.smf.co.uk/wp-content/uploads/2008/09/Publication-The-Flexible-New-Deal-Making-it-work.pdf>
- NAO (2015). Outcome-based payment schemes: government's use of payment by results.
- NESTA (2018). Funding innovation: A practice Guide. Making money work harder. As of 18/05/2022: <https://media.nesta.org.uk/documents/Funding-Innovation-Nov-18.pdf>

- Newton, B., Meager, N, and Sainsbury, R. (2014). Work Programme evaluation: the participant experience report Department for Work and Pensions research report 892.
- O'Halloran, D. et al, (2022). 'The development of the Australian Unemployed Workers Union Rating Scale (AUWURS) of employment service providers', *Australian Journal of Public Administration*, (Early View), p.3.
- OECD (1997). *Contracting out government services: Best Practice Guidelines and Case Studies*. Occasional papers No. 20.
- OECD (2006); in: Lindsay & McQuaid (2009).
- OECD (2021). *Applying Evaluation Criteria Thoughtfully*.
- Pastore, F. (2019). The quasi-market of employment services in Italy. IZA DP No. 12662
- Poutvaara, P. and Jordahl, H. (2020). Public sector outsourcing. The desirability of outsourcing the provision of public services depends on their characteristics and market conditions. *IZA World of Labor*. 65 (2).p1-11.
- Rees, J. et al. (2014). 'Support for All in the UK Work Programme? Differential Payments, Same Old Problem', *Social Policy & Administration* 48(2), pp.221-239
- Rehwald, K., Rosholm, M., and Svarer, M. (2015). Are Public or Private Providers of Employment Services More Effective? Evidence from a Randomized Experiment. *IZA Discussion Paper Series*. No. 9365.
- Sasse, T., Guerin, B., Nickson, S., O'Brien, M., Pope, T. and Davies, N. (2019). 'Government outsourcing: What has worked and what needs reform?' Institute for Government. As of 24.11.2021:
<https://www.instituteforgovernment.org.uk/sites/default/files/publications/government-outsourcing-reform-WEB.pdf>
- Scharle, A. (2011). *New Developments, Good Practices and Lessons Learnt. One year of the PES to PES Dialogue: Discussions at the Dissemination Conference*. DG Employment, Social Affairs and Inclusion.
- Scottish Government (2018). 'Report – to date – of the Strategic Public Social Partnership (PSP) Model in Scotland.
- Shutes, I. & Taylor, R. (2014). Conditionality and the financing of Employment Services - Implications for the Social Divisions of Work and Welfare. *Social Policy & Administration*. 48 (2), p204-220.
- Sienkiewicz, L. (in collaboration with ICF) (2016). *Measuring Customer Satisfaction with PES. Increasing PES effectiveness by meeting customer needs*. Employment, Social Affairs and Inclusion.
- Stafford, B., (2012). Supporting moves into work: New Deal for Disabled People findings. *Scandinavian Journal of Disability Research*, 14(2), pp.165–176. DOI:
<http://doi.org/10.1080/15017419.2011.558235>
- Stephan, G. (2016). Public or private job placement services: Are private ones more effective?. *IZA World of Labor* 2016: 285 doi: 10.15185/izawol.285

- Struyven, L. (2004). Design choices in market competition for employment services for the long-term unemployed. OECD Social, Employment and Migration working papers, no. 21.
- Struyven, L. & Steurs, G. (2004). 'Quasi-market reforms in employment and training services: first experiences and evaluation results' in Evaluation of Systems and Programmes: Third report on vocational training research in Europe, Luxembourg: Office for Official Publications of the European Communities, (Cedefop Reference Series 57)
- Sultana and Watts. (2005). Career Guidance in Europe's Public Employment Services
- Tubb, H., (2012). 'Activation and Integration: Working with Individual Action Plans', European Commission Mutual Learning Programme for Public Employment Services
- van Gestel, N. et al. (2019). From quasi-markets to public-private networks: Employers' engagement in public employment services. Policy Administration. 53, pp.434-448.
- Vegeris et al (2011). Flexible New Deal evaluation: Customer survey and qualitative research findings. Department for Work and Pensions.
- Warren, J., Garthwaite, K. & Bamba, C., (2014). 'After Atos Healthcare: is the Employment and Support Allowance fit for purpose and does the Work Capability Assessment have a future?', Disability and Society, 29:8, pp.1319-1323
- Weishaupt, T.J. (2014). Central Steering and Local Autonomy in Public Employment Services. PES to PES Dialogue.
- Welsh Government (2021). 'Draft Social Partnership and Public Procurement (Wales) Bill
- Whitworth, A., & Carter, E. (2018). 'Rescaling employment support accountability: From negative national neoliberalism to positively integrated city-region ecosystems', Environment and Planning C: Politics and Space 36(2), pp.274-289

Annex 1: Research framework

Topic / Question	Interpretation / How we plan to respond to the question	Review	Interviews	Workshop
Contract design				
1. 1.1 What is the best payment model for achieving the outcomes that DWP is seeking to achieve, when contracting out employment and health assessment services?	<p>Q1(1) is the main question in the “Contract design” group and will draw on evidence found in relation to Q1(2-5) to answer it.</p> <p>It is unlikely that a single model will be identified as the best solution. It is likely that models will have different advantages and disadvantages.</p> <ul style="list-style-type: none"> We will identify possible models through the evidence review and interviews. We will establish key parameters of these models (risks, benefits, costs, implications, etc.), their advantages and disadvantages through evidence review and interviews. We will develop a set of criteria to score each method. We will analyse and compare different models and aim to identify an optimum one during the workshop. We will identify ways to improve how employment and health services are contracted out in workshop discussions. 	■	■	■
1.2 How do we assess and, if possible, price risk ⁸⁷³ and how should that inform the contracting model?	<ul style="list-style-type: none"> We will seek evidence on the types of risk associated with different types of contracts, including the impact of factors such as uncertainty on volume and performance, or a pricing cap.⁸⁷⁴ We will seek evidence on managing risk when contracting out services and on the extent of transfer of risk to providers through the review and interviews. 	■	■	
1.3 What are the costs and benefits of payment by result (PbR) ⁸⁷⁵ ?	<ul style="list-style-type: none"> We will document costs (e.g. increased overhead, perverse incentives, loss of goodwill) and benefits of PbR models identified through evidence review and interviews. 	■	■	

⁸⁷³ Different contracts have different risks – and the greater the risk, the higher the price. For example, a high level of uncertainty about the volumes (targets) for the contract requires to price that risk; if volumes (targets) are fairly certain then the risk is low and does not have to be priced.

⁸⁷⁴ A price cap sets a limit on the prices that a provider can charge (taken from: <https://www.investopedia.com/terms/p/price-cap-regulation.asp>).

⁸⁷⁵ PbR is a model of financing public services where payments depend on the results achieved by service providers, thus providing a financial incentive for providers to deliver good services throughout the term of the contract (see Cabinet Office 2011 Open Public Services White Paper. Cabinet Office, London)

Topic / Question	Interpretation / How we plan to respond to the question	Review	Interviews	Workshop
1.4 How much specification ⁸⁷⁶ should be used in an employment contract?	<ul style="list-style-type: none"> It seems unlikely the level of detail in contract specifications will be discussed in the literature, but we will look out for it. We will ask views and opinions of interviewees on the essential, optional and redundant elements of employment contracts. We will ask views and opinions of interviewees on the optimum extent and detail of specification in contracts to balance effectiveness with opportunity for innovation. 	?	■	■
1.5 What lessons were learnt from using Cost Plus ⁸⁷⁷ contracts during the pandemic?	<ul style="list-style-type: none"> It not likely this information will be available in the literature. We will ask views and opinions of interviewees on the subject, informed by a review of Cost Plus contracts, if available and accessible to us. 	?	■	
2. What are the implications of using a prime/sub model to support clients with complex needs?	<ul style="list-style-type: none"> It seems unlikely that this information will be available in the literature, but we will look out for it. We will ask views and opinions of interviewees on this subject and, particularly, the risks of this model failing to connect specialist small providers with people with complex needs 	?	■	
3. What are the implications of different strategies for geographical division for Contract Package Areas ⁸⁷⁸ ?	<ul style="list-style-type: none"> We do not anticipate that there will be a large among of information in the literature, but we will look out for any analysis of Contract Package Areas. We will ask interviewees for their views on geographical division and particularly its implications for competition and (the limits of) economies of scale. 	?	■	
4. Are there optimal lengths for contracts?	<ul style="list-style-type: none"> It seems unlikely that this information will be available in the literature, but we will look out for it. We will ask views and opinions of interviewees on the subject. 	?	■	
Market competition				
5. What questions, criteria and scoring	<ul style="list-style-type: none"> We will review the literature to establish what was tried and pros/cons identified. We will seek opinions from interviewees of how these could be improved. 	■	■	

⁸⁷⁶ Specification here refers to what extent the way(s) how the service should be provided is detailed. No or little specification means 'black box' provision where providers can do whatever they want as long as they achieve outcomes – this should encourage innovation. Specifying every detail of the provision (e.g. in implementing evidence-based interventions) limits innovation but reduces risks to providers allowing the price to be based primarily on costs. Specifying some aspects of service provision but allowing providers to tailor other elements to customers according to evidence base is referred to by DWP as 'grey box'.

⁸⁷⁷ A cost-plus contract is an agreement to reimburse a company for expenses incurred plus a specific amount of profit, usually stated as a percentage of the contract's full price (taken from: <https://www.investopedia.com/terms/c/cost-plus-contract.asp>).

⁸⁷⁸ Contract Package Areas are the geographical areas across England and Wales with a single prime contractor managing a network of subcontractors delivering services in the area.

Effective contracting of employment and health services

Topic / Question	Interpretation / How we plan to respond to the question	Review	Interviews	Workshop
mechanisms are used in commissioning services and what effect do these have on market competition and (dis)incentives for efficacy and efficiency?				
6. 6.1 How does the type of contracts affect the dynamics of market competition in the short- and long-term?	<ul style="list-style-type: none"> We will review the literature to find evidence on market entry/adaptation/exit, on provider goodwill or capacity (and the role of a price cap on competition). We will seek views from service providers representatives about their experiences. 	■	■	
6.2 When designing contracts and running competitions, what are the trade-offs between optimising short-term and long-term outcomes?	<ul style="list-style-type: none"> It seems unlikely that this information will be available in the literature. We will seek views from service commissioners and experts on the trade-offs between factors such as 'value-for-money' in the short term and maintaining provider capacity or goodwill, in the long term, as well as other such examples 	?	■	
Live running				
7. How is service quality measured and what customer service standards are used?	<ul style="list-style-type: none"> We will review the literature to establish what was tried and pros/cons identified. We will seek opinions from interviewees of how quality could be measured and what elements customer service standards should include. 	■	■	
8. What incentives are used to encourage collaboration or sharing information among contractors?	<ul style="list-style-type: none"> We expect most evidence will come from interviews with service providers representatives or other commissioners, but we will look out for it in the review too. 	?	■	■
9. To what extent do service providers share contract resources with each other or other lines of business– and does this benefit the department?	<ul style="list-style-type: none"> We expect evidence will come from interviews only. We will ask interviewees for their views and opinions on the extent of resource-sharing by contractors and on its potential benefits or costs to the DWP. 		■	■

Effective contracting of employment and health services

Topic / Question	Interpretation / How we plan to respond to the question	Review	Interviews	Workshop
<p>10. Do providers bring their own knowledge and expertise to the area – and to what extent (if at all) does the DWP allow them to use it?</p>	<ul style="list-style-type: none"> We expect evidence will come from interviews only. 		<ul style="list-style-type: none"> ■ 	
<p>11. How (if at all) does contracting out encourage innovation in provision of services?</p>	<ul style="list-style-type: none"> We will review the literature to establish what worked or not and why. We will seek opinions from interviewees of how service innovation could be better facilitated when contracting out. 	<ul style="list-style-type: none"> ■ 	<ul style="list-style-type: none"> ■ 	<ul style="list-style-type: none"> ■

Annex 2: Search protocol

Databases

The search was carried out on:

- **Web of Science** (screening all results) and
- **Google Scholar** (screening the first 30 results) in order to identify relevant literature.

Search terms

We used a Boolean search string to ensure the relevance of results where the search functionality supported it⁸⁷⁹ (see below). We have drawn upon the search strings suggested in our proposal, refined in consultation with our expert team, and through testing.

Web of Science⁸⁸⁰
TS= (((("contract*" OR "outsourc*" OR "commission*" OR "payment by results") AND ("employment")) AND ("advantage*" OR "disadvantage*" OR "benefit*" OR "drawback*" OR "cost*")) AND ("UK" OR "United Kingdom" OR "England" OR "Scotland" OR "Wales" OR "Northern Ireland" OR "OECD")) ⁸⁸¹
Google Scholar advanced search
With all of the words: employment services
With at least one of the words: contract commission outsourc
Where my words occur: anywhere in the article (about 1,220,000 results)
Where my words occur: in the title of the article (53 results)
Return articles dated: 2001 – (onwards)

Search results were downloaded and combined with the literature identified in the preliminary search. After removal of duplicate studies, the titles and abstracts were screened. Full papers of potentially relevant studies were retrieved and a final judgement on eligibility will be made by RAND Europe researchers.

Data from relevant studies were extracted into a standardised template (see below).

All search results were screened by members of the research team against the inclusion/exclusion criteria (see below). Screeners consulted with each other during the early stages of the review to ensure consistency in the process.

⁸⁷⁹ The search terms differ due to the different functionalities that each search engine allows. Google Scholar allows far fewer terms to be included and does not support a Boolean search.

⁸⁸⁰ <https://www.webofscience.com/wos/woscc/summary/ce3a4387-8977-46aa-a2a3-62030bb4d6a3-19453c10/relevance/1>

⁸⁸¹ The inclusion of "OECD" facilitates the inclusion of literature on international contracting out without risking inclusion of conflict, post-conflict or otherwise fragile governments which are not relevant for comparison.

Inclusion and exclusion criteria

Included in the review	Excluded from the review
Academic articles and grey literature, including government documents or reports	Commentaries, editorials and features
Published in English	Published in languages other than English
Articles published after 2001 ⁸⁸²	Published prior to 2001
Research about the UK or other OECD country	Research about non-OECD countries

Snowball search

The bibliographies of studies identified via the search of databases that met the inclusion criteria and retained for full-text review were searched for potentially relevant sources. RAND Europe researchers compiled a list of all potentially relevant sources and took the same steps as those for the results of the database search, i.e. application of the above inclusion and exclusion criteria, title and abstract screen, full-text review and data extraction.

Targeted hand search

A targeted search included resources cited in the tender specifications that set out the principles for Government contracting⁸⁸³ and review Government outsourcing practices and results,⁸⁸⁴ including in employment contracts,⁸⁸⁵ as well as international evidence.⁸⁸⁶ In addition, RAND Europe researchers searched websites of the Institute of Labour Economics (IZA),⁸⁸⁷ Organisation for Economic Co-operation and Development (OECD),⁸⁸⁸ and Knowledge Centre of Public Employment Services (PES)⁸⁸⁹ for resources on contracting out (employment and health) services. The above inclusion/exclusion criteria were also applied in the targeted hand search.

Search results

Database	Number of results
Web of Science	271

⁸⁸² Could potentially be changed to 2011, depending on number and relevancy of search results.

⁸⁸³ HM Government (2021). 'The Sourcing Playbook. Government guidance on service delivery, including outsourcing, insourcing, mixed economy sourcing and contracting.' As of 24.11.2021: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/987353/The_Sourcing_Playbook.pdf

⁸⁸⁴ Sasse, T., Guerin, B., Nickson, S., O'Brien, M., Pope, T. and Davies, N. (2019). 'Government outsourcing: What has worked and what needs reform?' Institute for Government. As of 24.11.2021: <https://www.instituteforgovernment.org.uk/sites/default/files/publications/government-outsourcing-reform-WEB.pdf>

⁸⁸⁵ Department for Work and Pensions (2020). 'The Work Programme – A quantitative impact assessment.' As of 24.11.2021: <https://www.gov.uk/government/publications/the-work-programme-impact-assessment>

⁸⁸⁶ Stephan, G. (2016). Public or private job placement services—Are private ones more effective?. *IZA World of Labor*. 285 doi: 10.15185/izawol.285

⁸⁸⁷ <https://www.iza.org/>

⁸⁸⁸ <https://www.oecd.org/>

⁸⁸⁹ <https://ec.europa.eu/social/main.jsp?catId=1163&langId=en>

Google Scholar	53
Snowballing	40
Tender specifications	4
IZA, OECD, PES Knowledge Centre	59
TOTAL	427

Total number of search results: 427

Total number of results selected for screening after the removal of duplicates: 424

Total number of studies retained after title and abstract screening: 78

Total number of studies included in the review: 78

Data extraction template

Category	Description	Source 1
Corresponding specific question		
General information	Database	
	Short reference	
	Full reference	
	Abstract	
	Countries covered	
	Years covered	
Contract design	Benefits of PbR (incentives to maintain quality, achieve outcomes)	
1(3)		
1(3)	Costs of PbR (incl. increased overhead, perverse incentives, loss of goodwill)	
1(2)	Managing risks in contracts (types of risks, ways of assessing risks, transfer of risk)	
1(2)	Impact of uncertainty on volume and performance (incl. setting prices and targets)	
1(4)	Level of detail in contract specification ('black box', 'grey box', cost effectiveness, opportunity for innovation)	
2	Implications of using a prime/sub model	
3	Implications of geographical division (incl. competition, economies of scale)	
4	Length of contracts	

Effective contracting of employment and health services

Market competition		
5	Selecting contractors (questions, criteria and scoring)	
5	Effects of contracting on the market (competition, entry/adaptation/exit, efficacy, efficiency)	
6(1)	Effects of contract type (incl. a pricing cap) on market competition in the short- and long-term	
6(2)	Trade-offs between optimising short- and long-term outcomes	
Live running		
7	Measuring service quality and customers standards	
8	Incentives for collaboration and sharing resources among contractors	
9	Sharing resources with other lines of business	
10	Contractors' use of knowledge and expertise	
11	Facilitating service innovation	
Other		

Annex 3: Interview topic guide

Introduction/background

1. Could you please tell me about your organisation and your role?
2. Could you please ***briefly*** tell me about your experience with contracting (public) services?

The discussion today will focus on different models of commissioning public services. These are referred to as pay for performance schemes, payment by results, payment by outcomes.

3. Which model of (PbR) contracting are you most familiar with?
4. Which health assessment contracts are you familiar with?

Contract design

5. In your experience, what benefits are associated with different PbR contracts?
6. Based on your research, what costs are associated with different PbR contracts?
7. What other implications are associated with different PbR contracts?
8. What risks are associated with different PbR contracts?
9. Drawing on your experience, how do commissioners assess the risks in different PbR competitions?
10. How (if at all) did the uncertainty on (labour) market outlooks and expected volume of clients affect provisions in this contract?
11. How (if at all) did the uncertainty on contractors' performance affect provisions in this contract?
12. To what extent are pricing caps used in different PbR contracts?
13. Contract specification can detail closely how the service should be provided, or it can leave some or no detail at all ('black box' provision). What in your view is the optimal level of specifications in PbR contracts?
14. What in your view is the optimal length of a PbR contract?
15. CostPlus employment contracts were introduced during the pandemic. If you have any experience or knowledge of these, what aspects of these contracts would be useful for new contract designs in the future?
16. Services are often contracted out using a prime/sub model. What, in your opinion, are the positive or negative implications of using this model?
17. In the UK, providers have submitted bids for Contract Package Areas which divide service provision up geographically. What is your knowledge or experience with this strategy of geographical division?

Market competition

18. A variety of criteria, scoring mechanisms and questions can be used at the tendering and bidding stages to select contractors. For this contract competition, what was involved in this selection process?
19. How did this contract's provisions affect service providers over time?
20. This competition and contract provisions aimed to strike a balance between securing the 'value-for-money' for the commissioner in the short term and maintaining providers' capacity or goodwill in the long term. How well (if at all) has this balanced been achieved, in your view?

Live running

21. Based on your experience, how is service quality measured in different contract models?
22. What (if any) incentives does/did this contract contain to collaborate or share information with other contractors?
23. In your experience or knowledge, have you been aware of contractors sharing resources with each other?
24. Have you come across contractors sharing resources with other stakeholders in public service provision?
25. In your view, what specialist knowledge and expertise do contractors bring to the provision of these services, if any?
26. Commissioners may look to balance necessary specification with an aim to incentivise innovation and new approaches to the provision of services. How, if at all, was/is innovation encouraged in this contract/PbR models?

Conclusion, thanks and wrap up

27. Is there anything else that you would like to add, that we haven't mentioned so far?
28. Is there anyone you recommend that we should also try to talk to as part of this study?
29. Would you be available to participate in a workshop which we will be organising?
30. Do you have any questions for us?

Thank you very much for your time!

Annex 4: Synthesis workshop

List of workshop participants

Providers (4) – names withheld to ensure anonymity

Academics and researchers (3) – names withheld to ensure anonymity

Commissioners and other DWP staff (11) – names withheld to ensure anonymity

RAND Europe (6):

- Dr Chris van Stolk
- Michael Whitmore
- Dr Julia Doyle
- Jessica Dawney
- Dr Madeline Nightingale
- Joanna Hofman

Agenda

- | | |
|--------------------|--|
| 09.30-09.40 | Welcome and purpose of the workshop |
| 09.40-10.05 | Presentation of the study and early findings |
| 10.05-10.30 | Breakout groups |
| 10.30-10.35 | <i>Break</i> |
| 10.35-11.20 | Plenary: summary and discussion |
| 11.20-11.30 | Wrap up, next steps and thanks to participants |

Breakout groups topics and facilitation questions

Group 1: Costs and benefits associated with PbR contracts

- Is there a need for more/different research to better understand the benefits & costs of contracting out and the trade-offs involved?
- How can competitions/contracts be designed to enable DWP to benefit more fully from providers' specialist expertise and knowledge?
- What provisions in future competitions/contracts would reasonably limit costs for DWP or providers? What trade-off would these require?

- How can we help small/specialist providers better access the market within a prime model?

Group 2: Level of specification: balancing the 'black box' with minimum service requirements

- How – if at all – do minimum service requirements affect: (i) innovation? (ii) costs and pricing?
- What risks do minimum service requirements bear for providers?
- How are the minimum service requirements examined and how this can be improved?
- How are the minimum service requirements linked to performance management?

Group 3: Incentives to encourage collaboration, sharing information or resources amongst providers

- What incentives for collaboration could be trialled? How would these work, what are risks & costs involved?
- What strategies for incentivising collaboration have worked elsewhere (also in other contexts)?
- What are the different impacts of collaboration groups/schemes if they are optional or compulsory?
- What barriers have there been to sharing resources to best serve the clients (how can they be overcome)?
- How do we ensure that collaboration/sharing is fair and mutually beneficial?

Annex 5: Methodological challenges associated with estimating the costs and benefits associated with PbR contracts

Empirical analyses of the effects (costs/benefits) of contracting out provision face certain methodological challenges. The primary challenge is establishing a meaningful counterfactual to illustrate what would have occurred to participants if they have not participated in the programme, or if they had participated in a different programme.⁸⁹⁰

Some studies seek to estimate differences in outcomes (e.g., cost effectiveness, job outcomes) achieved by the PES and contracted providers. However, programmes delivered by the PES and contracted providers may not be fully comparable. Differences in outcomes may reflect variation in programme design and delivery rather than the type of provider (public, private, non-profit). The profile of participants may also differ, further complicating comparisons.

Even if the programmes are comparable, it is difficult to establish a counterfactual due to the risk of unobserved heterogeneity/selection bias.^{891,892,893} There may be differences in the unobserved characteristics of jobseekers who participate in programmes delivered by public/private (or non-profit) providers. In addition, contracting out may be more common in sectors or regions that have particular characteristics that affect programme cost and quality, confounding comparisons.

A number of studies use randomisation to ensure comparability between programmes. However, randomisation does not necessarily eliminate the risk of selection bias. Behaghel et al (2012)⁸⁹⁴ describe an RCT where randomisation was implemented in terms of which programme was *offered* (public or private intensive job assistance versus a control group). Participation was voluntary, uptake relatively low (less than half of those assigned to a programme entered it) and higher for private agencies than the PES, meaning that there could be differences in the

⁸⁹⁰ Langenbucher & Vodopivec (2022).

⁸⁹¹ Andersson, F., Jordahl, H., and Josephson, J. (2019). Outsourcing Public Services: Contractability, Cost, and Quality. IZA Discussion Paper Series. No 12401.

⁸⁹² Ayaita, A. et al, 2021, 'Job Placement via Private vs. Public Employment Agencies: Investigating Selection Effects and Job Match Quality in Germany', IZA Discussion Paper No.14024

⁸⁹³ Cockx, B., & Baert, S. (2015). Contracting Out Mandatory Counselling and Training for Long-Term Unemployed. Private For-Profit or Non-Profit, or Keep it Public? IZA DP No. 9459

⁸⁹⁴ Behaghel, L. et al, 2012, 'Private and Public Provision of Counselling to Job-Seekers: Evidence from a Large Controlled Experiment', IZA Discussion Paper No.6518

composition of the two groups of job seekers stemming from their willingness to sign up.

Methods used in observational studies to address the risk of unobserved heterogeneity include instrumental variable models, regression discontinuity design, or difference-in-difference estimates.⁸⁹⁵ Other studies use propensity score matching to maximise the comparability of the groups whose job outcomes are being compared.⁸⁹⁶ Selection bias can be addressed to some degree by using panel data to control for unobserved heterogeneity.^{897,898}

⁸⁹⁵ Langenbucher & Vodopivec (2022).

⁸⁹⁶ Langenbucher & Vodopivec (2022).

⁸⁹⁷ Andersson, F., Jordahl, H., and Josephson, J. (2019). Outsourcing Public Services: Contractability, Cost, and Quality. IZA Discussion Paper Series. No 12401.

⁸⁹⁸ Koning, P. & Heinrich, C. J., (2010), "Cream-Skimming, Parking and Other Intended and Unintended Effects of Performance-Based Contracting in Social Welfare Services", IZA DP No. 4801