



Making Markets in Employment Support:
Promises and Pitfalls in the Work Programme's
Private Power Market

Eleanor Carter

Department of Geography

June 2018

Submitted in partial fulfilment of the degree of Doctor of Philosophy

Acknowledgements

The heartiest of thanks go firstly to Adam Whitworth. The time, thought and effort that Adam has dedicated to both me and this PhD project have been immense and I am hugely grateful for his support. From my first inklings of interest in undertaking doctoral research Adam has been a tremendous source of encouragement and advice. I owe Adam a deep intellectual debt: his insights have been crucial in helping me work through ideas and approaches. Beyond this, Adam's persistent backing and reassurance has been a crucial underpinning to a new-found confidence in my own ideas and abilities.

My second supervisor Dan Vickers also played a formative role in sharpening my analytical skills. I am grateful for the support of other friends and colleagues across the Geography Department at the University of Sheffield and DWP teams based at Rockingham House. I would also like to thank the Economic and Social Research Council who funded the project through an advanced quantitative methods studentship.

I owe much to my friends and family. My parents in particular are key culprits in nurturing my inquisitiveness and determination. Thank you for your generosity, support, and care. Thanks also to my sister Rosie, who has reminded me throughout that it is important to have fun and who has provided plenty of entertaining distractions.

Finally, to Mark McCombs whose quiet and sturdy support I have cherished through it all. Thank you.

Abstract

Welfare-to-work services have been a key area of experimentation in quasi-marketised public service delivery. The British flagship Work Programme is seen as an international pioneer in its reliance on outsourcing, payment by results and provider flexibility allied to promises of innovation and performance improvement. Within schemes dominated by such marketised accountabilities there are well-known risks and tensions around creaming, parking and churning. International literature equally makes clear that the design specificities of programme governance and accountabilities can play a key role in either facilitating or buttressing against these negative provider practices.

In this context, the overarching question which animates this thesis is whether this crafted Work Programme design structure is sufficient and appropriate to steer its quasi-marketised providers to the achievement of the full suite of government policy objectives. Unprecedented academic access to the commissioning Department's administrative datasets alongside sophisticated and conceptually tailored multivariate quantitative analyses underpin the thesis' empirical contributions. The analysis is framed by an original multi-dimensional analytical framework articulating multiple potential alternative *types* of quasi-markets. This conceptually broad and empirically focused study provides a rare opportunity to trace outcomes directly from the plans and promises of a particularly bold quasi-market experiment and to consider the ways in which key design elements cascade through to, and are detectable in, the patterning of employment and earning outcomes of programme participants on the ground. The empirical analyses highlight myriad ways in which Work Programme promises end up in performance pitfalls despite, if not because of, its particular variety of quasi-marketised governance.

Table of Contents

Acknowledgements	3
Abstract	4
1 Introduction: making markets in employment support	21
1.1 Fracturing the understanding of marketisation in public services	22
1.2 British welfare-to-work in context.....	23
1.3 The Work Programme.....	25
1.4 Perennial risks and hubristic design promises: the research focus and chapter summaries	26
2 Making (quasi) markets in the welfare state	31
2.1 Chapter summary	31
2.2 Shifting modes of governance in public service delivery	31
2.3 Governance as a grounding concept	33
2.4 The rise and rise of market governance and accountabilities	37
2.4.1 Public service markets as ‘quasi’-markets	39
2.5 Fronting and fracturing understandings of market governance: the difference thesis	40
2.6 A differential view of markets	42
2.6.1 Dimensions of variation in public service quasi-markets – allocation	44
2.6.2 Dimensions of variation in public service quasi-markets – production	45
2.7 Varieties of quasi-market: where non-market accountabilities overlay marketised governance.....	47
2.8 Network accountabilities – the missing dimension.....	49
2.9 Conclusions and charting a course to operationalise the analytical framework	51
3 Evolution of the “welfare-to-work” agenda: shifting policy goal posts and applying new quasi-market approaches	54

3.1	Chapter summary	54
3.2	Policy trends in unemployment and “worklessness”	54
3.3	The activation turn	56
3.4	Beneath the banner of activation	57
3.5	British activation objectives	58
3.6	Carrots, sticks, or a bit of both? Activation trends and the rise of ‘welfare-to-work’ in the UK	60
3.7	Era 1: Tory beginnings (1989 – 1997)	61
3.8	Era 2: New Labour’s Third Way (1997 – 2010): ‘creeping conditionality’ and quasi-markets.....	61
3.9	Era 3: Conservative rocket boosters (2010 – 2018): enter the Work Programme	70
3.9.1	Introducing the Work Programme.....	70
3.9.2	The particularities of the quasi-market configuration embodied by the Work Programme: the allocation dimension.....	72
3.9.3	The particularities of the quasi-market configuration embodied by the Work Programme: production dimension	74
3.10	Work Programme as a private power market and anticipated challenges for this quasi-market form	76
3.11	Work Programme design as a moderator of gaming?	77
3.11.1	Work Programme design innovations: Differential payment groups.....	82
3.11.2	Work Programme design innovations: Large, regional contracting areas with specified minimum performance levels and market share shift.....	83
3.11.3	Work Programme design innovations: Sustainment payments for longer-term employment	84
3.12	Are these bold design promises kept?	85
4	Data and Methods.....	90
4.1	Chapter summary	90
4.2	Introduction to the research approach	90
4.3	Introduction to underlying datasets – an exciting in-house opportunity	93

4.3.1	National Benefits Database (NBD).....	94
4.3.2	Work Programme Analytical Database (WPAD)	95
4.3.3	Labour Market System (LMS) meetings, opportunities, and ‘client’ characteristics.....	95
4.3.4	P45/6 Employment data	96
4.3.5	Additional datasets capturing employment and earnings – Real Time Information sample	97
4.3.6	Additional external datasets to capture local economic context	99
4.4	Bringing the datasets together.....	100
4.5	Investigating the promise of differential payments: methods used in Chapter 5	101
4.6	Investigating the geography of incentives and performance: methods used in Chapter 6.....	102
4.7	Exploring the employment and earning trajectories of Work Programme participants: methods used in Chapter 7	103
4.7.1	Construction and coding of earnings trajectories for the sequence analysis	104
4.7.2	Measuring the distance between earning trajectories in the sequence analysis	107
4.7.3	Grouping similar trajectories together	110
4.8	Comparing employment and earnings outcomes for participants in Work Programme and Work Choice: methods used in Chapter 8.....	110
4.8.1	Identifying the counterfactual through propensity score matching..	112
4.8.2	Benefits of Propensity Score Matching compared to other available methods	117
4.8.3	Propensity Score Matching in Summary.....	117
4.9	Research considerations and positionality	117
5	Equality through inequality: pursuing differentiated universalism through the Work Programme payment structure	120
5.1	Chapter summary	120

5.2	Reconciling Work Programme objectives: the four E's of the differentiated universalism tightrope	120
5.3	A troubling pursuit? International experience in relation to the attainment of differentiated universalism in quasi-marketised welfare-to-work systems.....	124
5.4	Why would the Work Programme be any different? Expectations of programme design for the pursuit of differentiated universalism	125
5.5	Findings: Differential payments but <i>still</i> differential outcomes across payment groups	128
5.6	Moving beyond official 'payment groups'	132
5.7	Concealed characteristics: extending the analysis of differentiated universalism beyond payment groups	133
5.8	Getting personal	134
5.8.1	Benefit history in period prior to Work Programme	137
5.8.2	Health and disability.....	138
5.8.3	Demographics.....	140
5.8.4	Employment history	141
5.8.5	Household Characteristics.....	142
5.8.6	Local geographical context	143
5.8.7	Summary of model findings and implications for differentiated universalism	143
5.9	Considering the differentiated payment system as a 'model': possibilities for enhanced profiling in the UK context	144
5.10	An alternative payment system? – moving beyond crude payment groups 149	
5.11	Chapter summary	151
6	Jeopardising geography? The spatial patterning of Work Programme performance.....	154
6.1	Chapter summary	154
6.2	Reprising the spatial logic of the Work Programme's private power market 154	

6.3	The geography of Work Programme incentives: contracts to tempt international providers versus meaningful local labour markets	156
6.4	From paper to practice: is there spatial variation in performance?	160
6.5	Understanding contract performance variation or routes to ‘winning’ within the Work Programme incentive structure	164
6.6	Predictable patterns of success? Mapping Work Programme performance at local authority level.....	167
6.7	Making sense of the variation	174
6.8	Putting people in their place.....	179
6.9	Discussion: controlling for composition and context.....	180
7	Crafting incentives for sustained employment: Investigating employment and earning trajectories for Work Programme participants	186
7.1	Chapter summary	186
7.2	Introduction: The evolution of payable ‘outcomes’ within welfare-to-work payment-by-results systems	186
7.3	The low-pay no-pay cycle: a background to unstable employment outcomes 188	
7.4	Coherent incentives? A complex response from designers in steering for sustainment.....	191
7.4.1	Anticipated response to sustainment incentives – the ‘known unknown’ in Work Programme design	192
7.5	Putting employment trajectories in context: common earning patterns for Work Programme participants	194
7.5.1	The gulf between ‘job entry’ and ‘job outcomes’	195
7.6	Understanding employment and earnings over time.....	196
7.7	A typology of employment and earnings trajectories amongst Work Programme participants	201
7.7.1	Super-cluster 1: ‘slow and unstable’	201
7.7.2	Super-cluster 2: ‘stuck in low-pay’	203
7.7.3	Super-cluster 3: ‘general earnings success’	203

7.7.4	Super-cluster 4: ‘gold standard’	205
7.8	Reflections and considerations on employment and earning trajectories	206
7.9	Policy implications and challenges in predicting cluster membership	207
7.10	Conclusions	209
8	Making markets in employment support: does the variety of quasi-market matter for people with disabilities and health conditions?	212
8.1	Chapter summary	212
8.2	Introduction: the uncomfortable gulf in job outcome rates for people with health conditions and disabilities	212
8.3	The particularities of the quasi-market configuration embodied by the Work Programme and Work Choice	214
8.3.1	Applying dimensions of variation as a comparative tool: <i>variation on the allocation dimension</i>	214
8.3.2	Applying dimensions of variation as a comparative tool: <i>variation on the production dimension</i>	215
8.4	Pure provider-directed or hybrid forms of quasi-market?.....	219
8.5	An empirical exploration of different quasi-market configurations for employment and earning outcomes.....	220
8.5.1	Propensity scores and matching algorithm.....	225
8.5.2	Quality of the ‘match’	226
8.6	Findings: Employment and earning effects across the two programmes.	228
8.7	Discussion and conclusions.....	233
9	Rest In Peace Work Programme: Epitaph for a bold quasi-market experiment	236
9.1	Beginnings of the eulogy.....	236
9.2	Conceptual and analytic contributions	236
9.2.1	Critical reflections and limitations of the conceptual framework	239
9.3	Empirical contributions	244
9.4	Policy contributions.....	248

9.4.1	Stepping out of the private power market?	250
9.5	Burned, bruised or broken? Future policy and research directions.....	250
	References	255
Appendix 1	273
	Employment and benefit spell data cleaning for the production of variables on benefit claiming and employment histories	273
Appendix 2	276
	Preparation of administrative data	276
Appendix 3	284
	Full logistic regression output underpinning Chapter 5.....	284

Figure 1.1 Key British welfare-to-work arrangements 2011-2018	24
Figure 2.1 Governance and accountability types, adapted from Considine, 2001 and Carter, Forthcoming	35
Figure 2.2 Gingrich’s six quasi-market types. Source: Gingrich, 2011, p. 12.....	43
Figure 2.3 Visualising the allocation dimension in public service markets (an original illustration informed by and extending from Gingrich, 2011 and Wiggan 2015a, 2015b).....	44
Figure 2.4 Visualising the production dimension in public service markets (an original illustration informed by and extending from Gingrich, 2011 and Wiggan 2015a, 2015b).....	46
Figure 2.5 Overlaying non-market accountabilities on the production dimension	48
Figure 2.6 Considering horizontal network accountabilities: a hypothesised axis of variation introducing collaboration across multiple, interconnected provider organisations.....	51
Figure 3.1 Shifting functions of the welfare state. Source: extending from Peck and Jones, 1995	55
Figure 3.2 key features of activation ideal types. Reproduced from Carter and Whitworth, 2017.....	58
Figure 3.3 Available policy tools to activating labour market policy regimes. Reproduced from (Weishaupt, 2011, p. 69). Categories of NIF, NIN, PIF and PIN are used subsequently.....	60
Figure 3.4 Three welfare-to-work eras and the nature of key activating policy tools	66
Figure 3.5 Visualising the allocation dimension for the Work Programme quasi-market	72
Figure 3.6 Visualising the production dimension for the Work Programme quasi-market.....	75
Figure 3.7 Work Programme objectives and associated incentive design instruments	80
Figure 3.8 Work Programme Payment Groups and payment levels. Reproduction of previous work published in Carter and Whitworth, 2015	83
Figure 3.9 Illustration of Work Programme payment model, reproduced from DWP, 2010b, p. 11	85

Figure 3.10 The payment tail for the Work Programme: Duration of sustained employment required to trigger job outcome payments and maximum extent of sustainment payments	85
Figure 4.1 Contributing datasets used in the analysis	94
Figure 4.2 Connections across multiple administrative and externally held datasets	100
Figure 4.3 Data sources, sample size and methods used in each of the empirical chapters	101
Figure 4.4 Earnings trajectories as a sequence of distinct elements	105
Figure 4.5 Definition of element types used in sequence analysis	106
Figure 4.6 Optimal Matching operations	107
Figure 4.7 Weighting socioeconomic proximity. The ‘distance’ from no earnings (Z) to any earning level is given a higher cost than transitions between the earning states	109
Figure 5.1 Work Programme Payment Groups and payment levels. Reproduction of previous work published in Carter and Whitworth, 2015	126
Figure 5.2 Different job outcome rates between Work Programme payment groups	129
Figure 5.3 Divergent job outcome rates between Work Programme payment groups across the implementation period. Source: DWP Administrative data. Note Payment group 8 is not plotted due to insufficient population size in month-to-month cohorts for the production of percentage figures	130
Figure 5.4 Variation in job outcome rates for sub-groups within payment group 3. Source: DWP administrative data	132
Figure 5.5 Participants with disabilities and/or health conditions within the Work Programme payment groups: Source: DWP administrative data.....	133
Figure 5.6 Variation in job outcome rate across observable participant characteristics in publicly available data. Source: DWP administrative data.....	134
Figure 5.7 Explanatory variables used in logistic regression analysis.....	136
Figure 5.8 Predicting the achievement of job outcomes through administrative data: benefit history. Source: DWP administrative data.....	138
Figure 5.9 Predicting the achievement of job outcomes through administrative data: health and disability characteristics. Source: DWP administrative data.....	139
Figure 5.10 Predicting the characteristics of job outcomes through administrative data: demographic characteristics. Source: DWP administrative data	141

Figure 5.11 Predicting the achievement of job outcomes through administrative data: Employment history and household characteristics. Source: DWP administrative data	142
Figure 5.12 Predicting the achievement of job outcomes through administrative data: local geographical context. Source: DWP administrative data coupled with externally prepared claimant rate and IMD figures	143
Figure 5.13 Exploring the achievement of job outcomes through administrative data: distribution of participant predicted probabilities. Source: DWP administrative data	145
Figure 5.14 Predicting the achievement of individuals' job outcomes: a comparison of alternative models. Outcome variable is the achievement of the official programme job outcome	146
Figure 5.15 Claimant segmentation and predictive accuracy	148
Figure 5.16 From predicted probabilities to an improved differential payments system. Source: DWP administrative data	150
Figure 6.1 Work Programme Contract Package Areas and Prime providers.	157
Figure 6.2 Job outcome rates within alternative spatial scales: CPA, contract and local authority	161
Figure 6.3 CPA-level job outcome rates, descriptive statistics (N = 18).....	161
Figure 6.4 Job outcome performance gap between contracts within each CPA	162
Figure 6.5 Description of Local Authority job outcome performance variation (N = 378).....	163
Figure 6.6 Equivalent minimum performance benchmarks, reproduced from CESI, 2013	164
Figure 6.7 Within Prime contract local area performance variation for payment group 1	165
Figure 6.8 Within Prime contract local area performance variation for payment group 2.....	166
Figure 6.9 Within Prime contract local area performance variation for payment group 3.....	166
Figure 6.10 Within Prime contract local area performance variation for payment group 6.....	167
Figure 6.11 Local authority deprivation across Britain, standardised indicator sourced from CRESR and produced as described in Abel et al., (2016).....	169

Figure 6.12 Proportion of payment group 1 participants achieving job outcomes at local authority level.....	170
Figure 6.13 Proportion of payment group 2 participants achieving job outcomes at local authority level.....	171
Figure 6.14 Proportion of payment group 3 participants achieving job outcomes at local authority level.....	172
Figure 6.15 Proportion of payment group 6 participants achieving job outcomes at local authority level.....	173
Figure 6.16 Association between job outcome rates and i) local authority deprivation; and ii) proportion of working age population participating in the Work Programme, for payment groups 1 and 2.....	175
Figure 6.17 Association between job outcome rates and i) local authority deprivation; and ii) proportion of working age population participating in the Work Programme, for payment groups 3 and 6.....	175
Figure 6.18 Exploring local authority Work Programme performance. Outcome variable is job outcome rate *** $p < 0.001$	177
Figure 6.19 Variation in local authority payment-by-results 'success' by deprivation and proportion of programme participants amongst the working age population ...	178
Figure 6.20 Exploring variation in local authority payment-by-results 'success' by deprivation and proportion of programme participants amongst the working age population. Outcome variable is proportion of potential PbR achieved	179
Figure 6.21 Plotting job outcome rates for 'people in place'	180
Figure 6.22 Alternative measures of contract performance: a 'Contextual Value Added' for Work Programme Prime providers.....	184
Figure 7.1 Illustration of Work Programme payment model, reproduced from DWP, 2010b, p. 11	191
Figure 7.2 Earnings profile for participants attached to Work Programme between April 2013 and September 2014, final column presents overall percentage breakdown	195
Figure 7.3 Definition of distinct labour market elements within participant employment trajectories	197
Figure 7.4 Investigating employment and earnings pathways: Comparing cohort-level representation to sequence index plot for Work Programme participants	198
Figure 7.5 Earnings trajectories as a sequence of distinct elements	198
Figure 7.6 Optimal Matching operations	199

Figure 7.7 Identifying the optimal cluster solution: Caliński-Harabasz pseudo-F index for Ward’s method	200
Figure 7.8 Ward’s linkage identifies 9 clusters (blue) within 4 ‘super-clusters’ (grey)	200
Figure 7.9 Sequence index plot for super-cluster 1: ‘slow and unstable’ earning trajectories	201
Figure 7.10 Sequence index plot for super-cluster 2: ‘stuck in low-pay’ earning trajectories	203
Figure 7.11 Sequence index plot for super-cluster 3: ‘general earnings success’ earning trajectories	204
Figure 7.12 Sequence index plot for super-cluster 4: ‘gold standard ’ earning trajectories	205
Figure 7.13 Cohort representation of earnings trajectory clusters for Work Programme participants	206
Figure 7.14 Employment and earnings trajectory clusters and relative size	206
Figure 7.15 Demographic characteristics of programme participants within trajectory groups	208
Figure 8.1 Job outcome rates for Work Programme participants over time, by self-reported health condition and disability status	213
Figure 8.2 Visualising the allocation dimension for the Work Programme and Work Choice quasi-markets	215
Figure 8.3 Visualising the production dimension for the Work Programme and Work Choice quasi-markets	216
Figure 8.4 Contractually specified modular intervention structure of Work Choice (adapted from Purvis et al., 2013)	218
Figure 8.5 High-level programmatic comparison of Work Programme and Work Choice.....	221
Figure 8.6 Logistic regression model to predict Work Choice participation	224
Figure 8.7 Distribution of propensity scores for Work Choice and Work Programme participants	225
Figure 8.8 Propensity scores of Work Choice ‘treatment’ and matched ‘Work Programme ‘control’ group	226
Figure 8.9 Comparison of baseline and post-match characteristics across treated and untreated groups	227
Figure 8.10 Employment effect of Work Choice compared to Work Programme ..	229

Figure 8.11 Graphical representation of earning trajectories for matched Work Programme and Work Choice participants	230
Figure 8.12 Super-cluster 1: slow, chequered and unstable.....	231
Figure 8.13 Super-cluster 2: stuttering transitions and low-pay	231
Figure 8.14 Super-cluster 3: quick entry to stable employment	232
Figure 8.15 Work Programme and Work Choice participant presence within each of the earning trajectory clusters	232
Figure 9.1 Notional positioning of programme markers against allocation axes	240

1 Introduction: making markets in employment support

A fundamental shift has taken hold of welfare state governance in developed economies since the 1980s with the increasingly widespread use of privatisation and market-like methods for the delivery of public services. ‘Bureaucracy’ has become a pejorative term with few advocates and many critics and a path has been charted that seeks to introduce an ‘enterprising’ spirit to social provision (Considine, 2001; Fournier and Grey, 1999; Gingrich, 2011). Government purposes are increasingly deployed through an outsourced dynamic in which an ‘enabling state’ (Deakin and Walsh, 1996) becomes a purchaser and overseer of services rather than their provider (Le Grand and Bartlett, 1993). This increased use of market-like methods is justified by claims that markets increase efficiency, efficacy and service responsiveness (Le Grand and Bartlett, 1993; Osborne and Gaebler, 1992), although the evidence for such claims in the sphere of complex services is equivocal at best (de Graaf and Sirovátka, 2012; Warner and Hefetz, 2012).

In the field of (un)employment policy specifically, in parallel to these governance shifts there has been an important recasting of policy objectives with countries across the OECD making decisive shifts from a ‘passive’ to an ‘active’ welfare system in which eligibility for out-of-work social security benefits is tied increasingly tightly and explicitly to the stated obligation to seek paid work. Unemployment is no longer framed as a structural issue of demand management but has been reconfigured as a supply-side challenge of improving the employability of the unemployed. This decisive turn has ushered in an ‘activation orthodoxy’ based on a programme of policies which aim to propel working-age welfare recipients (back) into the labour market as quickly as possible through a combination of applying more stringent conditions to benefit receipt, improving financial returns to work, enhancing employability, and providing in-work support (Bonoli, 2010; Peck and Theodore, 2000). This is then a double dynamic where both citizens who are in receipt of out-of-work benefits and the system of service provision itself are being ‘activated’ through new tools and logics.

Throughout the thesis the term ‘activation’ is used to capture the overarching governmental ambition to move people away from ‘inactively’ claiming unemployment benefits and instead towards ‘active’ contributions through paid employment in the labour market. More specifically, ‘welfare-to-work’ and ‘employment support’ are terms used interchangeably to connote programmatic services and interventions directed at people who are unemployed to support them into employment.

At the nexus of these two trends – marketisation and activation – a small but rich seam of comparative literature explores the trends of governance reforms and activation policies and considers the influence of governance on the dynamics of national

activation approaches (Lødemel and Moreira, 2014; van Berkel et al., 2012; van Berkel and Borghi, 2007; Wiggan, 2015a, 2015b). Analytically, marketisation sits beneath an umbrella of wider governance configurations and options – procedural, corporate, market and network for Considine (2001) – of which marketisation has been a particularly potent trend internationally in recent decades. This literature acknowledges a distinction between the ‘substantial’ formal content of policy and the ‘operational’ level of governance – that is, the mode of coordination underpinning the provision of services – but considers these forms to be mutually constitutive and dependent (Carmel et al., 2003; de Graaf and Sirovátka, 2012; van Berkel et al., 2012). The result is a key discernment: that governance reforms cannot be without consequence for the content and effects of activation. This acknowledgement in turn sparks an important impetus both for this thesis and the wider field given that the logics of particular governance forms (and accountability levers) will inform the nature and content of social policies as well as having important (and heterogeneous) implications for service user experiences and outcomes. And if we wish to know and improve the latter, research must deeply engage with the former. In short, governance matters.

1.1 Fracturing the understanding of marketisation in public services

Alongside the pervasive application of market-like principles to the organisation of public services since the 1980s there is a broad acknowledgement that the term ‘marketisation’ is descriptively inadequate, indeed inaccurate, given that markets in public services are in important ways artificial *quasi marketised* constructs which diverge from a conventional understanding of market functioning (Bartlett and Le Grand, 1993; Le Grand and Bartlett, 1993).

Less widely acknowledged within this, however, is the degree to which a common language of ‘market’ reform has obscured the extent to which the process of market making in the sphere of public services in reality refers to a highly varied set of *quasi-market* forms and reform processes (Jacobs, 1998; Greener, 2008). The homogenised conception of a singular marketising endeavour has been powerfully deconstructed by advocates of ‘the difference thesis’ (Powell, 2015; Gingrich, 2011; Zehavi, 2012; Meagher and Goodwin, 2015). Briefly, this cluster of academic work suggests that ‘marketisation’ is not a singular phenomenon but rather that it captures a diverse range of “practices, rationales, trajectories, actors and impacts” (Meagher and Goodwin, 2015, p. 4). This is a key theoretical grounding for the thesis as it enables the discussion to progress from the usual polarising debate “between markets in services as “good” and markets in services as “bad” [which] *misses much of what markets in services are doing*. Markets have dramatically changed the way services operate, but not in a uniform way” (Gingrich, 2011, p. 7, emphasis added). Fundamentally, the differential nature and function of markets “empower different actors and thus trade off different incentives” (Gingrich, 2011, p. 3).

Central to this thesis is a consideration of how the quasi-marketised relationships are structured since “depending on how information, contracts, and demand are shaped,

the incentives that producers face in the delivery of services may promote attention to the buyer, or to the user, or even give new producers the scope to follow their own interests” (Gingrich, 2011, p. 9).

Empirically, there is considerable variation in the design of markets for the provision of welfare-to-work services internationally (de Graaf and Sirovátka, 2012; van Berkel et al., 2012) but, as van Berkel et al. (2012, p. 273, emphasis added) note, much of the current research “pays little attention to the issue of *diversity* in the design and functioning of markets”. There is in this context an analytic gap and a need for a systematic consideration of the particular formulations or varieties of quasi-markets and their links to user experiences and programme performance (van Berkel et al., 2012). The thesis takes this analytic necessity as an invitation for the development of an extended and multi-dimensional graduated analytical framework articulating multiple potential alternative *types* of quasi-markets.

The empirical aspects of the thesis examine the practice of marketised governance from the perspective of just one central public service – the contracted employment support (aka ‘welfare-to-work’) provided to people who are long-term unemployed – in Britain post 2010. This conceptually broad and empirically focused study provides a rare opportunity to trace outcomes directly from the plans and promises of a particularly bold quasi-market experiment and to consider the ways in which key design elements cascade through to, and are detectable in, the patterning of employment and earning outcomes of programme participants on the ground.

1.2 British welfare-to-work in context

In the UK overall policy for employment services and the administration of out-of-work benefits is the responsibility of the Department for Work and Pensions (DWP). Jobcentre Plus is the UK’s public employment service (PES) and DWP’s key operational agency tasked with delivering employment services across a large network of frontline offices. In contrast to decentralised arrangements which are more prevalent across the rest of Europe it is the highly centrist DWP which controls design and delivery of employment policy and, allied, social security policy, in terms both of the Jobcentre Plus PES and major contracted-out programmes (Wiggan, 2015a).

In the UK people who are unemployed and of working age may be eligible to receive means-tested social security benefits, typically either Jobseekers’ Allowance (JSA) or Employment and Support Allowance (ESA). JSA is the main out of work benefit for people who are assessed as being ready and able to seek employment and enter work. As a condition of their continued benefits eligibility recipients are required to sign a ‘claimant commitment’ setting out the steps they will take to actively look for work and must attend a fortnightly job-search review at a Jobcentre Plus office or face sanctions. People who have a disability or long-term health condition may apply for ESA and will undergo a ‘Work Capability Assessment’ to assess their ability to look for work and/or carry out preparatory work-related activities. People who do not meet

the criteria of the Work Capability Assessment are directly diverted to JSA and are required to seek work immediately. Dependent on the prognosis of the Work Capability Assessment those who are entitled to claim ESA may be directed to a ‘Work Related Activity Group’ (who are entitled to just 88 minutes a year of Jobcentre Plus support, LWI, 2016a) or, for more severe health conditions, a ‘Support Group’ who may volunteer but will not be mandated to participate in any work preparation activities (indeed they will not actively be offered any employment support). The recent introduction and ongoing roll out of Universal Credit – the rationalised integration of key working age benefits and tax credits – is gradually replacing these ‘legacy’ benefit types, but without fundamental reform to the underlying groupings, obligations and supports for people who are out of work.

Since the late 1990s the UK’s provision of employment services has shifted away from the state as the main provider (through Jobcentre Plus) to independent organisations in the voluntary and private sectors. Initially this shift was small scale, gradual and highly variable as partnership arrangements were struck between local Jobcentres and alternative provider organisations (Sainsbury, 2017; Sunley et al., 2006). Up until 2011 by far the main provider of employment support services was still the state via Jobcentre Plus (Sainsbury, 2017).

In 2011 the new Conservative-led Coalition Government dramatically expanded the UK’s market for contracted employment support by independent providers through its rapid introduction of the Work Programme, the largest outsourced ‘payment-by-results’ programme ever launched in the UK (NAO, 2015). In international perspective the UK is generally understood as a committed marketiser in welfare-to-work (Wiggan, 2015a, p. 119) and the size, scope and experimentation involved in Work Programme’s marketising efforts position it at the bleeding edge of international reform experiences.

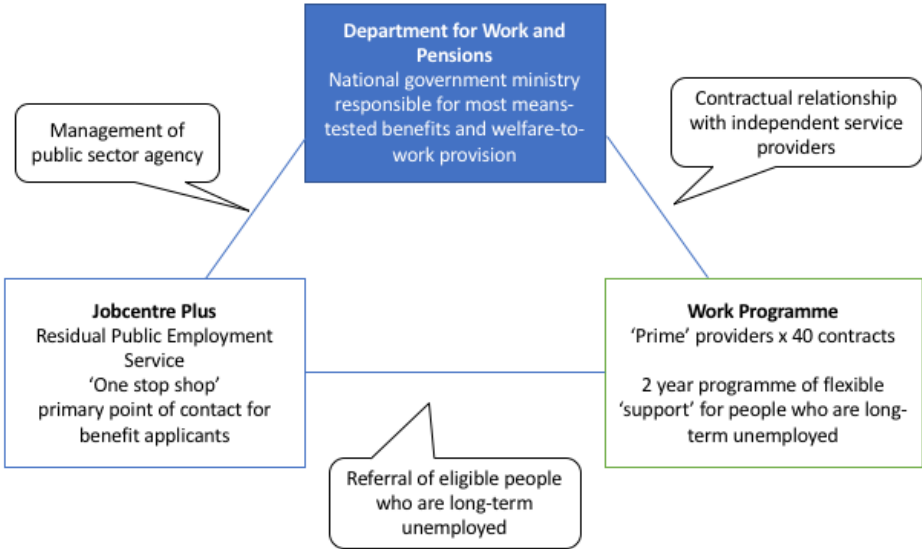


Figure 1.1 Key British welfare-to-work arrangements 2011-2018

1.3 The Work Programme

In 2011, the Coalition Government's cancellation of the still young Flexible New Deal contracts and the rapid introduction of the Work Programme at a stroke replaced virtually all welfare-to-work services administered by the DWP in England, Scotland and Wales. The programme offers support to people who are long-term unemployed or who are deemed at risk of becoming and aims to help them get, and keep, jobs. The Work Programme is a single scheme, larger than any previous British employment programme and serves an unprecedented range of unemployed people. It is a monolith in terms both of its size and its breadth: the Work Programme is expected to cost between £3 billion and £5 billion and could help up to 3.3 million people with diverse needs, circumstances and barriers (NAO, 2012).

As discussed further in Chapter 3, delivery of the Work Programme takes place through contracts between the DWP and large-scale, mainly private sector Prime providers which can both deliver services themselves and/or sub-contract to organisations within large and (sometimes) complex supply chains sitting underneath each Prime. The Work Programme is structured geographically into 18 large 'regional' Contract Package Areas with two or three Primes in each, to whom claimants are randomly allocated from Jobcentre Plus if they have not found work within an initial period of Jobcentre Plus provision, the duration of which depends largely on the type of out-of-work benefit received and the Work Programme 'claimant group' into which they are therefore placed. Figure 1.1 outlines these overarching relationships.

The Work Programme embraces fully market accountability principles. It is a pioneer and radical experimenter in the heavy extent to which payment to providers is predicated on the achievement of employment outcomes, with payments to providers tied entirely to job outcomes since 2014. The Work Programme implements a 'black box' delivery model so that providers have almost complete flexibility over their interventions. This flexibility is argued to be required since, unlike previous contracted employment programmes, Work Programme is 'universal' and therefore has to cater for the needs of a wide range of unemployed programme participants within a single scheme. In its size, breadth and extent of quasi-marketisation therefore, the Work Programme is a vanguard of radical experimentation in welfare-to-work quasi-markets not only within the British context but also within international comparative perspective (Finn, 2011; Mulheirn, 2011; this author in Carter and Whitworth, 2015). It is, for these reasons, and its importance within the UK employment policy landscape that it is an important case study to assess empirically in this thesis. More broadly, the Work Programme is a valuable case due to the wider lessons that can be drawn around the performance of alternative governance arrangements for employment polices beyond the UK and, indeed, social policies beyond the field of employment.

1.4 Perennial risks and hubristic design promises: the research focus and chapter summaries

Within public services dominated by marketised accountabilities – such as the Work Programme – it is expected that opportunism or ‘gaming’ will be a particular design challenge. Hart et al. (1997) contend that where there are incomplete contracts (which is almost inevitably the case in complex public services) and producers have great control then private providers demonstrate an ability to pursue cost-cutting innovations (i.e. improve crude efficiency) but that they are also more likely to do so at the expense of quality and other unspecified but potentially important aspects of provision. Specifically, there are risks within the Work Programme – as in all such quasi-marketised employment support programmes – that providers will seek to ‘cream off’ and take easy payments for participants close to the labour market whilst ‘parking’ those participants who face multiple barriers to work and whose payments are therefore substantially more difficult and/or more costly to achieve. Mitigating such risks is a perennial design challenge within quasi-marketised welfare-to-work schemes relying strongly on marketised accountability levers such as payment-by-results as demonstrated by the plentiful international evidence of creaming and parking in practice across such programmes (Struyven and Steurs, 2005; Bredgaard and Larsen, 2007; Finn, 2010a, 2010b; Considine et al., 2011; de Graaf and Sirovátka, 2012).

In addition, the international literature also makes clear that the specificities of programme design and payment structures can play a key role in either facilitating or buttressing against such risks and provider behaviours (Considine, 2000; Considine et al., 2011; Finn, 2009, 2011, 2012; Struyven and Steurs, 2005; van Berkel and van der Aa, 2005; Koning and Heinrich, 2013). The implication is that through careful contractual specification and regulation it may be possible to overcome, or at least mitigate, these behaviours and ensure that providers act in support of the full suite of programmatic objectives. The challenge then is for policymakers – through design savvy – to configure and implement a suite of steering tools that retain the perceived innovative and efficient impetus of quasi-markets whilst ensuring that a complex set of public sector objectives are met.

As discussed further in Chapter 5, for DWP beneath the headline objective of moving people who are long-term unemployed into employment the Work Programme aspires to achieve four main, and potentially conflicting, objectives: efficiency (purchasing employment outcomes at a lower unit cost than preceding schemes); efficacy (increasing aggregate time in employment through sustained jobs); economy (providing savings to HM Treasury through reduced benefit spending and increased tax revenue); and equality (reducing the gap in employment rates between disadvantaged groups and other participants).

Within DWP’s design response to the criticisms of previous quasi-marketised schemes there are three bold and untested marketised accountability elements within Work Programme which seek to mitigate these known quasi-market risks towards the delivery of this set of programmatic objectives: differential payments; the

geographical scale and structure of Prime contracts; and sustainment payments. The overarching question which animates this thesis is whether this crafted Work Programme design structure is sufficient and appropriate to steer its quasi-marketised providers to the achievement of DWP's full gamut of policy objectives. Enabled by the combination of unprecedented access to DWP's key administrative data alongside sophisticated and conceptually tailored multivariate quantitative analyses, the thesis' contributions both relate to and flow from its inter-related examination of conceptual debates around the governance of quasi-markets, systematic empirical understanding of the programmatic performance implications of Work Programme's internationally innovative quasi-marketised governance attempts, and new policy understandings around the nature, limits and interactions of alternative governance mechanisms within alternative varieties of quasi-marketised public services.

The remainder of the thesis is structured across eight chapters.

Chapter 2 introduces the key concepts of governance and accountabilities before setting out the thesis' original analytic framework that develops and critically extends Gingrich's (2011) typology for considering variation in the form of quasi-marketised public service. Specifically, the chapter develops an original multidimensional analytical framework which maps and illustrates graduated dimensions of *market allocation* and *market production* that serve as a scaffold against which to trace crucial aspects of variation in the formulation of quasi-markets in practice. A key justification for this conceptual approach is the expectation that the way the market is structured will have important implications for attuning the attentiveness of service providers to the priorities and preferences of either the state, service users, or providers themselves. This in turn is expected to have important implications for the quantity, quality and distribution of services and consequently on the lived experience and 'outcomes' of programme participants.

Chapter 3 provides a firmer grounding in the specific policy context of UK welfare-to-work. The chapter outlines the wider chronology of policy evolution since the early 1990s. The analytic framework developed in the preceding chapter is then used as a tool to systematically unpack the form of quasi-market instigated within the Work Programme. This reveals the Work Programme to embody a *private power market* in which the preferences of service providers are dominant.

Chapter 4 sets out the data and methods which underpin the empirical contribution. The chapter sets out the ways in which the thesis benefits from unique academic access to DWP's comprehensive administrative datasets combined with methodologically sophisticated and conceptually tailored multivariate quantitative analyses to explore in unprecedented depth, detail and specificity the key linkages between governance mechanisms and programme performance.

Chapter 5 assesses critically the differential payment structure and considers the success of this design feature in *narrowing the gap* in job outcomes between

disadvantaged groups and everyone else. The research question flowing from this then is: *Has the Work Programme's differential payment structure calibrated provider incentives within the private power market such that all programme participants have the same likelihood of entering and sustaining paid work, regardless of their characteristics and circumstances?*

Chapter 6 investigates the geography of incentives and performance in the Work Programme, assessing potential tensions between the spatially extensive contracting areas and local labour market contexts. The research question here asks *Is the geographical container of regional Contract Package Areas an appropriate spatial scale at which to uphold minimum performance levels and incentivise competition between Prime providers?*

Chapter 7 considers the use of sustainment payments and the structure of financial incentives which extend well beyond an initial period of employment. The research question asks *What are the employment and earning trajectories of Work Programme participants? Has the use of sustainment payments broken the low-pay no-pay cycle?*

Chapter 8 offers a more synoptic reflection on the Work Programme's private power market. The research questions whether the neglect of participants with health conditions and disabilities is inevitable, asking *Does the variety of quasi-market matter for people with disabilities and health conditions?*

In summary, therefore, the focus of the thesis is both narrower and broader than the Work Programme itself. It is narrower in the sense that the research does not seek to offer a full evaluation of the scheme (although the empirical work represents in various ways contributions that are original, novel and significant compared to existing analysis of the programme) but rather looks to assess – using a range of robust multivariate quantitative analyses – the very specific design promises made within the private power market.

And yet the empirical findings also take on a much broader application. In international perspective the UK is generally understood as a committed marketiser in welfare-to-work (Wiggan, 2015a, p. 119) and the scope and experimentation involved in Work Programme's marketising efforts position it at the bleeding edge of reform experiences. Relatedly, welfare-to-work can be seen as something of an unchartered pioneer in its extensive use of outsourcing and payment-by-results which is now increasingly being adopted within other policy spheres (Albertson et al., 2018; NAO, 2015), and Work Programme is at the vanguard of this internationally.

This thesis's original conceptual framework, in conjunction with the unprecedented depth and targeted sophistication of empirical scrutiny advanced, offers policymakers a powerful lens through which to consider the linked conceptual-empirical implications of variation in quasi-market types. The work offers critical reflections on the design fixes available to policymakers internationally who seek to better balance

the competing priorities of government, service users, and providers both within and beyond the field of welfare-to-work services.

Chapter 2 begins by setting out the thesis's conceptual foundations around governance and accountabilities and, building on these ideas, by advancing its original analytical framework to enable better understanding of varieties of quasi-markets and their implications for providers, commissioners and service users.

2 Making (quasi) markets in the welfare state

2.1 Chapter summary

Welfare-to-work services have been a key area of experimentation internationally in quasi-marketised public service delivery. Critical discussion of these varied market forms, however, and their differing implications for policy delivery and performance, are less well advanced in the academic literature and policy debates. This chapter builds the conceptual backbone to support the subsequent policy discussion and empirical chapters by developing an original analytical framework to guide thinking on the governance of welfare-to-work services which is sensitive to the *varieties* of marketisation on offer and the key dimensions on which the organisation of service provision might vary. Analysing the diversity in governance formulations – that is, the mode of coordination underpinning the provision of services – provides a key route to consider divergence in the priorities of service delivery and to unpack the potentially heterogeneous implications for service user experiences and programmatic employment outcomes.

The chapter begins with an introduction to the key concepts of governance and accountabilities before moving to a more subtle discussion than is currently seen in the literature around *varieties* of quasi-markets and the artificiality of those market forms which are constructed to underpin public service delivery. The chapter is informed by, and crucially develops and extends, Gingrich's (2011) framework for considering variation in the form of quasi-marketised public service. Specifically, the chapter develops an original analytical framework which maps and illustrates graduated dimensions of *market allocation* and *market production* which in turn serve as a scaffold against which to trace crucial aspects of variation in the formulation of quasi-markets in practice.

A key justification for this conceptual approach is the expectation that the way the market is structured will have important implications for attuning the attentiveness of service providers to the priorities and preferences of either the state, service users, or providers themselves. This in turn is expected to have important implications for the quantity, quality and distribution of services and consequently on the lived experience and (employment, but also broader) 'outcomes' of programme participants.

2.2 Shifting modes of governance in public service delivery

A fundamental shift has taken hold of welfare state governance in developed economies since the 1980s with the increasingly widespread use of privatisation, private firms and market-like methods for the delivery of public services. This conceptual chapter considers the broader shifts at play in the arrangement of public services (that is, the "how" in policy implementation processes) and introduces the changing levers through which governments configure the provision of social programmes.

Until recently, social policy scholars typically paid little attention to issues of governance in the study of social policy reforms, programme evaluation and welfare state transformations (van Berkel and Borghi, 2007). A growing number of studies however increasingly recognise the centrality of governance to policy design, experience and performance. Beyond this, researchers are crossing traditional disciplinary boundaries and integrating a discussion of ‘operational policy’ or governance – that is, the organisation and management of policy making and programme delivery processes – with that of the ‘formal’ programmatic intent and substance of those social policies (for example, Considine, 2001; Henman and Fenger, 2006; Rees et al., 2014; Sol and Westerveld, 2005; Struyven and Steurs, 2005; van Berkel and van der Aa, 2005; Wiggan, 2015a, 2015b). Put differently, these studies explore the connections between the governance of public services and the substantive aspects of their provision, content and emphasis of programme implementation.

Van Berkel and Borghi suggest that we should welcome research that sits across social policy and issues of governance “because governance reforms are not without consequences for the content of social policy programmes” and because the introduction of new modes of governance may have intended as well as unintended consequences for policy substance (2007, p. 280). What this emergent body of work suggests is that:

“The question of what mode of coordination of the provision of social services is used, is not simply a “technical” issue about what actors to involve in service provision, or how to structure the relationships between them. It is also based on, or has consequences for, a range of other issues, such as the values that will guide the service provision process, the role and position of service users in service provision, the accountability of service providers, the discretion of frontline workers involved in service provision, opinions on how the behaviour of service providers and service users should be steered...”

van Berkel and Borghi, 2007, p. 281

In short, the nature and content of social programmes are not simply a product of ‘official policy’. Programme governance sits as a key mediator between formal policy – as embodied in strategy documents and edicts – and the implementation, lived experience and ‘outcomes’ of programme participants on the ground.

In this chapter a deeper engagement with the theoretical considerations of governance regimes enables the anticipation of particular weaknesses and failings which may be experienced under alternative governance and accountability formulations. This enables the research to be more alert to tensions and potential unintended consequences when reconciling programmatic intent with the logics of particular modes of service coordination.

The chapter proceeds across five key sections. Firstly, it introduces the key grounding concepts of governance and accountabilities to trace the alternative modes of coordination underpinning the delivery of public services and sets out important

distinctions between these terms. The second section charts the rise of quasi-marketised forms of public service governance. Thirdly, the chapter problematises ‘marketisation’, suggesting that it is not a singular homogeneous phenomenon (as often discussed in the literature) and draws on the difference thesis to fracture and decompose this as a trend which encompasses multiple varieties of quasi-market. The fourth section provides an extension to previous work on the difference thesis by explicitly unpacking and reconceptualising the underlying dimensions against which quasi-marketised formulations might vary. This provides a new analytical framework through which to compare alternative quasi-marketised arrangements and, importantly, their expected differing impacts for service users and programme outcomes. Finally, the chapter progresses the conceptual understanding further by drawing new links between the literatures on accountabilities and quasi-market varieties in order to enable researchers and policy makers to better identify where alternative accountability structures might serve as a corrective to perceived weaknesses in current quasi-market programmes.

2.3 Governance as a grounding concept

The term ‘governance’ has become ubiquitous as a topic of academic discussions (Henman and Fenger, 2006; Jessop, 1998; van Berkel et al., 2012b) and reviews of the literature typically conclude that the term is used in a range of ways and with a variety of meanings (Rhodes, 1996; Stoker, 1998; Kooiman, 2003; Treib et al., 2007). As a result, some have reviewed core components of the literature and suggest cynically that: “if it means everything, perhaps it means nothing” (Prätorius, 2003, p. 237).

Despite this eclecticism and fluidity, governance continues to hold value as a grounding concept in this thesis through its capacity to provide a rich conceptual framework for understanding changing processes and forms of government function. It helpfully marks a distinction from conventional understandings of ‘commanding’ government and instead places a focus on ‘steering’ and ‘influencing’ mechanisms that do not rest on recourse to the crude authority and sanctions of government (Stoker, 1998). In this spirit, Jessop (1998, p. 30) traces the rise of the governance paradigm noting that a key factor in the escalating use of the term is “the need to distinguish between ‘governance’ and ‘government’”.

Governance then signifies a set of elusive but significant shifts in the way in which government seeks to govern (Pierre and Peters, 2000). Importantly the term *expands* the set of institutions and actors that are drawn from, but also beyond, government; *blurs* the boundaries and responsibilities for tackling social and economic issues; and *reformulates the roles and relationships* (particularly the power dynamic) between the institutions involved in collective action (Stoker, 1998). It is not that the state is without power or control, rather it draws on different, new repertoires through which to steer and achieve objectives (Jessop, 1998).

Treib et al., (2007) support the disambiguation of the term by outlining three conceptions or realms of governance. The first is governance as *politics*, where state

actors share power with private actors in the process of policy formulation. The second conception is *polity*, the governance of institutional rules, norms and actions and which for Treib et al., (2007) spans between the logics of the market and that of hierarchy. The third frame describes governance as the specific steering instruments of *policy* implementation, for example, command and control, incentive and supply, information, deliberation and persuasion. Most particularly for the present study, the focus straddles the second two conceptions of *polity* and *policy*, that is, subsequent to decision making on appropriate political objectives (which have indeed shifted over time, and which are elaborated in the subsequent chapter on activation policy).

The empirical focus of the thesis is informed by Considine's (2001; Considine et al., 2011) work on governance transformations in the sphere of welfare-to-work which similarly identifies two crucial aspects to the shifting governance dynamic. The first transformation involves organisational changes within the public service itself: "bureaucrats find themselves having to define their roles according to new demands" (Considine, 2001, p. 10). This parallels the *polity shift* and captures revised institutional logics and rationalities. The second process for Considine (2001) involves the development of new relationships – and, indeed, new *tools* for navigating relationships – with other public and private agencies. Considine refers to an 'enterprise toolkit' of now familiar devices including contracts, targets, and pricing which accumulate and interact to support a process of changed governance.

Although Considine does not adopt a separate nomenclature for these two discrete but related concepts, this thesis makes a key analytical separation between the underlying institutional logic – henceforth termed 'governance' or 'governance framework' – and the specific tools and devices for negotiating the changed relationships inferred by those particular logics and rationalities – henceforth referred to as 'accountabilities' or 'accountability levers'. These distinct ideas, and the interactions between them, are elaborated further throughout this chapter's later discussion. For Considine, governance takes four types: procedural, managerial, market, and network (2001). Of the first three, each "has different legislative enactments of public service regulations, different expectations about the roles of public managers, together with alternative strategies for resource allocation, organisational dependency and interdependency..." (Considine, 2001, p. 23). The fourth variety of 'network' governance was considered to be more of a hypothetical formulation at the time of Considine's 2001 study. Arguably, and as discussed further in the closing chapter of the thesis, since this variety of governance is understood as a "synonym for multi-agency co-ordination, reciprocation, lateral communication and discretionary bargaining by local actors" (Considine, 2001, p. 24) it remains an emergent and not fully prevalent form in the still highly centralised realm of UK welfare-to-work services.

Figure 2.1 reproduces and extends Considine's (2001) summary of the organisational character associated with each form of governance with the addition of the 'democratic' form informed by Jantz et al. (2015; see also Whitworth and Carter, 2018; Carter, Forthcoming). For each governance ideal Figure 2.1 conveys the key essence of each type: "its form of rationality, form of control, primary virtue, [as specified by

proponents] and the nature of its service delivery focus” (Considine, 2001, p. 24).

Governance mode	Source of rationality	Form of control i.e. core accountability levers	Primary virtue	Service delivery focus
Procedural	Law	Rules and regulations	Reliability	Universal treatment
Corporate	Management	Plans and targets	Goal-driven	Targets
Market	Competition	Contracts, competition and pricing	Cost-driven	Prices
Democratic	Citizen voice	User choice and voice	User-responsive	User satisfaction
Network	Relationships	Co-production	Flexibility	Brokerage

Figure 2.1 Governance and accountability types, adapted from Considine, 2001 and Carter, Forthcoming

Closely entwined with this distillation of governance models there is a small but growing number of academic articles (for example, Jantz et al., 2015; Jantz and Jann, 2013; Koppell, 2005; Romzek et al., 2012; including from this author, Whitworth and Carter, 2018) using the language of ‘accountabilities’ to capture the shifting instrumental toolkit prioritised under alternate governance conceptions. This thesis is informed in particular by Jantz et al., (2015, p. 5) who understand accountability as “a) a system of knowing and evaluating someone’s behavior [sic] according to some standards and b) as a system of rewards or sanctions that are depending on these evaluations”.

At first sight ‘accountabilities’ appears as a potentially perplexing duplication of aspects of ‘governance’. However, there are three core reasons for accountabilities being a valuable addition to this chapter’s conceptual lexicon. Firstly, Jantz et al., (2015) introduce a valuable separation between administrative (i.e. bureaucratic procedural) forms and democratic accountabilities. As will be discussed further, this additional ‘democratic’ service-user-centred form is a valuable lens through which to consider a crucial additional set of preferences and values to which service provision may be oriented.

The second key contribution is the emphasis this conception places on the range of accountability levers not as a singular institutional arrangement but as an agglomeration and interaction of types which sit together within a broader multi-dimensional ‘accountability framework’ (Whitworth and Carter, 2018) or “web of accountability relations” (Jantz et al., 2015, p. 20). Whilst one particular accountability dimension may be expected to dominate within given governance contexts, at particular points in time, it is understood that this will not necessarily be the *only* accountability lever in play. For example, whilst Great Britain as a “committed marketizer” of employment services (van Berkel et al., 2012a) “could be expected to have gone furthest in embedding market instruments as the dominant form of accountability ... this does not translate into the replacement of *all* alternative accountability instruments” (Jantz et al., 2015, p. 4, emphasis added). It is important then to tease apart any overarching and dominant logic or ‘governance type’ whilst

continuing to identify the on-going presence of alternative accountability levers and forms which fall outside of this hegemonic governance formulation. Importantly then, and as discussed further in relation to Figure 2.5 below, the suite of accountability forms in practice tend to co-exist and intersect with one another.

Related, this connects to the third and final benefit of ‘accountabilities’ which is to open up explicitly a more detailed recognition and discussion of governance hybridity, still a relatively neglected area of scholarly attention. Greater awareness of these hybrids facilitates a fuller investigation of how a range of accountability types mesh (or fail to mesh) together, with the potential for more supportive configurations where the strengths of some accountability forms have the potential to rectify or partially correct for weaknesses in others if their interactions are incorporated more explicitly and more fully into policy design and delivery.

Previous work published by this author (Whitworth and Carter, 2018) then outlines five accountability mechanisms:

- **Procedural** accountability draws on Weberian hierarchy, rule-based principles and norms of reliability and procedural fairness. Key devices include set processes, rules and requirements;
- **Corporate** accountability utilises contractualised performance targets to monitor and compare provider performance so as to reward or punish providers accordingly (e.g. contract renewals/terminations, additional/reduced referrals or payments);
- **Market** accountability steers using levers of price and competition between providers. Payment-by-results, accelerator pricing, or other financial incentives/sanctions for good/bad performance typify market accountability levers;
- **Democratic** accountability responds transparently to the views of citizens and users. Whilst Jantz et al. (2015) focus ‘top-down’ on the accountability of politicians through the electoral cycle, Whitworth and Carter (2018) also include possible ‘bottom-up’ democratic accountability through user experience via levers of ‘voice’ (e.g. feedback, complaints, service rating) and ‘choice’ (e.g. provider and/or intervention selection and exit) to drive service quality (Hirschman, 1970);
- **Network** accountability coordinates provision across multiple, inter-dependent service providers using relationships based on trust. Softer informal levers around the collective need for organisations to protect reputations and foster on-going relationships of co-operation and co-dependence offer informal and collectively-enforced network accountability mechanisms (Olson, 1965).

Importantly, the core values and approaches associated with each of these governance types and operationalised through accountability mechanisms are more than “transitory rhetorical flourishes” but rather emerge as very real engagements enacted

at the core of public services (Considine, 2001, p. 18). The behaviour, relational dynamics and experiences of staff members, agencies and service users are profoundly shaped by the degree to which differing combinations of accountability levers are leant on within any given programme or service stream. Any public programme on paper may incorporate elements from across each of these five accountability types (Whitworth and Carter, 2018) and in practice the accountability matrix – the respective mix and weighting of those accountability levers – may shore up or displace the ability for a programme to successfully deliver policymaker (or service user) ambitions. Furthermore, as is ever the case there may also be a mismatch between a programme’s accountability devices on paper and its operational reality once implemented. Indeed, it is arguably the perceived weaknesses within one governance logic (underpinned by a dominant set of accountabilities) which make way for the ascendance of an alternate governance form, as succinctly described with particular attention to market governance below.

2.4 The rise and rise of market governance and accountabilities

Various scholars have helpfully traced the evolution of these governance forms (Newman, 2001; Tenbenschel, 2005; Knuth, 2014; Considine, 2001). Considine (2001) notes that procedural, corporate and marketised accountabilities broadly correspond to the development of public bureaucracy from its post war origins to more recent waves of reform. The procedural dimension captures ‘old fashioned’ notions of public bureaucracy where services are delivered according to rule-based principles and norms of reliability and procedural fairness with public servants following set processes, rules and requirements. Corporate governance can be seen to align with the rise of New Public Management (Hood, 1991; Pollitt, 2003) as a response to the rigidities and lack of responsiveness of traditional bureaucratic levers. Under this dimension contractualised plans and performance targets are used to monitor and compare service provider performance.

The rise of market – or more precisely quasi-market – forms of governance and accountability since the 1980s has been a particularly potent trend. The growing use of market-inspired modes of coordination is of specific interest for the current study since British welfare-to-work policy has been at the leading edge of these reforms internationally (Lødemel and Trickey, 2001; van Berkel et al., 2012a). There are a range of arguments utilised by proponents for the introduction of market-like arrangements into welfare systems traced typically to the economic and political challenges of the post-1970s era which both challenged debates over spending (calls for small government), promoted a desire for greater choice, and emphasised constraints on traditional policy instruments (Gingrich, 2011).

Gingrich (2011) distils two dominant ‘pro-market’ arguments belonging respectively to public choice theorists and advocates of ‘New Public Management’. Public choice theory suggests that elected politicians are rent-seekers and that “...public provision

creates a bureaucracy out to service its own needs and not those of users” (Tullock et al., 2002 in Gingrich, 2011, p. 27). The recommendation then is that markets ought to limit the size and scope of the state and that by shrinking public provision services will be replaced with a superior *marketised* form of allocation and production.

A second line of ‘pro-market’ argument is generally understood as nestling within New Public Management (NPM) thinking, though arguably NPM is a fairly unwieldy and imprecise project with no fixed agreement as to what precisely it encompasses or entails (Hood, 1991). Proponents here suggest that government ought to be managed like a business either by liberating bureaucrats to act like business managers, by breaking bureaucracies into agencies responsible for pursuing targets (as per the corporate accountability lever in Figure 2.1), or by ‘contracting out’ and using independent providers to deliver services (steered by marketised accountabilities). Here the argument is not necessarily that state spending on public services should be reduced (the government is not presumed to be a predator or leach) but rather it is argued that government bureaucracies operate inefficiently and are unresponsive to citizen’s demands and desires in their delivery of public services (Hood, 1991; Kettl, 1993).

Osborne and Gaebler’s (1992, p. 14) seminal work argued against traditional, procedural Weberian rule-based democracy which had “a distinct ethos: slow, inefficient, impersonal”. The promise from within this frame is that markets are a technocratic means for improving public services, not a replacement for them (Lundsgaard, 2002). It is this argument which has reverberated most prominently through recent waves of public service ‘reform’ in the UK.

Marketised coordination is instilled by splitting apart the ‘purchasing’ function from that of ‘providing’ public services in the belief that government should ‘steer not row’ (Osborne and Gaebler, 1992). Under this marketised approach the state retains important responsibilities around financing and regulating services even if the public sector no longer provides those services directly. In explaining early moves to marketization in the UK in the late 1980s Le Grand and Bartlett (1993, p. 3) note that “...the state was to become primarily only a purchaser of welfare services, with state provision being systematically replaced by a system of independent providers competing with one another in internal or ‘quasi’-markets”. The behaviour of service providers is steered through economic stimuli such as payment incentives and any regulation principally takes place through contracts with public commissioners (typically central or local government).

The dominant arguments in support of the marketised approach suggest that this mode will improve the efficiency, responsiveness and quality of services (Osborne and Gaebler, 1992; Le Grand and Bartlett, 1993; Savas, 2000; Lundsgaard, 2002). The improving power of markets is thought to be driven by two broad processes. Firstly, open competition between multiple independent service providers is expected to result in improved service *efficiency* such that the total costs of service delivery are minimised. Secondly, choice between multiple providers, exercised either by service

users or by government purchasers, is expected to instil greater responsiveness to need, increase entrepreneurialism and is expected to contribute to more tailored, better *quality* service solutions. A further benefit is thought to flow from a process of “de-bureaucratisation” since marketised models are seen as more flexible and capable of responding to changes in the business cycle and the needs of target groups (Bredgaard and Larsen, 2007, p. 289).

2.4.1 Public service markets as ‘quasi’-markets

Crucially, for the promises of improved efficiency and quality to hold, important assumptions must be met, namely that competition and choice function as in theoretical markets of neoclassical economics. Specifically, “For the allocation of a service by a conventional market to be efficient, responsive and to offer genuine choice, the market concerned has to be *competitive*” (Bartlett and Le Grand, 1993, p. 19). The work of Julian Le Grand (Le Grand, 1991; Bartlett and Le Grand, 1993; Le Grand and Bartlett, 1993) has for these reasons popularised the use of the term ‘*quasi*’-market since in important ways the introduction of market-like mechanisms in the field of public service diverges from a conventional understanding of market functioning.

Le Grand and Bartlett (1993) emphasise that in public service markets deviations from a conventional market structure occur on both the supply and demand side. On the demand side, consumer – that is service user – preferences are not expressed directly in money terms through consumer purchasing power as in the key ‘invisible hand’ of pure markets. In contrast, within public service ‘markets’ service users are typically not able to express choice over their provider at all and where service users do have choice it tends to be highly constrained in its remit and operation. Instead, ‘choice’ is dominated by commissioners via purchasing decisions that are taken by a government agency or are delegated to a third party acting on the public sector’s behalf.

Conventional market relationships, or even government procurement relationships in basic goods (stationary etc), typically have one ‘principal’ (the contractor) and one ‘agent’ (the producer). Contrastingly, in more complex public service market arrangements there are three key sets of actors: payers, users, and producers of services. The presence of separate payers and users fragments demand, “raising the question of who the principal is to whom the agent is supposed to respond – the payer of the service or the user” (Lowery, 1998; in Gingrich, 2011, p. 9).

In terms of supply, in contrast to conventional markets, not all provider organisations are necessarily operating in direct pursuit of profit maximisation. Independent, non-profit providers also seek to act in the space of public service provision and their objectives and priorities in this space – as well as their ownership structure – is not always clear (Bartlett and Le Grand, 1993).¹

¹ Work conducted by the Third Sector Research Centre (Rees et al., 2013) suggests that ownership structures and charitable purpose may – in the context of cost pressurised welfare-to-work provision – do little to differentiate provider practice.

Beyond these supply and demand side peculiarities in public service quasi-markets, Gingrich (2011) goes on to identify information asymmetries and incomplete contracts as further market ‘problems’ which pose challenges for structuring effective marketised arrangements within public services. By way of information asymmetries, Gingrich suggests that “where the supplier of the service, such as a doctor or social worker, knows more about the cost and quality of the service than the user [and perhaps also more than the payer, this], make[s] it difficult for consumers to shop around for producers or insurers” (Gingrich, 2011, p. 9). The implications of asymmetries in information – where greater insight is held by private providers than by the government payer – features in Chapter 5. The performance challenges of information asymmetries in the field of employment support services specifically has been identified in previous work commissioned by the DWP which found that in schemes seeking to implement a degree of user choice, the absence of relevant information made selecting a provider particularly challenging (Conolly et al., 2010). Relatedly, specifying ‘quality’ in outsourced service contracts is notoriously challenging. Whilst it may be possible for policy makers to specify some outcomes of interest it is highly questionable as to whether such a specification could fully capture all aspects of service quality, meaning that public services are plagued by incomplete contracting (Hart et al., 1997). This serves to compound principal-agent problems and underscores frequent challenges in quasi-marketised approaches to public service design and delivery, including in welfare-to-work services.

As a result, the pursuit of market mimicking mechanisms in public service governance must always be considered as a trend to ‘quasi’-markets. These ‘quasi’-markets are not instinctive, automatic ‘natural’ market systems, but rather they are artificial constructs: outcomes are specified; prices must be set; contracts written; competition between providers arranged; ‘winning’ providers selected; performance overseen and steered; the ‘market’ regulated, refined, maintained. In short, the rules of the ‘game’ are designed by policy-makers (re-badged as ‘market stewards’) and “there is not a single free, competitive, benchmark against which to assess the introduction of market incentives in services” (Gingrich, 2011, p. 9; Gash et al., 2013).

2.5 Fronting and fracturing understandings of market governance: the difference thesis

This discussion of public service markets as ‘quasi-markets’ helpfully identifies the ways in which we must consider their introduction as an artificial endeavour and, as an implicit consequence, how there are important distinctions in the design and functioning of these systems when compared to theoretical pure markets of neo-classical economics. What is perhaps less explicit in much of the literature, and indeed in the highly polarised academic and political debates, is the degree to which the common *language* of ‘market’ reform has obscured comparative conceptual (and linked, empirical) analyses of what are, in reality, a varied set of quasi-market forms and reform processes (Jacobs, 1998; Greener, 2008).

Greener (2008) problematises the term ‘market’ perceptively noting that the concept is used to describe complex and varied forms. Powell (2015, p. 110), summarising this

discussion, suggests that the “phrases *quasi-market* or *internal market* have been used promiscuously to describe any system in which providers compete for resources. A convergence of terminology has created the illusion of a convergence of policy”.

This lumping together, or convergence, around a singular marketising endeavour has been powerfully deconstructed by advocates of ‘the difference thesis’ (Powell, 2015; Gingrich, 2011; Zehavi, 2012; Meagher and Goodwin, 2015). In short, this cluster of academic work suggests that ‘marketisation’ is not a singular phenomenon but rather that it captures a diverse range of “practices, rationales, trajectories, actors and impacts” (Meagher and Goodwin, 2015, p. 4). It is therefore possible – and indeed more appropriate – to consider the analysis of marketisation as a comparative endeavour and consequently to draw comparisons both *within* and *across* country contexts to look at multiple ‘varieties’ of quasi-marketisation.

The work of Jane Gingrich (2011) provides a powerful schematic through which to sharpen understandings of the – now dominant – market governance arrangements and fracture what has elsewhere been considered as an undifferentiated lump of governance reforms. Powell (2015) identifies Gingrich’s work as the most recent, detailed, and developed account of the ‘difference thesis’ and Wiggan (2015a, 2015b) has successfully used the foundations of Gingrich as a device through which to consider market developments in the context of UK welfare-to-work. The headline argument of Gingrich’s work is that:

“markets empower different actors and thus trade off different incentives, that markets in public services vary systematically and that policymakers can manipulate this variation strategically”.

Gingrich, 2011 as précised by Powell, 2015, p. 110

This is a key theoretical grounding for the chapter, and indeed the thesis as a whole, as it enables the discussion to progress from the usual polarising debate “between markets in services as “good” and markets in services as “bad” [which] *misses much of what markets in services are doing*. Markets have dramatically changed the way services operate, but not in a uniform way” (Gingrich, 2011, p. 7, emphasis added). Fundamentally, the differential nature and function of markets “empower different actors and thus trade off different incentives” (Gingrich, 2011, p. 3).

This chapter thus recognises and conceives of variation in market structures not as an artefact of chance or state-specific teleology but as the result of particular active reform choices made by politicians and officials. Central to the thesis is a consideration of how the marketised relationships are structured since “depending on how information, contracts, and demand are shaped, the incentives that producers face in the delivery of services may promote attention to the buyer, or to the user, or even give new producers the scope to follow their own interests” (Gingrich, 2011, p. 9).

What then are the options on the table in terms of the form that quasi-markets can take and the dimensions on which they may be manipulated?

2.6 A differential view of markets

As a route to unpack the specific market configurations advanced in particular areas of outsourced public service provision, a key contribution of this chapter is the development of an original analytical framework that significantly extends Gingrich’s (2011) contribution and starting point by more comprehensively mapping and illustrating the dimensions that underpin the varieties of public service markets on offer. In essence, this creates a richer analytical tool through which to understand the particular flavour of quasi-marketisation advanced in any given context and anticipate impacts for service users, providers and commissioners.

Before introducing the chapter’s original conceptual contributions this section firstly outlines the typology of quasi-markets advanced by Gingrich, at the heart of which is a consideration of variation on two dimensions – how access to services is *allocated* (capturing financing and regulation of service provision) and how they are *produced* (referring to the structure of competition and choice in the market). On the allocation dimension, Gingrich separates those services that are collectively financed through taxation from those where individuals bear costs and risks. The production dimension “refers to the structure of competition and choice in the market, as how this is organised affects whether it is the state, the provider or the service user who is best placed to advance their preferences and thus shape service production” (Gingrich as summarised by Wiggan, 2015a, p. 118).

Based on these dimensions Gingrich identifies three broad categories of state-, user- or provider-driven markets that essentially identify “who is the ‘piper calling the tune’” (Powell, 2015, p. 110). State-driven markets increase the power of government to set incentives for cost-efficiency; consumer-led markets empower the users of services and create incentives for producers to respond to users’ preferences of high-quality provision; producer-driven markets allow independent service providers to pursue profit maximisation (Gingrich, 2011, pp. 13–17). Service providers are strongest in situations where service users have a constrained ability to exit the market or where the contract defining the relationship between the government purchaser and provider has either weakly defined outputs/outcomes and/or limited monitoring of service quality.² “In either circumstance the provider enjoys greater scope to direct the market and prioritise their preference for profit maximisation over cost control or service responsiveness and equity” (Wiggan, 2015a, p. 118).

		Production dimension: who has effective control?		
		State: "Efficiency aims"	Users: "Quality aims"	Producers: "Profits and rents"
Allocation Dimension: Responsibility for Access	Collective	Managed Market <i>Recent English contracting in education</i>	Consumer-Controlled Market <i>Swedish health care market in the early 1990s</i>	Pork Barrel Market <i>English elderly care market in the 1980s</i>
	Individual	Austerity Market <i>Dutch health care markets</i>	Two Tiered Market <i>English education market</i>	Private Power Market <i>English elderly care market since the mid-1990s</i>

Figure 2.2 Gingrich’s six quasi-market types. Source: Gingrich, 2011, p. 12

As outlined in Figure 2.2, for Gingrich (2011) the destination for the allocation and production dimensions is a hard classification of six types of markets. The development of these ‘ideal types’ marks definitive progress in our understanding of varieties of quasi-markets in public services. Nonetheless, there is the potential to enhance Gingrich’s work, specifically in addressing the ‘boxing in’ of distinct types of quasi-markets in her classification, which is limiting in two key ways both analytically and for policy.

Firstly, the classification is not particularly strong on identification strategies. In a line-up of different quasi-marketised configurations *what are the key traits that flag a marketised programme or policy domain as exemplifying one particular form or another?* For example, how strong does user choice and voice need to be before a

² Gingrich suggests that pressures toward profit maximisation will be amplified where providers are privately owned and responsive to shareholders.

market becomes classified as ‘consumer-controlled’? Essentially, for each of the quasi-market types in Figure 2.2, *how do we know it when we see it?* A second pertinent limitation of Gingrich’s discussion is that of boxing in and marking an analytically unnecessary and unhelpful series of binary distinctions between particular forms of quasi-market. To some degree in any heuristic device this scalpel-sharp cleanliness of separation between different forms is needed for conceptual clarity and aids in conveying heuristically a new argument. Yet in this respect Gingrich’s (2011) significant original framework misses the degree to which there may be both gradation in dimensions, hybridity across dimensions, and a degree of fuzziness at the edges of each ‘type’.

To progress further the conceptual understanding of quasi-market varieties this chapter then responds directly to weaknesses in the forerunner framework, firstly by offering and extending a clear articulation of the underlying dimensions and the traits that align with particular quasi-market types (by developing the labelled axes in Figures 2.3 and 2.4) and secondly by conceiving and visualising these axes as continua such that the framework can be more alive to subtle variations in the forms of quasi-market under study.

2.6.1 Dimensions of variation in public service quasi-markets – allocation

The analytical framework developed within this thesis, and accepted for publication elsewhere (Carter, 2018), begins by considering the allocation dimension (Gingrich, 2011). The thesis makes an original contribution by separating the allocation dimension into two new underlying components: financing and regulation. These are visualised in Figure 2.3 below.

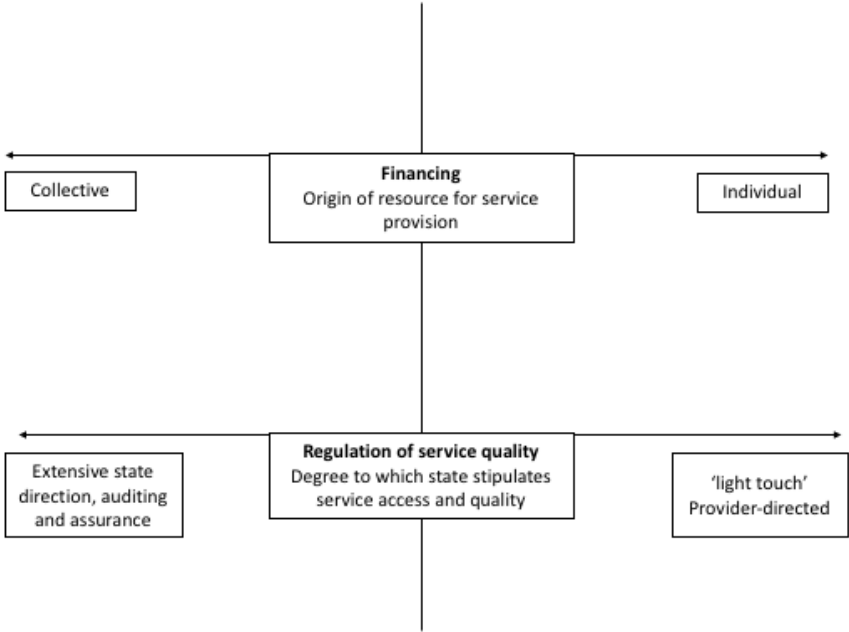


Figure 2.3 Visualising the allocation dimension in public service markets (an original illustration informed by and extending from Gingrich, 2011 and Wiggan 2015a, 2015b)

The first allocation dimension relates to whether services are collectively financed by society (i.e. provision costs are publicly funded through general taxation) or whether provision relies, in full or in part, on private resources of individuals (for example, through co-payments, means-testing, or incentives for private consumption). This financing dimension is shown in the upper portion of Figure 2.3 and the continuous scale seeks to reflect the ways in which the resource for service provision may be structured in a range of ways. Unlike Gingrich's table in Figure 2.2, which proposes a binary distinction between 'collective' and 'individual' financing, Figure 2.3 suggests instead that there are a range of graduated potential options. At the extreme poles of the axes sit the 'pure' forms of wholly collective (leftmost) and individual (rightmost) financing, but between these bounds services may be funded by incorporating collective *and* individual resources to varying extent.

The ultimate source of service funding, while a key dimension of variation across alternate market structures, is however not the only facet which can be understood as undergirding the pattern of service allocation.

The second allocation component captures the degree and strength of regulation for provider activity, with an axis that spans from extensive state directed auditing and service assurance through to a light-touch regulation framework. This component therefore incorporates the suggestion from Wiggan that a high level of state-led regulation "ostensibly promotes equitable access to services and underpins quality" (2015b, p. 155) whereas weak regulation enables providers to more easily pursue profits by targeting the most lucrative service users and reducing costs by scaling back the quality of provision or reducing services for riskier and/or poorer populations (Gingrich, 2011; Wiggan, 2015a). The lower portion of Figure 2.3 draws out a spectrum of regulation ranging from extensive state-directed auditing and assurance on service access (leftmost) to, at the other extreme, a very light-touch stance on service quality and access (rightmost) and seeks to capture the degree to which the 'cost' of poor or partial provision may be cascaded to service users.

Overall, Figure 2.3 expands on previous work by visualising the underlying allocation concept as a graduated dimension with 'more market-like' arrangements as the rightmost options on the spectrum, whilst the left-hand side is further from the market and actively seeks to mediate the pure market lever. In sum, the allocation dimension captures "whether price and selection mechanisms are used to allocate services, or whether services preserve a strong collective guarantee of access and funding" (Gingrich, 2011, p. 10).

2.6.2 Dimensions of variation in public service quasi-markets – production

Alongside considerations of allocation, the second key route through which quasi-market arrangements vary is in the realm of 'production'. The production dimension covers choice and competition: the way that a service is 'produced' within the quasi-market. In his application of Gingrich's framework, Wiggan (2015b) has argued that different combinations of competition and choice will influence the priorities of

service production. Depending on the configuration of provider competition, commissioner control and choice function services will be differently responsive to the needs and (dis)empowerment of either the state, service users or providers (Wiggan, 2015b).

Figure 2.4 expands and visualises the discussion offered by Gingrich (2011) and Wiggan (2015a, 2015b). The first sub-component (upper portion of Figure 2.4) relates to how ‘open’ the market is to new provider entrants, and captures the degree to which the quasi-market arrangements service the hegemony of specific or incumbent providers. This spans from an open, highly contestable and accessible market (far left) to a situation where access is constrained or limited (and therefore appears as an oligopoly), be that through restrictions on the size or sector of ‘eligible’ providers or other facets (such as contract duration, likely sanctions for failing to meet contractual obligations, or pre-qualification for new provider status).

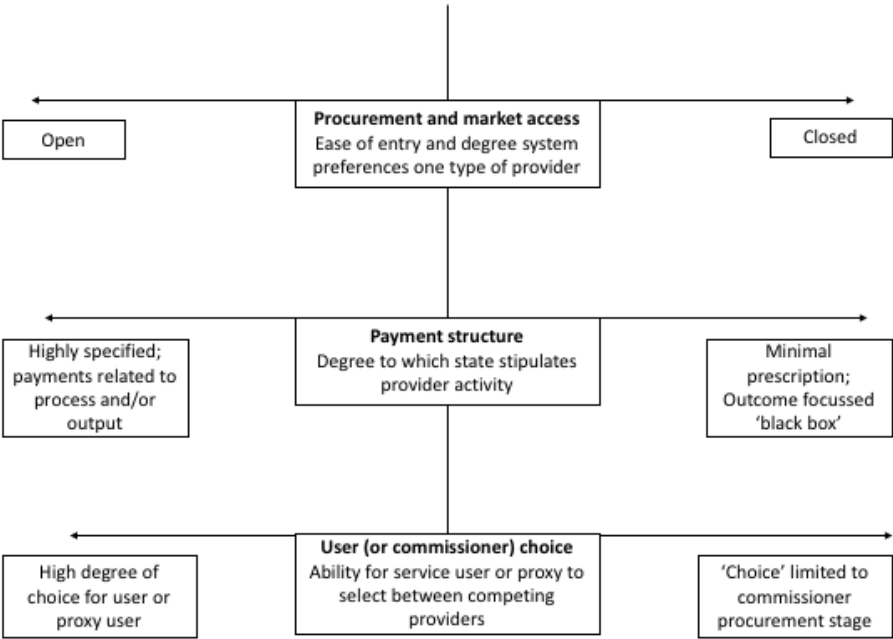


Figure 2.4 Visualising the production dimension in public service markets (an original illustration informed by and extending from Gingrich, 2011 and Wiggan 2015a, 2015b)

The second axis (in Figure 2.4) captures the control apportioned to the state (vis-à-vis service providers) in the design and stipulation of provider activity. At the left-most extent, there is a high degree of public sector control over provider activity since payment to providers is explicitly tied to the delivery of fixed service components. Shifting rightwards, the dimension portrays an increasingly outcomes-led approach whereby service providers are afforded the freedom to innovate and design flexible services geared to pursuing specified outcomes but where the public sector takes a position of agnosticism on the *means* of securing these outcomes.

The final axis in the lower portion of Figure 2.4 captures the function of user choice and voice within service production (Hirschman, 1970). Here, the left-most portion of

the axis conveys a situation where service users are powerful agents and where the choice and/or voice of those citizens who use a service provide a powerful steering rein to the behaviour of providers. By contrast, at the right-hand extent service users have limited if any ‘choice’ and their perspectives therefore have no traction for service provider behaviour. For each axis within Figure 2.4 a position closer to the extreme right-hand side reflects a stronger *provider* position, be that through minimal competition (upper axis), minimal state control (middle axis) or minimal responsiveness to user preferences (lower axis).

2.7 Varieties of quasi-market: where non-market accountabilities overlay marketised governance

What becomes apparent in mapping out the dimensions of variation in quasi-markets is the degree to which ‘variation’ can in part flow from the ways in which non-market accountabilities overlay and intersect with the ‘marketised’ forms. For example, in relation to the allocation dimension it is possible for the quasi-market to take a formulation that is a closer approximation of natural clearing markets (the right-most portion of Figure 2.3) but other *non-market* accountabilities can be used to ensure service access, for example, with the accompaniment of procedural rules, norms and regulation – perhaps through the regime of minimum service requirements or via corporate accountabilities, with targets set around user access and waiting lists.

Relatedly, on the production dimension each axis can be seen to pivot from a relatively unfettered, provider-dominant situation (right-most portion of Figure 2.4) through a ‘softened’ or *hybridised* form (in the middle portion of the axes) to a situation (in the left hand portion) where another stakeholder’s set of preferences are more fully met.

Considering the middle axes of Figure 2.4, at the left-most extreme state preferences for control over service provision, equity, and efficiency are prioritised. Specifying core components of services that are to be received by all participants holds up procedural fairness (i.e. overlays procedural accountabilities). In situations where it is possible to tightly specify service components which are known to correlate with programmatic outcomes this position can drive efficiency by minimising activities which are not known to influence the public service outcomes of interest (this might be achieved through overlaying corporate accountabilities). This can be considered as equivalent to coupling procedural or corporate accountabilities to the marketised arrangement.

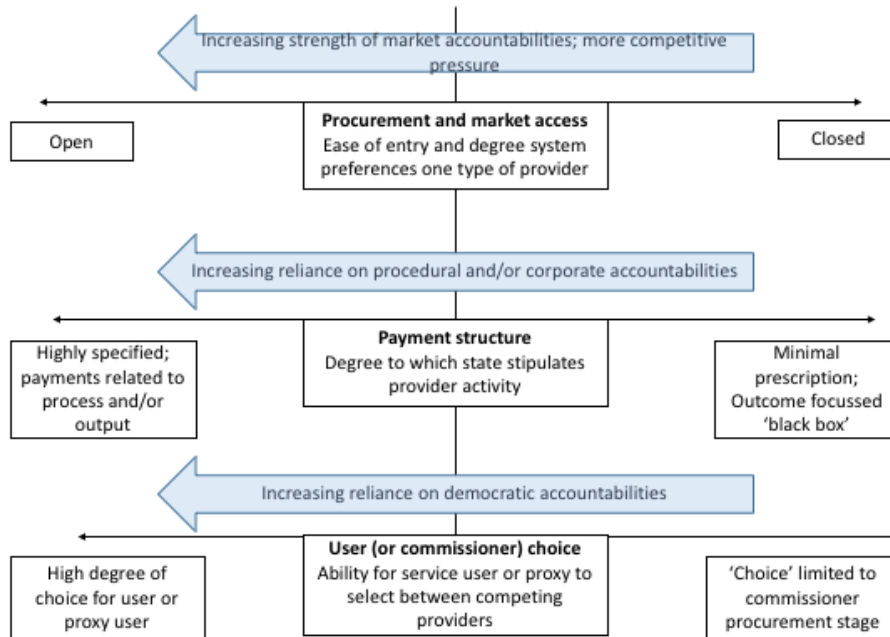


Figure 2.5 Overlaying non-market accountabilities on the production dimension

On the bottom axis of Figure 2.5, the left extreme of the chart strongly exerts service user preferences for quality, responsive services. This can be understood as analogous to the overlaying of democratic accountabilities to the otherwise marketised arrangement such that providers are strongly beholden to user priorities.

The upper portion of Figure 2.5 makes for a more challenging connection to an alternate *non-market* form. At the right most portion of this axis incumbent providers are able to pursue their own preference for profit maximisation relatively freely. Arguably at the left-hand portion there is ‘more market’ accountability as providers effectively become responsive to the actions of *other* providers and compete to remain in the market. This competitive pressure would be expected to act as a counterweight to the more cynical or malfeasant routes to profit maximisation (setting inflated, non-competitive prices and/or poor or partial performance) as incumbent providers with poor performance or expensive provision would need to adapt to become more efficient or else be forced to exit from the market place (through non-renewal of contracts and loss of market share).

Stimulating and maintaining competitive pressures within a *quasi-marketised* arrangement, however, is likely to require considerable ongoing ‘market stewardship’ on the part of policy makers, who will likely need to correct as they go in order to ensure the maintenance of competitive pressures across providers on an ongoing basis. The irony then is that considerable government activity may be required in order to amplify the competitiveness of a quasi-market, for example, by facilitating market share shift to better performing providers, or by calculating performance tables. In a highly competitive market it may also be the case that government efficiency preferences will be met.

In this way the variation in quasi-markets can be connected back to the wider discussion on accountability types whereby the presence of non-market accountability levers can be understood as introducing slightly different *forms* of quasi-market. A system which embodies market governance takes on a particular character or formulation in accordance with the accompanying accountabilities. This chimes with the description offered by Jantz et al., (2015, p. 19) who indicate that “[M]arket accountability is not, per se, a displacement of other accountability forms, rather it seems to co-exist and intersect with administrative and democratic accountability as policymakers oscillate between encouraging competition and freedom in delivery with control to cope with the unwanted consequences of market mechanisms”.

It is only by reading the full web of accountabilities at play that we can fully understand the pressures and imperatives driving the shape and form of service provision in any given context, including more narrowly in the various potential forms of quasi-marketised service provision. The analytic framework developed here then offers a substantive extension and improvement to Gingrich’s forerunner framework as it: firstly, functions as an explicit tool for identifying particular quasi-marketised forms; secondly, supports analysis to unpick the likely dominance of any stakeholder – be that state, user, or provider – in shaping provision; and, thirdly, hints at accountability levers that might be used as correctives to recalibrate quasi-market systems in situations where perverse behaviours, as a result of a single stakeholder’s dominance, are leading to unintended and undesirable outcomes. The utility of the analytic framework will be demonstrated by mapping British employment support programmes against these axes in the subsequent policy chapter whilst the framework’s comparative credentials will be explored in Chapter 8.

2.8 Network accountabilities – the missing dimension

The obvious accountability elephant in the room within this discussion is the *network* dimension. This network accountability lever would in principle function in such a way that provider behaviour is responsive to collaborative norms and shared obligations. The literature refers to network governance as capturing a ‘relational’ dynamic between service purchasers and providers, as per Williamson’s (1975, 1979) work on relational contracting. According to Considine, (2001, p. 30) network governance attempts to avoid the potentially predatory environment of marketised relations by proposing interdependence as a “binding characteristic”. This form of network accountability can then be considered as replacing the *vertical* hierarchy between what might otherwise be principal-agent relations. Considine (2001) additionally introduces the notion of network governance as co-operation and the co-production of results across multiple provider organisations. The survey tools used in Considine’s study to detect this form of governance at the frontline capture situations where a high value is placed “on trust and maintaining good contacts with clients and other service providers, including competitors. Adherents to this approach were presumed to be most likely to say ‘when you get good results with clients it’s usually a joint effort by yourself, the training person, the employer, etc’” (Considine, 2001, p. 34). This can then be seen as *horizontal* network formation, where a marketised

dynamic of competition is superseded by a networked logic of co-production as a process for assembling complex services.

Notably, network accountability is wholly absent from Gingrich's typology and does not feature anywhere in her discussion of quasi-market variation. This absence may be expected on three fronts. Firstly, it is not immediately clear how the utilisation of network accountabilities might be squared with her highly partisan account of market making endeavours. Secondly, there is also the issue that network governance has generally been understood as more hypothetical than real or as an 'emergent' public sector governance model (Bovaird and Löffler, 2009), including in the case in British welfare-to-work provision (Whitworth and Carter, 2018). Therefore, given that Gingrich's account is highly reliant on real-world case studies there is arguably less ability to integrate a consideration of networks. Thirdly, for Gingrich and others who have applied her framework the *production dimension* captures the intersection of choice and competition. A network accountability lever – in hypothetical form – does away with 'competition' and, indeed, is understood to have emerged "partly in response to concerns about NPM's hard-edged competitive outcomes" (English and Skellern, 2005, p. 2). The absence of network accountabilities is however problematic for the original analytic framework developed in this chapter given that it seeks more broadly to comprehensively trace the key dimensions of variation between *potential* quasi-market types. Unlike Gingrich, the device developed here aspires to provide a frame by which key varieties of quasi-markets and their hybrids can be considered *even where these do not exist in practice*. To achieve this it is crucial for the dimensions not to be restricted in conception to what is currently present but rather to probe and frame in such a way that we can trace future potential quasi-market possibilities. The ability to identify potential, but as yet, not implemented quasi-market forms is expected to be particularly valuable in supporting policymakers to identify additional accountability levers which may overcome weaknesses experienced in the quasi-market designs of the present.

In response to this absence of network accountabilities, Figure 2.6 outlines a potential additional dimension against which public service formulations may be considered. The concluding chapter of the thesis returns to more fully consider network accountabilities and address the absence of this axis both in the quasi-market framework and, as will be seen, though the dearth of networked considerations in the specific case of the British welfare-to-work quasi-market.

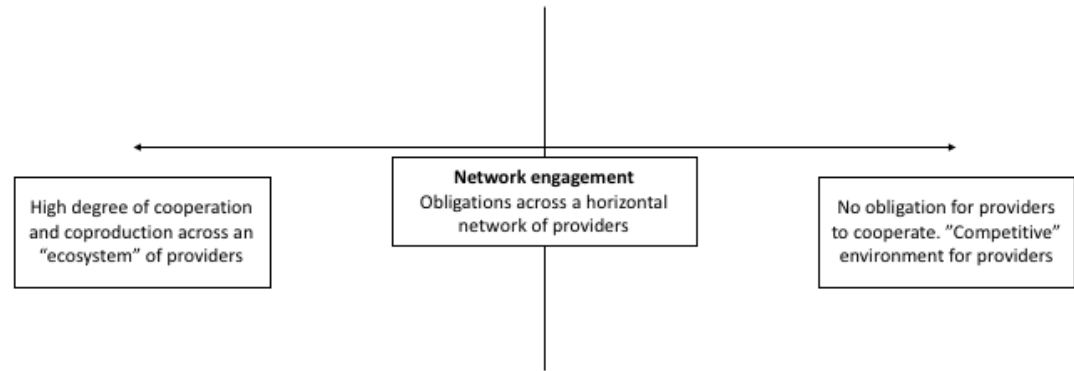


Figure 2.6 Considering horizontal network accountabilities: a hypothesised axis of variation introducing collaboration across multiple, interconnected provider organisations

The relational dynamic between multiple provider parties is anticipated to be particularly relevant in situations where citizens experience ‘complex’ and compound social challenges which sit across several commissioner domains, for example, in the arena of welfare-to-work where childcare challenges, housing insecurity, health issues, and unemployment intersect. Following from previous work by this author (Whitworth and Carter, 2018) the concluding chapter poses network accountability as a viable route to holistically combine streams of support in an ecosystem that ‘wraps around’ service users as a potential solution to some of the weaknesses encountered in current quasi-marketised governance configurations.

2.9 Conclusions and charting a course to operationalise the analytical framework

This chapter draws on the lens of governance and accountabilities to register the alternative modes of coordination underpinning the delivery of public services. This focus stems from an acknowledgement that the nature and content of social policies are crucially informed by their governance, that the logics of particular governance forms and accountability levers will have important implications for service user experiences and outcomes. The ascendance of market-inspired modes of governance is documented as a particularly dominant and heady trend. While previous studies have considered the rise of ‘quasi-markets’ as an undifferentiated and unified set of reform processes the discussion here is informed by and extends from the ‘difference thesis,’ which suggests that it is more appropriate to fracture the marketisation trend and introduce varieties of quasi-marketisation.

The original analytical framework developed across the chapter makes three particular academic contributions. Firstly, the chapter advances from Gingrich’s (2011) forerunner framework by unpacking aspects of public service allocation and provision to more comprehensively outline the underlying dimensions – that is – acknowledging a more detailed set of concepts which comprise each of these. The allocation dimension captures the form of financing for services (ranging from wholesale public funding to fully private provision) and the regulation of provider activity by the state (stretching from very light-touch or negligible regulation through to comprehensive

scrutiny and service quality control). The production dimension captures three underlying axes:

- i) How ‘open’ and competitive the market is, spanning from an open, highly competitive and accessible market to a situation where access is constrained or limited;
- ii) The specification of provider activity, ranging from a high degree of public sector control over provider activity to an increasingly outcomes-led approach, where the public sector takes a position of agnosticism on the means of securing specified outcomes; and
- iii) The role of service user choice and voice, spanning from a situation where the preferences and concerns of participating citizens act as a powerful steering mechanism for provider behaviour to a state where service user choice and/or voice is constrained to such a degree that it has no traction for service provider actions.

This offers a clearer articulation of the differentiators of alternate market varieties and the traits that align with particular quasi-market types. Depending on the design of particular quasi-markets and their configurations with respect to these dimensions, providers and their service delivery activities are expected to be differently responsive to the priorities and preferences of either the state, service users or providers themselves. This in turn is expected to have implications for the quantity, quality and distribution of services and consequently on any social outcomes achieved through such provision.

The second contribution flows from visualising these dimensions as continua such that the framework can be more attentive to subtle variations in the forms of quasi-market under study. The framework moves beyond Gingrich’s clear but unrealistic binary distinctions, for example, between fully public or fully private financing, to consider a continuous scale of alternative positions where the implementation of any axes is blended and partial. Rather than the scalpel sharp distinctions in the grid offered by Gingrich, this visualisation of the underlying concepts opens up a tool through which to draw out *hybridity* across multiple quasi-market forms.

Finally, this analytical tool pivots readily between theoretical discussions and applied work. As demonstrated in the subsequent policy chapter, it is a powerful tool for operationalising the design features of particular marketised forms and systematically comparing across these – as illustrated in Chapter 8 which uses the framework to compare two quasi-marketised programmes. The tool then is accessible not only to *ideal* types but can be used to map the messiness of quasi-markets in practice and is a route to investigate hybridity, where quasi-markets appear to straddle multiple ‘market types’ as conceived under Gingrich’s preceding framework.

Given previous difficulties in making systematic comparison across multiple marketised formulations it is hoped that this analytical framework can be used as a foundational tool for future work which seeks to learn lessons by reading *across* quasi-

market forms. Chapter 3 moves next to discuss the recent evolution of welfare-to-work policies in the UK context and, within this, utilises this chapter's original conceptual framework to unpack the key employment support programme at the heart of the study's empirical interest: the Work Programme.

3 Evolution of the “welfare-to-work” agenda: shifting policy goal posts and applying new quasi-market approaches

3.1 Chapter summary

In common with other developed economies, since the 1980s the UK has made a decisive shift from a ‘passive’ to an ‘active’ welfare system in which eligibility for out of work benefits is tied increasingly tightly and explicitly to the stated obligation to seek paid work.³ Unemployment is no longer framed as a structural issue of demand management but has been reconfigured as a supply-side challenge of improving the *employability* of the unemployed, thus beckoning in activation as the policy prescription.

In order to set an understanding of the Work Programme in context this chapter begins with a brief discussion of the wider transitions that have taken hold in terms of the overarching policy objectives in the field of unemployment protection and employment support. The conceptual framework developed in the preceding chapter is then used as a tool to systematically unpack the form of quasi-market instigated within the Work Programme. This reveals the Work Programme to embody a *private power market* in which the preferences of service providers are dominant. This market form promises innovation but introduces particular risks where providers defend profits through cost-cutting at the expense of effective or high-quality services. Programme designers anticipated these pressures and sought to implement a suite of steering tools that retain the perceived innovative, effective and efficient impetus of provider-directed quasi-markets whilst ensuring that a complex set of public sector objectives are met.

This discussion sets up the rationale and need for the empirical work developed across the remainder of the thesis, which questions whether these design tools facilitate programme objectives to support *more* participants into work, *faster*, *for longer* and whilst *reducing gaps in performance outcomes between the easier- and harder-to-help* (DWP, 2010a).

3.2 Policy trends in unemployment and “worklessness”

Subsequent to its post-war origins there have been important changes in what Torfing (1999, p. 7) describes as the “basic contents” of the modern welfare state. Credence has been placed on Bob Jessop’s description of the shifting functioning of the welfare state from the traditional post-war ‘Keynesian welfare state’ to that of a ‘Schumpeterian workfare state’ (Torfing, 1999; Peck and Jones, 1995; Dingeldey, 2007; Deeming, 2015). The key concerns of the Keynesian welfare state were to maintain full employment, primarily through demand-side management (Jessop, 1993) and to provide security and protection from Beveridge’s ‘five giants’ (Deeming,

³ Indeed, under Universal Credit in-work benefits are now also subject to conditionality and recipients are obliged to seek more and/or better payed employment.

2015). Here citizenship captures not only civil and political rights, but social rights as well (Marshall, 1950).

For Jessop the crisis of the Fordist mode of production and Keynesian welfare state are intimately entwined and he suggests a shift to an emergent Schumpeterian workfare state. Peck and Jones (1995) elaborate on the key distinctions between the Keynesian welfare state and Schumpeterian workfare state and features of the transition between the two modes (as outlined in Figure 3.1). The shift between the two models is signalled by three overarching characteristics: firstly, the promotion of product, process, organisational, and market *innovation*; secondly, the enhancement of the structural competitiveness of open economies mainly through *supply-side intervention*; and, thirdly, the *subordination of social policy* to the demands of labour market flexibility and structural competitiveness (Jessop, 1993, p. 9, emphasis added). What becomes increasingly clear is that the shift in accumulation regimes and modes of social regulation is aligned with substantively altered policy understandings and prescriptions in relation to (un)employment.

	Keynesian welfare state characteristics	Crisis	Schumpeterian workfare state characteristics
Macroeconomic policy orientation	Countercyclical demand management. Interventionist policies supporting economies of scale and productivity growth	Fiscal crisis of the state triggered by falling profitability, declining tax base, and growing welfare demands	Supply-side innovation through goals of structural competitiveness, economies of scope, and continuous, flexible innovation
State discourse	Productivity and planning	Monetarism	Flexibility and entrepreneurialism
Wage relations	Collective bargaining. Promotion of norms of mass consumption	Crisis of profitability triggered by international competition. Pressure on wage rates	Flexible wage systems. Income polarisation. Heightened segmentation
Social policy	Progressive redistribution. Based on universalist welfare rights and social needs	Fiscal crisis of the state and welfare retrenchment	Productivist cost-saving concerns. Subsumed to goals of labour flexibility and business needs
Unemployment challenge	Lack of income		'Social exclusion' and limited employability
Policy response to 'unemployment'	Income protection for male breadwinners		(Re)insertion into labour market

Figure 3.1 Shifting functions of the welfare state. Source: extending from Peck and Jones, 1995

The unemployment policy challenge under the Keynesian welfare state was seen as cyclical and to be tackled as a demand side issue. The policy 'fix' is the replacement of male-breadwinner earnings via (presumed to be temporary) income protection measures (Deeming, 2015).

Contrastingly, under the Schumpeterian workfare state, unemployment is considered a structural challenge to be tackled by incentivising and enabling people who are unemployed to move (back) into the labour market (Deeming, 2015). This diagnosis

of unemployment suggests that it is damaging not simply because of the loss of earned income for people who are out of work but because of the way it inflicts a longer term ‘scar’ through the increased future incidence of unemployment and lower subsequent earnings in work (Arulampalam et al., 2001; Gregg, 2001). Unemployment is separately acknowledged as being detrimental to health and wellbeing (Clark and Oswald, 1994; Dolan et al., 2008; Paul and Moser, 2009; Waddell and Burton, 2006). The phenomenon of hysteresis (Blanchard and Summers, 1987) has gained traction, whereby longer-term unemployment (and the depreciation of human capital associated with this) serves to reduce “the effectiveness of the unemployed as potential fillers of vacancies” with the result that the (natural) equilibrium rate of unemployment is expected to track upwards (Layard et al., 1991, p. 4; see also Arestis and Sawyer, 2001). The role of the state consequently moves away from protection and is redefined in order to secure labour market participation. The recommended response is a suite of policies “where people are helped to sustain their own incomes rather than depending on long-term transfers” (Boeri et al., n.d., p. 3). Social policy pivots away from the satisfaction of income needs, redefines unemployment as a challenge of improving the *employability* of the unemployed, (McQuaid et al., 2005; Lindsay et al., 2007) and beckons in ‘activation’.

3.3 The activation turn

Across the OECD there has been an increased acceptance since the 1990s of the need for activation measures to strengthen the link between social protection, labour market policies and employment (Escudero, 2015). Activation has been promoted by both the OECD (subsequent to the OECD Jobs Study, 1994) and European Commission (following the White Paper on “Growth, Competitiveness and Employment,” 1993). Promotion of this policy approach emerged from cross-national comparative work which assessed relative levels of joblessness and economic performance and how they were influenced by labour market institutions and welfare state arrangements (Martin, 2015; OECD, 2013). This particular assessment suggested that variation in unemployment rates across developed economies was related to their alternative approaches in response to economic recession, specifically that higher levels and persistence of long-term unemployment were associated with weaker and less well co-ordinated delivery of employment services and the absence of benefit conditionality (OECD, 2013).

Activation policies respond to this assessment and are defined by their intent to help people of working age (back) into the labour market (Daguerre and Etherington, 2009). At a high level the solution is to re-orient welfare states towards becoming ‘enabling states’ (Gilbert, 2002) or from ‘passive’ to ‘active’ welfare systems (Lødemel and Trickey, 2001). This involves a shift in formal policy where employment, participation and activation have become the leitmotiv (van Berkel and van der Aa, 2005). While ‘passive’ measures are intended to alleviate temporary economic hardship during periods of unemployment, ‘active’ approaches aim instead to remove obstacles to

employment, motivate jobseekers to actively seek and accept employment, and/or to retain or improve jobseekers' employability (Weishaupt, 2011; OECD, 2013).

There are a range of reasons why Active Labour Market Policies (ALMPs) have become such a popular policy tool. Not only does activation align with the pursuit of non-inflationary growth by increasing labour supply and driving down the non-accelerating inflation rate of unemployment (NAIRU), it promises to reduce the number of income maintenance recipients, thus keeping social spending in check (Daguerre and Etherington, 2009). Activation also tallies with additional concerns of Western European governments at this time: it addresses the 'new social risk' of skills obsolescence in a post-industrial era; relatedly, it engenders a flexible and adaptable workforce better able to respond to the challenges of globalisation; it drives down the dependency ratio in a 'greying' Europe; and neatly aligns with the consensus that "paid employment is the best way out of poverty" (Daguerre and Etherington, 2009, despite, in practice, the persistence of in-work poverty, see Kemp et al., 2004).

3.4 Beneath the banner of activation

Whilst there has been international convergence to activating labour market policies across the developed economies (Bonoli, 2010; Lødemel and Trickey, 2001; Peck, 2001), Barbier (2004) notes that 'activation' has become a fashionable policy slogan whilst concealing the diversity of associated reforms. ALMPs do not represent a homogenous set of policies but, rather, reflect a diverse array of potential approaches encompassing varying possible aims, actors, instruments and consequences (Theodore and Peck, 2001; Barbier, 2004; Levy, 2004; Bonoli, 2010). In seeking to summarise this variability, the literature typically distinguishes between two stylized ideal types that are described as "human capital development" and "work first" approaches (Theodore and Peck, 2001; Lindsay et al., 2007) and that Levy (2004), with similar meaning, describes as 'thick' and 'thin'. Figure 3.2 summarizes the key, and somewhat caricatured, qualitative distinctions between these two ideal type activation regimes.

Dimension	Thin – ‘Work First’	Thick - ‘Human Capital Development’
Aims	Quick return to labour market	Greater focus on employability (i.e. up-skilling), job quality and in-work progression
Programme targets	Job transitions	Sustained employment transitions; reducing distance to labour market for the ‘harder to help’
Intervention model	Job-search, basic skills training and focus on rapid transitions into jobs	More intensive, longer-term and personalised training and supports
Relationship to labour market	Demand-side focus on inserting jobseekers into available opportunities quickly. Work experience limited, often mandatory and typically unpaid where exists	Supply-side focus on up-skilling jobseekers to improve their short and long-term labour market prospects. Work experience likely to be available, claimants having greater choice and often paid (whether by employers or government subsidies)
Relationship with individuals	Emphasis on pushing claimants into rapid job transitions largely irrespective of quality or suitability	Greater emphasis on pulling claimants into employment via building employability and higher quality job opportunities (in terms of pay, conditions, progression, etc)

Figure 3.2 key features of activation ideal types. Reproduced from Carter and Whitworth, 2017

The UK regime is typically identified – even sometimes by government itself (DWP, 2003) – as a ‘work first’ regime situated as a close fit with the thin, liberal type (Lødemel and Trickey, 2001; Torfing, 1999). The enduring work-first orientation for UK activation policies is underscored through international longitudinal comparative work: the UK’s low levels of overall spend and, within this, its lower spending on training – which would be expected to underpin the human capital development approach – is particularly striking (Bonoli, 2010, 2013).

3.5 British activation objectives

A critical reading of the British activation model has argued that its main aim is to ‘encourage’ the unemployed and inactive to enter the labour market as quickly as possible via hassle-based push strategies, potentially by accepting low-paid or inappropriate jobs (Dean, 2003). This strategy has been pursued under academic guidance that work-first (as opposed to training-first) produces better labour market outcomes (Layard, 2004). Under this approach it is argued that “almost any job is better than no job” (Layard, 2004, p. 1).

Importantly, to probe beneath this conception of the UK, Weishaupt (2011) provides a summary of the range of policy tools at the intersection of labour market strategy, employment support, and benefits which may be brought to bear on the ‘subjects’ of activation, that is, the detailed policies and programmes subsumed beneath the broad brush of activation. Figure 3.3 summarises these ideas and importantly separates out more ‘stick-like’ push factors – the negative incentives both financial and non-financial, which respectively draw on hardship and hassle of unemployment to ‘push’

people into the labour market – from the more ‘carrot like’ elements – the positive incentives and supports which may ‘pull’ individuals into work. Importantly, these policy tools may not perfectly align with the idealised descriptions of thick and thin approaches provided above. Whilst activation sits as an overarching endeavour it rarely embodies a pure form of the ideal *thick* and *thin* approaches in practice but, rather, typically combines carrot-like and stick-like interventions. Even a ‘thick’ regime may use some ‘negative’ elements, such as conditionality. Indeed the presence of negative incentives may be core to a justification for ‘thicker’ positive supports. A range of policies, programme levers and ambitions are layered in a form of decoupage.

Negative Incentives (financial) NIF	Negative Incentives (non-financial) NIN	Positive Incentives (financial) PIF	Positive Incentives (non-financial) PIN
Short duration of unemployment benefit payments	Job-search requirements	In-work subsidies	Job counselling
Low, means-tested social wages	Broad definition of a "suitable" job offer	Tax credits	Training courses (soft and occupational skills)
Unavailability (or phasing-out) of "early exit options" e.g. early retirement or disability pensions	Strict eligibility criteria	Benefit top-up for participation in public work schemes	Childcare support
	Benefit sanctioning when in breach of benefit conditions	Self-employment start-up subsidies	Other services (mobility, mental health, debt, substance abuse)

Figure 3.3 Available policy tools to activating labour market policy regimes. Reproduced from (Weishaupt, 2011, p. 69). Categories of NIF, NIN, PIF and PIN are used subsequently

In the discussion below, by using Weishaupt’s classification of activation tools in Figure 3.3 which separate out the carrots (positive incentives both financial and non-financial) from the sticks (the slightly oddly labelled “negative incentives”) alongside policy documents which explicitly spell out the intent of particular devices it is possible to trace the evolving emphasis of recent UK activation policy.

3.6 Carrots, sticks, or a bit of both? Activation trends and the rise of ‘welfare-to-work’ in the UK

Commentators have been quick to note that there has always been a whiff of activation in the UK system (White, 2000; Trickey and Walker, 2001; Sainsbury, 2014). As a phrase ‘welfare-to-work’ has been used by governments since the 1980s to express a desire to replace ‘passive’ income replacement support for people who are unemployed with ‘active’ employment support measures to help and encourage people to enter paid work (Hirsch and Millar, 2004). Across the past four decades however there have been important shifts in the specificities of policy intent and approach where, for example, there are subtle but important differences between reducing ‘unemployment numbers’ (i.e. cutting the ‘claimant count’ of people receiving unemployment benefits) and increasing ‘employment rates’, which can be achieved through a range of means, including supporting those previously considered ‘inactive’ to enter the workforce. The following discussion segments these shifting logics and policy proposals according to three key administrative eras (although noting that there have been shifts in emphasis even within these, under different ministerial teams, for example).

Before entering this discussion, it is pertinent to note that whilst the literature frequently uses ‘activation,’ ‘ALMPs,’ and ‘welfare-to-work’ interchangeably, in the remainder of the thesis a distinction is made between activation (and ALMPs) and ‘welfare-to-work’. Throughout, the term ‘activation’ is used to capture the overarching governmental ambition to move people away from ‘inactively’ claiming unemployment or other social insurance and towards the labour market. More

specifically, ‘welfare-to-work’ is used to connote programmatic services and interventions directed at people who are unemployed to support/encourage/hassle them into employment. Welfare-to-work services have historically been provided in the UK context by staff within the Public Employment Service (Jobcentre Plus) and provision now increasingly draws in employment advisors from a range of wider private, public and voluntary provider organisations.

3.7 Era 1: Tory beginnings (1989 – 1997)

The Conservative administration of the 1980s and early-mid 1990s can be considered a period of nascent activation: “conditionality was emphasised, and measures were introduced to *encourage people to leave benefit*” (Trickey and Walker, 2001, p. 186 emphasis added). The left hand column of Figure 3.4 sets out the high level programmes and policy changes associated with this administrative era. By classifying these according to Weishaupt’s frame it can be seen that “negative incentives” (NIN) dominate in this period. Legislative changes were used to tighten benefit eligibility and strengthen conditionality.

Subsequent readings of this policy arrangement, which sought to reduce the number of people claiming unemployment benefit, suggest that a significant number of unemployment benefit claimants were diverted onto disability related benefits (Alcock et al., 2003; Beatty and Fothergill, 2005). Trickey and Walker (2001, p. 187) note that it was in this period where the term “welfare dependency” was introduced into the British lexicon and this was used as a tool to build consensus that “expenditure on benefits was too high”. Reducing the generosity of out-of-work benefits and bolstering benefit conditionality was framed by policy makers as an appropriate response to removing an assumed group of ‘voluntary’ claimants from the register (Tonge, 1999). These shifts towards a supply side understanding of unemployment are understood to be important precursors to the New Labour strategy that followed (Trickey and Walker, 2001; Tonge, 1999).

3.8 Era 2: New Labour’s Third Way (1997 – 2010): ‘creeping conditionality’ and quasi-markets

The election of the New Labour government in May 1997 represents something of a ‘big bang’ moment in welfare-to-work. Blair’s first administration saw the reform of the welfare state to be one of its major tasks (Powell, 2000) and a central objective was to shift people off the social security rolls and into paid employment with an active role for government in this process (Sheldrick, 2000). The approach to achieve this was distinct. New Labour’s ‘Third Way’ or “modernised social democracy” sought to reconcile liberal economics and social justice in a transformed international economy (Driver and Martell, 2000, p. 150; Clasen and Clegg, 2004).

A ‘Third Way’ approach to public policy encompasses a number of distinct features, including:

- Government redrawing the social contract with rights to welfare matched explicitly and tightly by responsibilities, especially regarding paid work (White, 2000);
- The welfare state working proactively to help individuals off social security, away from ‘dependency’ and into work. The rise of “employment-centred social policy” (Driver and Martell, 2000, p. 152) consolidates the shift from ‘passive’ to ‘active’ benefits;
- In the area of welfare-to-work, provision is increasingly delivered in partnership with private and voluntary sectors, with government in an enabling and steering rather than direct delivery role.

Conservative 1979-1997 Objective to reduce the number of people claiming unemployment benefits; Targeting support through means testing (Hills, 1998)	New Labour 1997 – 2010 Increasing employment rate; Poverty reduction; “Making work pay” (Evans and Millar, 2006; Hirsch and Millar, 2004)	Conservative-led Coalition 2010 – 2015 “Fairness for taxpayers” – i.e. households claiming out-of-work benefits should not have higher income than those in work; “Better off in work”; Reducing “worklessness” (McKnight, 2015)
Early 1980s price-linking social security benefits rather than maintaining relativities with incomes, eroding the value of out-of-work benefits (Hills, 1998; Sutherland et al., 2008). <i>NIF</i>	1998 Introduction of New Deal for Young People (a programme including subsidised jobs for young people who are unemployed). Followed by a series of other “New Deals” for Lone parents, long-term unemployed people, partners of unemployed people, disabled people and people over 50. (Hirsch and Millar, 2004; Sunley et al., 2006). <i>Principally PIN but NDYP backed by sanctions for non-participation.</i>	2011 Work Programme replaces all existing welfare-to-work programmes and has ambition to support more people to move into work. <i>PIN</i>
1986 Introduction of Restart Interviews where people unemployed for more than six months were required to demonstrate the active steps they were taking to find work. (Wilson, 2013). <i>NIN</i>	April 1999 National Minimum Wage becomes law and is followed with frequent uprating (Bennett and Millar, 2005). <i>PIF</i>	Announced 2010 and incrementally implemented from 2012 Universal Credit replaces six of the main means-tested benefits/tax credits for working age people with a single benefit paid to people out of work and in low-paid work. UC applies tougher and more extensive conditions and sanctions than previous systems extending conditionality to those who are in paid work (Dwyer and Wright, 2014). <i>Originally PIF, now less clearly so.</i>
1996 Introduction of ‘Jobseeker’s allowance’ replacing income support and unemployment benefit (Strickland, 1996). Claimants must agree to sign a jobseeker’s agreement as a condition of receiving benefit (ibid). <i>NIN</i>	1998 National Child Care Strategy promises early education places for all 3- and 4-year-olds by 2004 and good quality affordable childcare provision (Hirsch and Millar, 2004). <i>PIN</i>	2012 JSA revised sanctions regime extends the severity of sanctions. People who are deemed to have repeatedly breached the conditions of benefit receipt could forfeit their benefits for up to three years (DWP, 2013a; McKnight, 2015). <i>NIN</i>
	1999 Working Families Tax Credit, implying significant real increase in means-tested support for families with children (Bennett and Millar, 2005; Brewer et al., 2005). <i>PIF</i>	2010 onwards Replacement of Incapacity Benefit and related benefits by Employment and Support Allowance (ESA), with more stringent medical tests, greater conditionality and time limiting of non-means tested entitlement for all but the most severely ill or disabled (Scottish Parliament, 2013). <i>NIF</i>
	2001 Mandatory Work Focused Interviews introduced for lone parents (Whitworth, 2012). <i>PIN/NIN</i>	
	2008 Replacement of Incapacity Benefit with Employment and Support Allowance (ESA) for new claimants brings compelled work preparation to those placed in the work-related activity group. <i>NIN</i>	

Figure 3.4 Three welfare-to-work eras and the nature of key activating policy tools

New Labour's approach to welfare-to-work was also premised on a particular understanding of the unemployment problem. As indicated by Lister (in Powell, 2000, p. 46), New Labour "is concerned with the 'supply side' of employability rather than the demand side of unemployment" which is rejected as 'old Left' and not feasible in today's global economy (Blair and Schroeder, 1998). Connecting the workless with labour markets – indeed, with society (Finn, 2000) – is therefore the policy problem to overcome.

The 'New Deals' – by providing a mixture of training measures, socially useful work experience and targeted subsidies to private employers – were presented as the flagship welfare-to-work policy mechanism to reattach those without work to the labour market. This was coupled with a fiscal strategy to 'make work pay' via policy measures such as the National Minimum Wage, Working and Family tax credits (Brewer and Shephard, 2004; Bennett and Millar, 2005) and additional support for those in work through the National Childcare Strategy (Whitworth and Griggs, 2013). These enhanced supports were used to ethically justify the more stringent conditionality placed on the receipt of out-of-work benefits.

In turn, over the course of the late 1990s and 2000s welfare-to-work policymaking has been characterised by a 'creeping conditionality' (Dwyer, 2004). The 'ratcheting up' of work-related behavioural requirements means that conditionality was, by the end of the New Labour era, radically wider, deeper and more quickly applied than in 1997 (Griggs and Bennett, 2009). Work-related obligations have been stretched to encompass traditionally inactive groups – such as lone parents and disabled people – with the size of the group who 'cannot work' diminishing over time (Dwyer, 2004; Griggs and Bennett, 2009).

The publication of the Blair-commissioned Freud Report in 2007 forms a further critical juncture in the evolution of welfare-to-work policy under New Labour (Rafferty and Wiggan, 2011) and the subsequent reinvigoration of this policy area under the Coalition government. Grover (2007) suggests that the report dovetails neatly with New Labour's fixation on an individualistic supply-side conceptualisation of the unemployment problem and resultant policy focus on enhanced conditionality and 'market-making' amongst service providers. Freud's (2007) report is an important galvanizing set-piece as it braids together the policy 'ends' of welfare-to-work with a set of recommendations which position large-scale constructed quasi-marketised arrangements as the key implementation 'means'.

At the heart of the report is a recommendation for restructuring and outsourcing the provision of welfare-to-work services by splitting provision such that Jobcentre Plus (the UK's PES) provides "streamlined, mass market provision" (Freud, 2007, p. 10) for claimants in the first year in which they are workless and considered to be able to work and relatively close to the labour market. Freud then proposes outsourcing and competition for "the hardest to help" (2007, p. 10). The report suggests a significantly expanded role for contracted return-to-work providers from the private and voluntary

sectors with contestability through competition in delivery. Beneath the overarching proposal for greater outsourcing Freud also makes particular recommendations for how this external provision ought to be arranged:

- **Personalised support** A non-prescriptive or ‘black box’ approach to commissioning means that providers have the freedom to innovate and provide tailored help to each individual;
- **Holistic single platform** Merging all welfare-to-work provision into a single platform with a ‘universal’ programme for all unemployed people;
- **Value for money and payment-by-results** An outcome focussed approach where spending on support services is justified through the subsequent welfare spending that could be ‘saved’. A central suggestion is that the majority of payment to providers should be made once they have successfully supported a move into employment;
- **Complex supply chains** Large regional contract areas with the Prime contract holders responsible for “marshalling an appropriate blend of subcontractors” with specialist voluntary sector organisations seen as crucial in providing expertise for supporting the hardest to help (Freud, 2007, p. 7).

Freud suggests that this reformed approach to welfare-to-work provision ought to sit within a wider activation landscape shaped according to enhanced conditionality and a rationalised benefit system. Asserting both that unemployment is now frictional rather than structural and a paternalistic belief that ‘work is good for you’ in all circumstances, Freud declares that enhanced conditionality and sanctions are needed to tackle the alleged behavioural causes of worklessness and the existence of a ‘dependency culture’, despite compelling evidence to the contrary (DWP, 2010b; Macdonald et al., 2014). Freud also supports the introduction of a single benefit for working age people that will provide a safety net and ‘incentivise’ the transition to work – later to become Universal Credit – the radical transformation/heavy millstone around the neck of the UK benefits system.

The logic which underpins Freud’s paper is informed by perceived benefits asserted to stem from the use of quasi-marketised services. Freud (2007) strongly advocated for provider-directed market arrangements, suggesting that the involvement of private sector and voluntary sector providers, competition, outcome-based payments and ‘black box’ delivery models would unlock innovation and lead to higher levels of performance compared to forerunner schemes. The report underscored the importance of provider discretion, suggesting that flexibilities will enable providers to “focus on the most efficient, and work-focused, interventions – in particular around intensive adviser support” (Freud, 2007, p. 44). Freud also advocated the use of largescale provider contracts in the belief that “[T]he large scale of the enterprises, and their management capabilities, are more likely to provide the conditions in which innovation can take place and quality be assured” (2007, p. 62) (2007: 62).

Importantly, the DWP’s (2008) ‘Commissioning Strategy’ largely replayed Freud’s recommendations and set out four key principles for the commissioning of future

welfare-to-work programmes: 1) Prime providers; 2) payment-by-results; 3) minimum service prescription; and 4) larger, longer contracts. These principles – particularly the outcome-focused payment model – were subsequently absorbed in New Labour’s Flexible New Deal (Wiggan, 2015). This scheme, however, would just months later be hastily scrapped and, within its Work Programme successor, rapidly be seen as a tentative step towards Freud’s recommendations.

3.9 Era 3: Conservative rocket boosters (2010 – 2018): enter the Work Programme

The UK has since the late 1990s seen remarkable cross-party agreement in approach – both in terms of the overarching objectives of policy (the activation ends) and through the quasi-marketised approach to service coordination (the activation means). This consensus has been aided by the transfer of key policy actors such as Lord Freud from Labour to the Conservatives in 2009 (the former banker was recruited and ennobled by the Conservatives and was subsequently made Minister for Welfare Reform). Hence, the arrival of a Conservative dominated Coalition government in the 2010 general election after 13 years of Labour government, and roll-out of its flagship Work Programme welfare-to-work scheme⁴ “does nothing to break out of the policy paradigm established by Labour; it simply takes it further and faster” (Lister and Bennett, 2010, p. 102).

3.9.1 Introducing the Work Programme

Although reflecting continuation from the Flexible New Deal in terms of underlying principles, the Work Programme, introduced in June 2011, is in comparative perspective a genuine revolution – if not large-scale experiment – in employment support policy, most notably in terms of the extent of sub-contracting and payments weighted to job outcomes as well as the ‘black box’ model of delivery.

The Work Programme replaced virtually all contracted welfare-to-work programmes run by the DWP in England, Scotland and Wales. It offers support to people who are long-term unemployed or who are understood as at risk of becoming so and aims to help them get, and keep, jobs. It is a monolith in terms both of its size and its breadth. The Work Programme is expected to cost between £3 billion and £5 billion and could help up to 3.3 million people (NAO, 2012). The Work Programme closed for referrals in 2017 and at the time of writing is finishing delivery for participants on the programme.

Delivery of the Work Programme takes place through contracts between the DWP and large-scale, mainly private sector Prime providers which can both deliver services themselves and/or sub-contract to organisations within large and (sometimes) complex supply chains sitting underneath each Prime. The Work Programme is structured

⁴ New Labour’s Flexible New Deal scheme was swiftly cancelled following the election.

geographically into 18 large ‘regional’⁵ Contract Package Areas (CPAs) with two or three Primes in each CPA to whom claimants are randomly allocated from Jobcentre Plus if they have not found work within an initial period of Jobcentre Plus provision, the duration of which depends largely on the type of out-of-work benefit received and the Work Programme ‘claimant group’ into which they are therefore placed. Unlike the Flexible New Deal, which contained mandatory service components, a ‘black box’ delivery model operates in the Work Programme so that providers have almost complete flexibility over their interventions. This flexibility is argued to be required since, unlike the various group-specific New Deals, Work Programme is ‘universal’ and therefore has to cater for the needs of a wide range of unemployed programme participants within a single employment scheme. The majority of Work Programme participants are in receipt of Jobseeker’s Allowance (JSA) which is the mainstream unemployment benefit claimed by people actively looking for work. Participants may also be claiming Employment and Support Allowance (ESA) or other incapacity related benefits where they have a health condition or disability recognised by DWP and affecting their ability to work.

Marketised accountabilities dominate and an aggressive payment-by-results model is the central tool through which DWP seeks to steer provider activity and performance. Providers are paid primarily – and since April 2014 exclusively – for the results they achieve in supporting people into sustained employment so what the provider earns is tied to how well they perform in terms of aggregate job outcomes. All payments are made on an individual, i.e. a per participant basis. There are three core components to the payment structure:

Attachment payment The attachment payment is a form of ‘start fee’ paid when participants join the programme. Each year the value of the attachment fee is reduced and by the start of the fourth year no attachment payment is made;

Job Outcome payment There is no payment made when a participant enters work. Instead, job outcome fees are claimed by providers once a participant has been in work for either a continuous or cumulative period of employment of either three or six months dependent upon the payment group. To be eligible for payment work must be for at least 16-hours per week;

Sustainment outcome payment Sustainment payments are made every four weeks following a job outcome payment for keeping a participant in employment for up to 26 months.

DWP, 2012; NAO, 2012

In its size, breadth and extent of marketisation the Work Programme is a vanguard of radical experimentation in welfare-to-work quasi-markets within comparative

⁵ The North East; Scotland; and Wales are examples of the scale of single CPAs.

perspective (Finn, 2011a; Mulheirn, 2011; this author in Carter and Whitworth, 2015). But as van Berkel et al., (2012b, p. 273) note, much of this research “pays little attention to the issue of diversity in the design and functioning of markets”.

There is then a gap and a need for a systematic consideration of the particular formulations or varieties of quasi-market and their links to user experiences and programme performance. It is to the task of opening up and systematically examining the specific market form constituted in the Work Programme that this chapter now turns. Utilising the rich analytical framework developed in the previous chapter provides a considerably more detailed conceptual lens than has been available previously in considering the practical implementation of the marketised form.

3.9.2 The particularities of the quasi-market configuration embodied by the Work Programme: the allocation dimension

The assessment of the flavour of quasi-market embodied by the Work Programme begins by investigating the allocation dimension. Against the formal financing component the Work Programme is collectively funded – as in previous UK employment support programmes – and is therefore positioned at the left-hand side of the upper axis in Figure 3.5. In principle, the Work Programme has implemented an innovative funding arrangement that draws from both the fixed Departmental Expenditure Limit and the Annually Managed Expenditure from which future benefit payments are made. This enables upwards flexibility in the financial envelope available to the Work Programme in the event of stellar programme performance. In practice this was never operationalised as performance never reached these levels, and in any case would not shift the arrangement away from collectively financed provision.

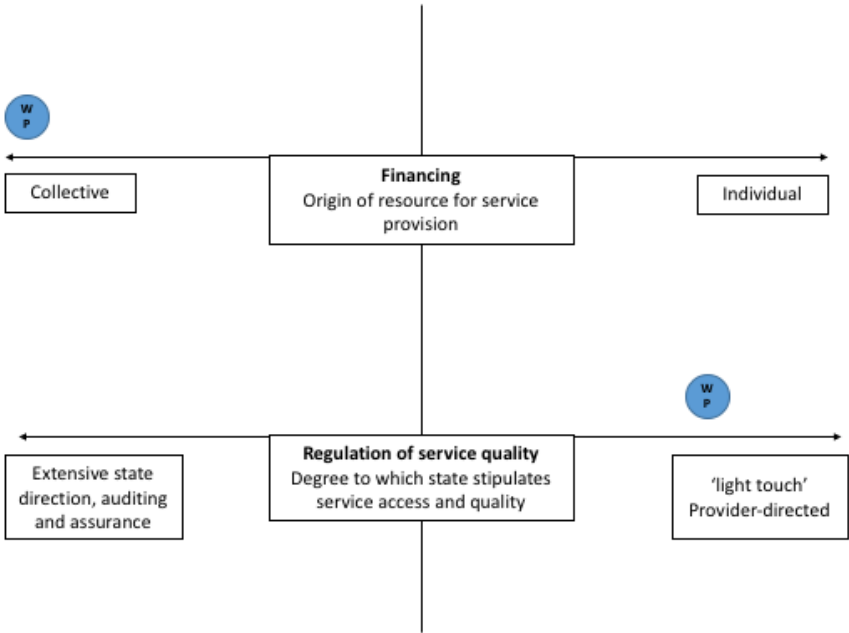


Figure 3.5 Visualising the allocation dimension for the Work Programme quasi-market

Turning to the regulation of service quality, there are two key programme components that relate to access and quality of provision. Firstly, in terms of entry to the programme, participants are referred to the Work Programme if they have not found work within a variable initial period of mainstream Jobcentre Plus provision. Timing of referral depends on participants' circumstances (such as age and benefit type). For the majority of those referred to the programme, participation is mandatory and failure to engage can result in benefit sanctions being applied (WPSC, 2013). Referral for mandatory participation is triggered automatically by Jobcentre Plus management information systems. Where claimants have the option to volunteer Jobcentre Plus work coaches refer by discretion. All claimants due to join the programme are then randomly referred to one of 2 – 3 Prime providers within their regional CPA.

Primes are then responsible for engaging with all participants and conducting attachment activity within 15 working days of the participant being referred by Jobcentre Plus (DWP, 2011a). In practice this entails an initial meeting and preparation of an action plan, after which the attachment fee is paid. Around 98% of referrals successfully achieve programme attachment, although some groups experience a lower conversion rate: only around 9 in 10 prison leavers are successfully attached for example.⁶

It is not clear whether DWP has taken any steps to ensure universal programme access for all those referred. The overwhelming majority of participants are (at least) *connected* to a Work Programme provider. At this level then, the procedural accountability of Jobcentre Plus underscores due process in 'access' to the programme.

Secondly, subsequent to formal attachment service access and quality are intended to be held up by 'minimum service guarantees' (MSGs), which are designed and published by Prime providers themselves. These service guarantees vary from provider-to-provider. There is no common floor standard specified by the DWP and many guarantees are vague, non-measurable and hence in practice, unenforceable. There are also questions as to how clearly these minimum expectations have been communicated to programme participants and there are concerns as to whether people participating in the Work Programme are aware of these service obligations (Whitworth, 2013). As argued by this author (Carter and Whitworth, 2015) and elsewhere (Finn, 2012; NAO, 2012) the well-known weaknesses in the MSGs undercut their effectiveness both for upholding service user experiences and commissioner value for money.

Consequently, the lower axis of Figure 3.5 conveys the limited extent of Work Programme design features geared to the assurance of service quality and access. The programme marker sits at the provider-directed, light touch end of the service regulation spectrum. This contrasts with previous (New Deals) and parallel (Work Choice) programmes – discussed in Chapter 8 – which enforce a higher degree of

⁶ Author calculation based on publicly available data accessed through DWP Stat-Xplore.

clarity over contractual service minima and which are therefore more congruent with state-directed standards and more robust quality assurance.

3.9.3 The particularities of the quasi-market configuration embodied by the Work Programme: production dimension

Turning to the production dimension, on the first component of market access the Work Programme was procured through a ‘Prime Contractor’ model which limits the direct contractual relationship between DWP to a relatively small number of large top tier ‘Prime’ providers who are then responsible for managing supply chains of provision (Armstrong et al., 2010; DWP, 2010a). The programme is structured geographically into 18 large regional CPAs with two or three Primes operating across each CPA.

Under the Work Programme particular forms and sizes of provider have a stronghold over provision as Prime contractors, largely due to financial stipulations set out by the DWP during commissioning. There were tight requirements for bidding as the DWP sought to ensure that Prime contracts were held by organisations capable of financing upfront investment in services and shouldering the financial pressures of back-ended outcomes payments. The DWP stipulated that potential Work Programme Prime organisations had to demonstrate an annual turnover of at least £20 million and be accepted onto the Framework of Employment Related Support Services (ERSS) (Wiggan, 2015a). This created a pool of large multinational ‘preferred suppliers’ exclusively entitled to bid for Work Programme contracts (Gash et al., 2013) and resulted in many organisations being unable to compete. Prime providers were required to include supply chains in their bids – that is a set of allied (usually smaller and more specialist) subcontractors – whom the Prime manages to deliver the contract (Lane et al., 2013). However, there is no stipulation for the use of subcontractors post-tender, and much discussion in the early years of delivery centred on whether smaller – particularly voluntary sector – organisations had been used as ‘bid candy’ when they subsequently failed to see the referral volumes that their Primes may have promised pre-bidding (Damm, 2014; Rees et al., 2013a). Some stakeholders have raised concerns about the creation of “mono-cultures” or “hyper-primes” with apprehension about the potential implications for service users from the squeezing out of specialist provision (Fuentes et al., 2014, p. 80; see also Damm, 2012).

The Work Programme is therefore positioned close to the ‘closed’ end of the market access dimension (upper portion of Figure 3.6) since tendering requirements significantly reduced competition for contracts and situated large existing private sector providers (and those on the ERSS framework) in a dominant position to tender for any future contracts that might become available as a result of market exit. Furthermore, assessment from the OECD (2014) suggests that ongoing competition between providers is unlikely to have been a major driver of performance.

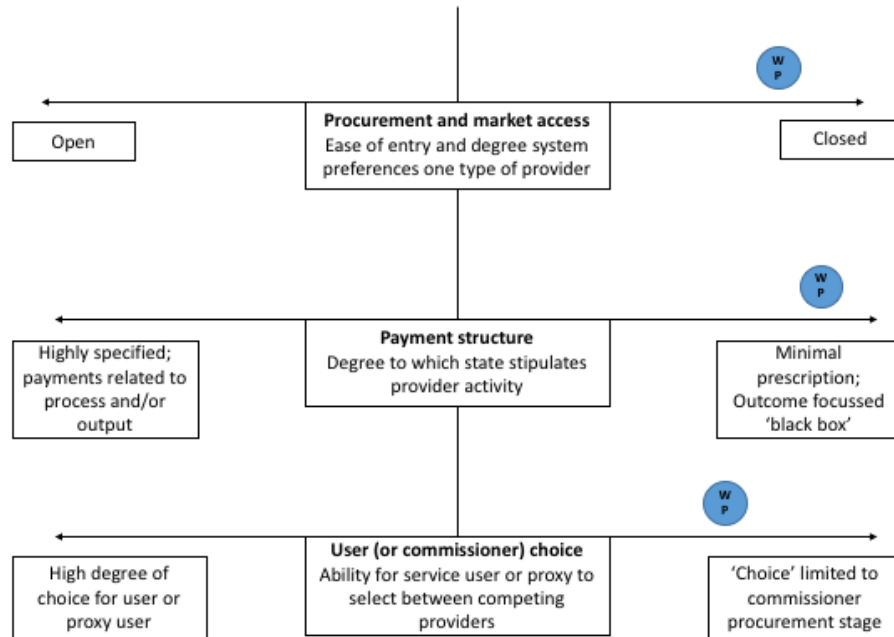


Figure 3.6 Visualising the production dimension for the Work Programme quasi-market

The Work Programme embraces fully the principle of payment-by-results and is – in both international and historic comparative perspective – a pioneer and radical experimenter in the heavy extent to which payment to providers is predicated on the achievement of outcomes. Its use of an extended period over which payments are made for ‘sustained’ employment is particularly novel and untested. The Work Programme fully embodies the ‘black box’ delivery model whereby Primes have almost complete discretion over the nature and extent of their intervention and where provider payments are based almost entirely on job outcome results rather than inputs (Rees et al., 2014). The extreme outcomes-orientation of funding can be seen by comparing the payment configuration to the immediate predecessor – Flexible New Deal – where the ratio between the initial joining fee, a successful transition into work and a sustained job outcome was roughly 40:30:30 (Vegeris et al., 2011, p. 13). In the Work Programme the ratio is in contrast closer to 10:25:65 (authors’ calculations cited in Rees et al., 2014) although it varies somewhat across the nine claimant groups. Importantly, the fixed attachment fee reduces to nil across the early years of the programme so since 2014 the Work Programme has unusually been operating under ‘pure’ payment-by-results: no portion of provider payment is guaranteed, even for participants with complex and/or multiple barriers to employment. Correspondingly, Work Programme is positioned at the extreme right-hand position on this payment structure axis.

On the final dimension relating to the function of choice, it is notable that user choice within the market *between* providers has been almost non-existent as a feature in UK welfare-to-work provision (Wiggan, 2015a). The majority of Work Programme participants are ‘mandated’ to engage as a condition for their ongoing receipt of unemployment benefits, albeit with a significant minority of people with long term health conditions afforded the option of voluntary participation. Participants are randomly allocated, without choice, to a Prime provider (and to clarify, this random

allocation is not part of a Randomised Control Trial: DWP commissioned *no* quantitative impact evaluation of the Work Programme). User choice/voice is exerting very limited, if any, influence on the shape of service provision such that Wiggan (2015a) astutely clarifies that the ‘customer’ in the Work Programme market is the DWP as opposed to programme participants. Though the Department seeks to animate ongoing competition by using ‘market share shift’ to reallocate a portion of the caseload to better performing providers, an OECD review suggests that this has in practice been a weak tool for instigating competition and ‘choosing’ efficiency (OECD, 2014). In the Work Programme ‘choice’ is thus very much limited to the selection of winning Prime providers made by civil servants at the procurement stage, placing Work Programme towards the far right of this dimension in the lower axis of Figure 3.6.

3.10 Work Programme as a private power market and anticipated challenges for this quasi-market form

Against these dimensions of quasi-market variation the Work Programme clearly sits as the embodiment of the provider-directed market. As discussed in the preceding conceptual chapter, for each axis in Figure 3.6 a position closer to the extreme right-hand side on the production dimension reflects a stronger provider position, be that through minimal competition (upper axis), minimal state control (middle axis), or minimal responsiveness to user preferences (lower axis).

Since the Work Programme is collectively financed (as shown on the upper axis of the allocation dimension in Figure 3.5) it may initially be expected that the scheme would tally with Gingrich’s Pork Barrel market. Figure 2.2 in the preceding chapter identifies this form at the intersection of producer dominance and collective financing. However, the development of the previous chapter’s more nuanced analytical framework facilitates a more subtle reading.

Although the Work Programme is collectively financed the quasi-market functions in a situation of strong financial constraints. Tighter than expected cost-pressures are exerted on Primes due to a combination of unrealistically high performance assumptions and, as a consequence, highly challenging average unit costs given the extent of outcomes payments, a more difficult than expected macroeconomic environment which affected job outcome (and hence payment) levels, the strong prevalence – and apparent success of – discounting practices at the bidding stage and lower than expected caseloads within some payment groups (Inclusion, 2011; Rees et al., 2013b).

In short, margins for the overwhelmingly large private sector Primes (39/42 contracts) were squeezed from the start: Primes are strongly motivated to retain these margins and providers’ behavioural responses to drive down the cost of provision were acute from the outset. These financial pressures are coupled with weak protections for service access and quality (lower axis of Figure 3.5) meaning that the Work

Programme appears closer to a *private power market*, where “tighter fiscal constraints push producers to pursue profits through user fees [not possible in the Work Programme] or reduced *quality*” (Gingrich, 2011, p. 17 emphasis added).

Under producer-driven markets the firms which receive public funding are the clear winners. There is some ambiguity as to the implications of private power markets for the state. In crude terms Gingrich suggests that the state ‘wins’ as costs are cut. Importantly, however, cost cutting may be pursued in ways which work against other policy objectives – including value-for-money or longer-term savings – and, indeed, against the preferences of service users. Individual service users are particularly vulnerable under this form of quasi-market as the quality and performance of services may decline in response to cost-cutting by providers and participants will have little scope through which to exert their preference for quality through either market (exit) or democratic (voice) means.

The particular quasi-market variety of the Work Programme as a private power market then sets up a specific constellation of risks. Hart et al. (1997) contend that where there are incomplete contracts (which is almost inevitably the case in complex public services) and producers have great control then (private) providers demonstrate an ability to pursue cost-cutting innovations (i.e. improve crude efficiency) but that they are also more likely to do so at the expense of quality. The lens of Transaction Cost Economics (Coase, 1937; Williamson, 1975) also points to such risks since providing an agent with incentives to achieve a specific outcome is likely to result in the shirking of other — mandated but unrewarded — outcomes such as quality (Holmstrom and Milgrom, 1991; Hill, 2013). As a result, such market forms “promise innovation but face the risk of rent-seeking and uncontrolled cost-cutting at the expense of efficient or high-quality production” (Gingrich, 2011, p. 18).

In the field of welfare-to-work a considerable body of international evidence tells of such risks in practice (Struyven and Steurs, 2005; Bredgaard and Larsen, 2007; Finn, 2010a, 2010b; Considine et al., 2011; de Graaf and Sirovátka, 2012). The particular cost-cutting endeavours of outsourced providers even has its own nomenclature – ‘creaming and parking’ – since a perennial problem within marketised frameworks is the difficulty in incentivising providers to support all claimants effectively given their differing support costs and varied likelihoods of realising employment outcomes. In the case of the Work Programme specifically, an ample body of evidence (Newton et al., 2012; PAC, 2012, 2013; WPSC, 2011, 2013), including from this author (Rees et al., 2014; Carter and Whitworth, 2015), points strongly to endemic problems of creaming and parking designed into the scheme.

3.11 Work Programme design as a moderator of gaming?

Within public services dominated by marketised accountabilities – such as the Work Programme – it is expected that opportunism or ‘gaming’ will be a particular design challenge, and the DWP equally expected as such. The expanded use of outcomes-

based contracts in public services has intensified concerns about their potential unintended effects (Koning and Heinrich, 2013). Studies which have focused on provider behaviour in quasi-marketised employment support systems have identified this behaviour as taking four distinct forms: ‘cherry picking’; ‘creaming’; ‘parking’; and ‘churning’ (Carter, Forthcoming; Finn, 2000, 2009, 2010a; Koning and Heinrich, 2013; Peck and Theodore, 2000; van Berkel and Borghi, 2007; van Berkel and van der Aa, 2005). These processes can be summarised as follows:

- ‘Cherry picking’ refers to the recruitment of easier to support individuals from within a wider pool of those potentially eligible i.e. behaving selectively pre-referral in situations where there is a rationing of programme places (this is also referred to as ‘cream skimming’);
- By the process of ‘creaming’ the literature refers to providers prioritising participants who are closer to the labour market (subsequent to programme entry) as well as targeting services on them in the expectation that they are more likely to trigger an outcome payment and that services required to facilitate this will be relatively low-cost;
- ‘Parking’ is the inverse process experienced by those participants who are understood to be ‘harder-to-support’ (Considine et al., 2011). These participants are deemed to be less likely to achieve an outcome payment and/or the support costs to achieve this are understood to be higher and they are therefore de-prioritised, receiving the minimum possible service (Rees et al., 2014);
- ‘Churning’ or ‘cycling’ is a dominant concern in work-first programmes which do not attend to the quality and duration of jobs (Finn, 2009). It refers to the repeated movement of individuals between low-paid, low-skilled jobs and unemployment and may also capture repeated participation on welfare-to-work programmes where this is punctuated by unstable periods in work (Sunley et al., 2006).

The international literature also makes clear that the specificities of programme design and payment structures can play a key role in either facilitating or buttressing against such risks and provider behaviours (Considine, 2000; Considine et al., 2011; Finn, 2009, 2011a, 2012; Struyven and Steurs, 2005; van Berkel and van der Aa, 2005). Crucially, the institutional setting and wider set of accountabilities informing stakeholder behaviours will interact with the potential for unintended practices such as creaming and parking (Koning and Heinrich, 2013). The implication is that through careful contractual specification and regulation it may be possible to overcome, or at least mitigate, these behaviours and ensure that providers act in support of the full suite of programmatic objectives. The challenge then is for policymakers – through design savvy – to configure and implement a suite of steering tools that retain the perceived innovative and efficient impetus of provider-directed quasi-markets whilst ensuring that a complex set of public sector objectives are met.

Beneath the headline objective of moving people who are long-term unemployed into employment, and the rhetorical simplicity of payment-by-results, the Work

Programme has a more subtle set of underpinning ‘critical success factors’ (DWP, 2010a; NAO, 2012). Notably these ambitions are framed as a direct response to perceived weaknesses of preceding welfare-to-work programmes. Specifically, for DWP as the service commissioners the intention in the Work Programme is to support *more* participants into work, *faster*, *for longer* and whilst *reducing gaps in performance outcomes between the easier- and harder-to-help*. Figure 3.7 sets out the DWP’s array of policy objectives and the bold set of design promises aligned to the pursuit of each of these.

Work Programme critical objectives, defined in invitation to tender	How objective relates to criticism of preceding programmes and/or anticipated opportunism in outsourced delivery under payment-by-results	Key programme design feature to address criticism	An "innovative" tool? Or examples available of where used previously?
Move more people into work ("increase off flow rates..." (DWP, 2010a, p. 3)).	Private providers seriously underperformed against contractual targets on preceding schemes (NAO, 2012).	The payment-by-results regime rewards providers for achieving sustainable outcomes and the 'black box' gives them greater freedom to develop more effective services. Work Programme is backed by contractual minimum performance levels. Market share shift is intended to instil ongoing competition between providers and further incentivise high performance. Furthermore, providers are able to work with participants for up to two years - longer than previous schemes (NAO, 2012).	"Payment-by-results has been used in Australia, the Netherlands, the United States, and the United Kingdom, but nowhere has the proportion based on outcomes been as large..." (NAO, 2012, p. 18). Minimum performance levels, the black box, market-share shift and a two year intervention window are all new to the programme.
Move participants into work sooner ("decrease average time on benefit" (DWP, 2010a, p. 3)).	Previous schemes paid for processes and activities, which are not necessarily strongly associated with improved employment outcomes (Freud, 2007).	Black box and outcomes-oriented payment structure to incentivise providers to implement efficient 'work-first' interventions.	As above, noting that the 'black box' is a tool which is new to the Programme
Move participants into work for longer ("increase average time in employment ... longer sustained jobs" (DWP, 2010a, p. 3)).	Employment outcomes resulting from previous 'work first' interventions resulted in short, unsustainable job placements, with programme participants quickly returning to claim out-of-work benefits (NAO, 2012).	Payment-by-results regime rewards providers for achieving sustainable outcomes. Providers are paid for sustaining claimants in employment over a longer period (up to two years) (DWP, 2012; NAO, 2012).	New: in previous schemes providers were only paid for up to six months of sustained employment (NAO, 2012).
Narrow gap between employment rates for disadvantaged groups and everyone else (DWP, 2010a, p. 4).	Disadvantaged claimants were overlooked by providers in favour of those easier-to-help (Finn, 2012; NAO, 2012).	Differential pricing within the payment model seeks to reflect the varying level of support that people with different needs require (DWP, 2012; NAO, 2012).	Differential payments have been used within previous schemes, e.g. Employment Zones partially adopted this approach. Nevertheless, differential payment is seen as an innovative tool in recent international perspectives (Finn, 2009).
Decrease numbers of workless households (DWP, 2010a, p. 4).	Unclear, but potentially related to perceived inability of previous schemes to support people who are most disadvantaged and distant from the labour market	No specific design features or incentives are designed to meet this target.	NA

Figure 3.7 Work Programme objectives and associated incentive design instruments

Within the designers' response to the criticisms of previous quasi-marketised schemes there are then three bold design elements which add subtlety to the potentially blunt schematic of 'payment for outcomes' – differential payments; the geographical scale and structure of Prime contracts; and sustainment payments. These can be seen as design specifics which *overlay* the provider-directed market without fundamentally disrupting the private power market formulation. These are tools which aim to steer within the world of provider dominance in order to moderate pressures for creaming, parking and churning so as to facilitate the achievement of the programme's public sector objectives (which are about more than crude efficiency). Here the Work Programme design uses further marketised accountability tools to correct for anticipated provider behaviours which run counter to programmatic objectives. The first of these, and the lever which has been most widely critiqued in select committees and academia is the *differential payment system*.

3.11.1 Work Programme design innovations: Differential payment groups

The Work Programme seeks to mitigate creaming and parking through minimum service guarantees (MSGs) for providers and, particularly, via the use of nine separate payment groups with differential payments across these groups (Figure 3.8; Carter and Whitworth, 2015). As described above, weaknesses in the substance, awareness and enforceability of the MSGs mean that in practice there is a strong reliance on the effectiveness of the payment groups and the differential pricing structure to defend against creaming and parking (Carter and Whitworth, 2015; Finn, 2011b). The intention with this pricing structure is that providers are incentivised to work meaningfully with all unemployed participants irrespective of the level of their support needs/costs because the outcome payments attached to claimants are designed to vary in line with these support needs/costs (Carter and Whitworth, 2015).

Claimant Group number	Description	Time of referral	Basis for referral	Max Payment Year 1 starters (£)	% of Work Programme population
1	Jobseeker's Allowance aged 18 to 24	From 9 months JSA	Mandatory	3,810	20.09
2	Jobseeker's Allowance aged 25 and over	From 12 months JSA	Mandatory	4,395	44.94
3	Jobseeker's Allowance early entrant	From 3 months JSA	Mandatory or voluntary depending on prognosis	6,600	24.65
4	Jobseeker's Allowance ex-Incapacity Benefit	From 3 months JSA	Mandatory	6,600	0.71
5	Employment and Support Allowance volunteer	At any time from WCA	Voluntary	3,700	1.76
6	New Employment and Support Allowance claimant	Mandatory when expected to be fit for work in 3-6 months; voluntary from WCA	Mandatory or voluntary depending on prognosis	6,500	5.92
7	Employment and Support Allowance ex-Incapacity Benefit	Mandatory when expected to be fit for work in 3-6 months; voluntary from WCA	Mandatory or voluntary depending on prognosis	13,720	1.03
8	Incapacity Benefit and Income Support (England only)	At any time	Voluntary	3,285	0.23
9	Jobseeker's Allowance prison leavers	On release	Mandatory	5,500	0.67

Figure 3.8 Work Programme Payment Groups and payment levels. Reproduction of previous work published in Carter and Whitworth, 2015

3.11.2 Work Programme design innovations: Large, regional contracting areas with specified minimum performance levels and market share shift

The next design departure relates to the geographical scale and structure of the provider market. The Work Programme Prime provider model divides Great Britain into 18 CPAs. The CPAs are geographically large: all of Scotland sits as a single CPA, for example. These CPAs represent the scale at which the competitive procurement for Prime providers took place with two or three Prime providers contracted to operate within each CPA, resulting in a total of 40 contracts (4x3 + 14x2) (OECD, 2014). Unlike the differential payment system, the logic associated with the *scale* of contracts is not clearly articulated in public-facing documents. DWP's own pamphlet on the Work Programme (2012) which seeks to explain the design and delivery arrangements and which describes the innovative aspects of the scheme, for example, offers no

discussion as to how or why the geographical scale of contracts was deemed most appropriate. Freud's report suggests that the large geographic extent of these contracts was intended to offer the large financial scale "appropriate to attract major players from around the world" who would have sufficient scale to arrange private finance and tap into investment from the banking community (2007, p. 63). This large regional geography was designed with an eye to attracting capitalised providers in the belief that it is this which will introduce high quality contract management and drive up *aggregate* programme performance, rather than because CPAs are the relevant, internally coherent and meaningful economic geography of Britain. Fewer, larger contracts also minimise DWP's contract and performance management needs and costs.

The scale of the CPAs is nonetheless an important programme facet since it is at this scale that two major design features intended to stimulate high performance operate. The first of these is the use of contractual 'Minimum Performance Levels'. Here DWP published estimates of the non-intervention rate – that is the expected ratio of job outcomes to participant starts – that would be expected for each of the main participant groups if they had not joined the programme. Primes are required to deliver results at least 10 percent higher than the 'non-intervention level' or risk losing their contracts (DWP, 2012). The second innovation is referred to as 'market share shift' and is intended to foster ongoing competition between Primes. Under this approach a higher proportion of new referrals are directed to the better performing Prime in each CPA.

3.11.3 Work Programme design innovations: Sustainment payments for longer-term employment

The third of the major incentive-adjusting design innovations in the Work Programme relates to the emphasis on *sustained* employment outcomes within the payment mechanism. Both in terms of a more stretching initial job outcome payment (which is only made following 3 or 6 months of cumulative employment, *not* job entry) and its heavy payment weighting to lengthy sustainment payments (up to 2-years beyond work entry), the Work Programme has more stretching 'outcomes' than forerunner schemes.

A schematic for the payment structure illustrating the tail of sustainment payments is reproduced in Figure 3.9 whilst Figure 3.10 provides detail on the sustainment payments for each of the payment groups. DWP's intention via these sustainment payments is to incentivise ongoing in-work support and help employment last, a perennial problem in the UK labour market and welfare-to-work context (Oakley, 2015).

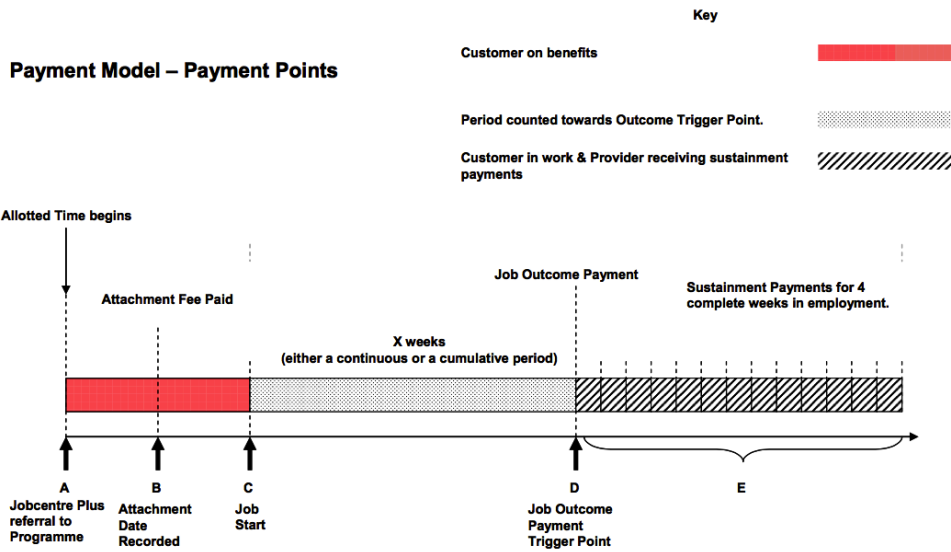


Figure 3.9 Illustration of Work Programme payment model, reproduced from DWP, 2010b, p. 11

Payment Group	Description	Duration of sustained employment required to qualify for job outcome (weeks)	Attach fee, year 1 maximum (£)	Job Outcome Fee, year 1 maximum (£)	Sustainment fee, paid at four week intervals (£)	Maximum available number of sustainment payments	Maximum fee available for year 1 participants achieving sustained employment (£)
1	JSA 18-24	26	400	1,200	170	13	3,810
2	JSA 25+	26	400	1,200	215	13	4,395
3	JSA Early Access	13	400	1,200	250	20	6,600
4	JSA ex-IB	13	400	1,200	250	20	6,600
5	ESA volunteers	13	400	1,000	115	20	3,700
6	New ESA claimants	13	600	1,200	235	20	6,500
7	ESA ex-IB	13	600	3,500	370	26	13,720
8	IB/IS (England only)	13	400	1,000	145	13	3,285
9	JSA prison leavers	13	300	1,200	200	20	5,500

Figure 3.10 The payment trail for the Work Programme: Duration of sustained employment required to trigger job outcome payments and maximum extent of sustainment payments

3.12 Are these bold design promises kept?

The overarching question which animates this thesis is then whether the Work Programme's crafted design structure is sufficient and appropriate in steering the private power market to the achievement of the full gamut of government policy objectives as DWP intended and have defended (as outlined in Figure 3.7). Specifically, the thesis empirically appraises key facets of the success (or otherwise) of the Work Programme in balancing the cost-cutting proclivities within a private

power market via the design savvy of a configured set of additional financial incentives and competitive stimuli.

The empirical work in Chapters 5 – 8 centres on these particular design ‘innovations’ which mark a strident departure from previous UK and international approaches to steering incentives in quasi-marketised systems. Reviews of international welfare-to-work practice suggest that nowhere have these approaches been used in such extreme forms or within such lightly regulated outcomes-based private power markets (Finn, 2012; NAO, 2012).

Chapter 5 assesses critically the differential payment structure and considers the success of this design feature in *narrowing the gap* in job outcomes between disadvantaged groups and everyone else. The research question flowing from this then is: *Has the Work Programme’s differential payment structure calibrated provider incentives within the private power market such that all programme participants have the same likelihood of entering and sustaining paid work, regardless of their characteristics and circumstances?*

Chapter 6 investigates the geography of incentives and performance in the Work Programme, assessing potential tensions between the spatially extensive CPAs and local labour market contexts. The research question here asks *Is the geographical container of regional Contract Package Areas an appropriate spatial scale at which to uphold minimum performance levels and incentivise competition between Prime providers?*

Chapter 7 considers the use of sustainment payments and the structure of financial incentives which extend well beyond an initial period of employment. The research question asks *What are the employment and earning trajectories of Work Programme participants? Has the use of sustainment payments broken the low-pay no-pay cycle?*

Chapter 8 offers a more synoptic reflection on the Work Programme’s private power market. The research questions whether the neglect of participants with health conditions and disabilities is inevitable, asking *Does the variety of quasi-market matter for people with disabilities and health conditions?*

In summary, therefore, the focus of the thesis is both narrower and broader than the Work Programme itself. It is narrower in the sense that the research does not seek to offer a full evaluation of the scheme (although the empirical work represents in various ways contributions that are original, novel and significant compared to existing analysis of the programme) but rather looks to assess – using a range of robust multivariate quantitative analyses – the very specific design promises made within the private power market.

And yet the empirical findings also take on a much broader application. In international perspective the UK is generally understood as a committed marketiser in welfare-to-work (Wiggan, 2015a, p. 119) and the scope and experimentation involved

in Work Programme's marketising efforts position it at the bleeding edge of reform experiences. Relatedly, welfare-to-work can be seen as something of an uncharted pioneer in its extensive use of outsourcing and payment-by-results which is now increasingly being adopted within other policy spheres (Albertson et al., 2018; NAO, 2015), and Work Programme is at the vanguard of this internationally. The conceptual framework, in conjunction with the empirical scrutiny advanced in the thesis, offers policymakers a powerful lens through which to consider the linked conceptual-empirical implications of variation in quasi-market types. The work offers critical reflections on the design fixes available to policymakers internationally, who seek to better balance the competing priorities of government, service users, and providers both within and beyond the field of welfare-to-work services.

4 Data and Methods

4.1 Chapter summary

This chapter outlines the data and methods underpinning the four empirical chapters. The research uses a series of quantitative research methods to investigate the patterning of employment and earning outcomes experienced by Work Programme participants. Each of the four empirical chapters uses a different form of quantitative analysis tailored specifically to respond to the questions described above, which arise from the core Work Programme design features and their anticipated implications for the employment outcomes of programme participants.

The sophisticated quantitative analyses presented throughout the thesis are enabled by the author's rare in-house access to key DWP administrative datasets. These data, rarely accessible to non-DWP researchers, are ideally placed to provide rich new insights to the research's analytical and empirical interests. Across various linkable administrative datasets these data contain a host of individual-level data on socio-demographic characteristics of programme participants, geographical identifiers about where they live, and capture key details on employment outcomes and earnings. The data benefit from being comprehensive (i.e. all benefit recipients and Work Programme participants nationally are included); extensively validated; geocoded; rich in their inclusion of rare key outcomes of interest including official Work Programme outcomes as well as earning profiles; and contain a wide array of relevant explanatory variables.

4.2 Introduction to the research approach

The overarching question which guides this thesis is whether the Work Programme's designed accountability levers are sufficient and appropriate in steering the private power market to achieve the full suite of government policy objectives (as outlined in Figure 3.7). Beneath this overarching question the empirical chapters below consider specific aspects of the Work Programme's design and whether these accountability devices unlock the programmatic ambition: *to support more participants into work, faster, for longer and whilst reducing gaps in performance outcomes between the easier- and harder-to-help*. The policy's critical objectives are themselves framed by, and detectable through, the achievement of employment outcomes. The research scope is therefore centred on the individual and geographical patterning of performance (that is employment and earnings outcomes) across different programme participants, geographical spaces, and time horizons.

Trying to tackle this research intent through a qualitative approach would be problematic. Such qualitative investigatory methods are prone to criticism of the 'picking and choosing' of cases and would not be sufficient in identifying *systematic* patterning in the achievement or non-achievement of employment outcomes for particular programme participants. This study's specific set of research questions lends itself to a research method where the distribution of employment outcomes can be

detected with confidence and which is therefore robust in defence against issues of sampling strategy or limitations on the researcher's access to programme participants (sample size or selection bias). A large-scale quantitative approach is more adept in offering this comprehensive and systematic perspective where suitable data can be found, and in such circumstances enables strong inferences to be drawn from the data as they can be subjected to statistical analysis (although, as noted above, this research is unusual in having access to comprehensive administrative rather than survey data). The quantitative approach adopted here is richer in its ability – given the appropriate data – to detect and quantify large scale patterns as well as to discuss their statistical significance, generalisability and systemic implications against programmatic objectives.

As outlined above in Section 3.12, each of the empirical chapters centre on a particular Work Programme design 'innovation' where the approach adopted by the DWP marks a strident departure from previous UK and international approaches to steering provider behaviour in quasi-marketised systems:

Chapter 5 assesses critically the differential payment structure and considers the success of this design feature in narrowing the gap in job outcomes between disadvantaged and less disadvantaged participants. The chapter asks: *Has the Work Programme's differential payment structure calibrated provider incentives within the private power market such that all programme participants have the same likelihood of entering and sustaining paid work, regardless of their characteristics and circumstances?*

The 'promise' of the differential payment structure is that it will compensate providers for the additional support costs associated with those participants who are 'harder to help' or further from the labour market. The analysis investigates the patterning of job outcomes *subsequent to the application of the differential payment schedule* and uses logistic regression to assess whether variation in employment outcomes is associated with participants' characteristics and previous labour market experiences.

Chapter 6 investigates the geography of incentives and performance in the Work Programme, assessing potential tensions between the spatially extensive CPAs and the much smaller local labour market contexts in which Work Programme participants exist. This chapter asks: *Is the geographical container of regional Contract Package Areas an appropriate spatial scale at which to uphold minimum performance levels and incentivise competition between providers?*

The analysis investigates variation in Work Programme performance between different geographic areas and at different geographic scales. In doing so these analyses respond to the frequent neglect of spatial issues and impacts within welfare-to-work analyses. Where previous studies have investigated spatial aspects of employment programmes they have found that work-first schemes entrench pre-existing spatial inequalities (Sunley et al., 2006) and this would be contrary to Work Programme objectives. The analysis therefore assesses whether there is spatial

variation in programme performance and if there is a relationship between local labour market context (considered at a range of geographic scales) and programme performance. This is investigated using descriptive statistics, mapping, and regression based approaches to deliver a spatialised ‘contextual value added’ performance metric for Prime contractors.

Chapter 7 explores the Work Programme’s use of sustainment payments as a design tool to overcome fractured and low-paid employment experiences. The research asks: *What are the employment and earning trajectories of Work Programme participants? Has the use of sustainment payments broken the low-pay no-pay cycle?*

Employment sustainment is both a key challenge in the UK labour market (Oakley, 2015; Shildrick, 2012; Shildrick et al., 2012) and a key objective – as illustrated through the designed heavy weighting and lengthiness of sustainment payments – for the Work Programme. Despite this policy need and programmatic emphasis on sustainment, however, much current analysis continues to focus crudely on binary indicators of employment outcomes that conceal as much as they reveal. By using innovative sequence analysis and optimal matching of participants’ employment sustainment and earnings trajectories the analysis for the first time in the UK context (and indeed internationally) facilitates an unprecedentedly rich longitudinal investigation of participants’ (un)employment and earnings rhythms which are not visible within the official programme job outcome metric.

Chapter 8 takes a more synoptic approach and looks to innovatively compare the employment and earnings outcomes achieved by matched programme participants in two parallel employment schemes which embody different forms of quasi-market configuration. In doing so Chapter 8 uses considered statistical approaches to take advantage of this natural experiment opportunity. The chapter explores indicatively for the first time in the UK context whether there are any apparent links between alternative quasi-market configurations and the employment and earnings outcomes across matched cohorts. The chapter asks: *Does the variety of quasi-market matter for people with disabilities and health conditions?*

The analysis compares employment and earnings outcomes for people participating in one of two government employment support programmes: Work Programme and Work Choice (where the smaller Work Choice programme adopts a hybridised and softer form of provider-directed quasi-market for its cohort of participants whose participation is based on long-term health conditions and disabilities). The analysis implements a novel quasi-experimental design using propensity score matching to identify variation in outcomes which is attributable to programme participation whilst controlling for variation which might be associated with participant characteristics. The natural experiment is then supplemented with sequence analysis to investigate variation in the employment trajectories of matched participants in the two schemes.

The remainder of this chapter is structured across three core sections. The first portion provides an overview of the data used and offers a brief discussion of the underlying datasets. The second section describes the methods adopted within each of the empirical chapters. Chapters 7 and 8 implement more unusual and more technically demanding analytical approaches (sequence analysis and propensity score matching) and greater attention is therefore paid to outlining these methods rather than dwelling on more widely used regression based approaches. The final section offers reflections on research considerations and positionality.

4.3 Introduction to underlying datasets – an exciting in-house opportunity

Although there are publicly available datasets on the Work Programme (first accessible online through DWP's Tabtool and subsequently through Stat-Xplore) the aggregated and limited cross-tabulations available through these platforms do not produce data with sufficient granularity to support the proposed analysis. De-anonymised individual-level participant information is not available in the public domain nor is it made available to researchers through data archiving centres.

A networking event between the University of Sheffield and members of the DWP labour market analysis directorate in 2012 opened up a constructive dialogue between University researchers and senior civil servants. A series of conversations identified shared areas of research interest. Civil servants with responsibility for this domain were interested in the proposed analysis but were uncomfortable passing any individual-level datasets outside the Department's secure operating system. DWP has a legal and ethical responsibility to make sure the Work and Pensions Longitudinal Study (WPLS) is used appropriately (DWP, 2005). To facilitate data access an arrangement was made for Dr Adam Whitworth (primary supervisor) and Eleanor Carter to access the DWP under 'secondment' arrangements (detailed in a Memorandum of Understanding between the Department and the University). In 2014 the University researchers were inducted as pseudo-staff members, were given DWP laptops and staff passes and were supported to complete compulsory data handling awareness training courses.

This facilitated rare in-house access to a suite of DWP administrative data and enabled embedded in-house working in DWP offices throughout the duration of the PhD research. A 2-hour introduction to the datasets was the maximum extent of pre-analysis support formally provided by the Department. No detailed analysis plans were agreed and whilst there was on-going interest in the research and its emerging findings from DWP policy and analysts colleagues (for example, with internal seminars and slide packs requested and delivered) there was very limited formal ongoing oversight by DWP. There was minimal support from local DWP analysts, despite the complexity of the IT systems, software and multiple complex datasets involved. This meant that considerable work was required by the researcher to access and manipulate the underlying datasets which are stored remotely and are accessed via Base SAS software environment. A new programming language (SAS) had to be learned in order to clean and prepare the data, alongside a raft of more technically complex software interfaces

to enable various databases and servers across the UK to interact in order to access various differently located datasets. The subsequent analysis was conducted in Stata 10 (the most up-to-date version available on DWP systems).

In terms of the administrative data used, the Work and Pensions Longitudinal Study (WPLS) is described by the DWP as its core and comprehensive analytical dataset. WPLS links all Benefit and Programme information held by DWP with employment records from Her Majesty’s Revenue and Customs (HMRC). It provides “a single client view, with history, allowing client group and across client group longitudinal analysis” (McIvor, n.d., p. 3). It has comprehensive coverage of all DWP clients receiving benefits or participating in programmes at 6th April 1999 or starting after that date.

Prior to analysis it was anticipated that the WPLS would serve as the dataset to underpin the research. In practice the WPLS is not a single dataset but rather represents a series of different underlying administrative datasets which can be connected through the use of individual claimant or programme participant identifiers. By connecting across multiple datasets it is possible to investigate longitudinal, spell-based information on individuals’ work and benefit histories.

Figure 4.1 outlines the various contributing datasets used in the research which are then described in further detail below.

Dataset	Data source	Key items
National Benefits Database	DWP	Details of historical and current benefit claims
Work Programme Analytical Database	DWP	Work Programme participants and programme Management Information (e.g. referral and attachment dates, achievement of job outcomes)
Labour Market System	DWP	Socio-Demographic information on individuals who engage with Jobcentre Plus
P45/46 Employment data	HMRC via DWP	Employment spells prior to programme participation
RTI	HMRC via DWP	Month-to-month earning information
Official Labour Market Statistics	NOMIS	Local authority claimant rates
Standardised Index of Multiple Deprivation	Centre for Regional Economic and Social Research	Relative deprivation at small area level

Figure 4.1 Contributing datasets used in the analysis

4.3.1 National Benefits Database (NBD)

This dataset contains information on claims made to key DWP benefits at an individual level. It contains clean, quality assured information on the claim characteristics, including amongst other variables: the start date; end date; and type of claim or benefit received (including all benefit types claimed by those eligible for the Work Programme: Employment Support Allowance; Jobseekers Allowance; Incapacity Benefit; and Income Support). An individual claimant identifier (encrypted National

Insurance Number ‘NINO’) is both unique and linkable to other DWP datasets. The data structure is one row per claim, so an individual who has claimed different types of benefit, at different time points, will have records held in multiple rows. NBD has 100% coverage and captures all claims for benefits made in Great Britain since 1999. It contains over 34 million separate benefit claims. Given that it is based on financial benefit payments it is comprehensive, validated extensively, and exceptionally high quality.

The NBD aggregates data from the underlying DWP payments systems and is not ‘live’. It offers a cleaned snapshot and has lags in completeness. Analysts suggest that data is lagged by around 4 months and this is considered in the timing of the analysis. This data is used to build variables which capture benefit claim durations, for 2-year and 5-year periods prior to Work Programme participation. Generally dates are accurately recorded but there are occasionally small lags in the dates recording spell end. This is corrected for by re-scaling any claim durations which extend beyond the 2-year or 5-year tracking windows.

4.3.2 Work Programme Analytical Database (WPAD)

This dataset contains information on all Work Programme participants and outcomes. Within the DWP it is the best single source of information on those participating in the scheme and is used to produce official programme performance statistics. Key variables include referral date, attachment date, job outcome date and sustainment payment dates.

Core information is sourced from the Provider Referrals and Payments system (PRaP) which is used to exchange information and payments between DWP and Prime providers in a secure environment. Given its core purpose is to validate claims and make payments to Primes WPAD is understandably comprehensive, validated extensively and exceptionally high quality. WPAD now holds information on 1.95 million Work Programme participants.

The data is held as one-row per participant and ORCID is used as a unique anonymised participant identifier which can be used to connect participant details to other datasets. There are a very small number (<1%) duplicate ORCID records (i.e. a single participant with multiple records) and this was corrected for in the cleaning process where only the most complete and/or most recent record was retained. Only a limited number of socio-demographic variables are also included in WPAD (age, sex, and specific identifiers of disadvantage collected prior to programme participation, Appendix 2 provides a full description of available and derived variables).

4.3.3 Labour Market System (LMS) meetings, opportunities, and ‘client’ characteristics

The Labour Market System (LMS) is the key IT system used by the Public Employment Service on the front line in Jobcentre Plus offices across the country in their interactions with people claiming benefits. It is used to manage and input a range

of ‘customer’ information, including: socio-demographic characteristics; the booking and administering of appointments; and referring customers to ‘opportunities,’ that is, external employment support services beyond Jobcentre Plus. The dataset has 100 per cent coverage of individual level data for people who are eligible for and referred to ‘opportunities,’ such as the Work Programme.

Data on socio-demographic characteristics from the Jobcentre Plus LMS contains a number of missing values as advisers do not routinely fill in *all* of the fields during client interviews and claimants are not compelled to disclose all information. Amongst variables of interest to this research (and, more broadly, to the DWP itself) missing information is problematic for variables identifying low qualification levels.

4.3.4 P45/6 Employment data

The P45/6 datasets contain employment information provided by HMRC data systems. This dataset is used to produce variables conveying employment histories (spells and duration) across 2-years and 5-years *prior* to programme entry which previous studies have suggested are key variables for assessing employment support programmes (Bryson et al., 2002). Not only are employment histories understood to be important predictors of employment outcomes in their own right, Bryson et al., (2002, p. 13) suggest that they can also help capture “otherwise unobservable characteristics, such as motivation,” which might also influence programme outcomes.

In most circumstances, employers are obliged to notify HMRC when an employee starts or ends a spell of employment and this data is collected via P45 and P46 forms respectively. This employment spell information is then passed to DWP and can be connected to other administrative data using encrypted NINOs. The dataset holds over 143 million separate linked employment records.

Importantly, the data do not cover all employees and there are a number of documented issues with the comprehensiveness and quality of this data. In terms of coverage, there is no requirement to supply information if the individual is below PAYE tax thresholds and is not going to be claiming new tax credits through the employer. Some employment spells, such as those corresponding to self-employment and where individual earning levels are lower than the PAYE threshold, are therefore not recorded. However, some employers do send in all details regardless and HMRC put them on the system in these circumstances. Despite the limitations in coverage this is by far the most comprehensive dataset on employment and earnings available to analyses of this sort at present, although the slowly growing availability of HMRC’s Real Time earnings data (see below) is gradually superseding this dataset.

There are also issues with unreliable dates: “Where HMRC do not know the date on which an employment spell started, they assign a start date of the 6th April in the year that they become aware of the employment spell” (DWP, 2013b, p. 7). A similar process occurs when there is uncertainty over the date on which an employment spell ended. The recording of end dates is known to be particularly problematic in this

dataset. Details on the actions taken to overcome the limitations of the employment data are outlined in Appendix 1.

At the time of conducting the analysis the time lag between a person starting a job and DWP being notified of the record through HMRC data transfers was not clear to analysts. A best estimate suggested that the majority of records were received within 6-months. As a result, WPAD gives a much more timely, cleaner and more reliable indication of the achievement of Work Programme ‘job outcomes’. This P45/6 dataset is therefore only used to construct employment histories prior to programme participation.

4.3.5 Additional datasets capturing employment and earnings – Real Time Information sample

The author was also granted access to additional HMRC Real Time Information (RTI) data containing granular information on earnings and employment which were newly available to the Labour Market Analysis team. This data was available for a large sub-sample (those referred to the programme between April 2013 and October 2014) of Work Programme and Work Choice participants.

These datasets provide unique academic access to a set of RTI earnings on the month-to-month income from employment for a sample of participants. These are enormously rich data that are gradually growing in availability and importance within UK government analysis and policy making. No published Work Programme analysis has ever utilised these rich RTI earnings data prior to this research project.

The RTI has been collected by HMRC’s new system for collecting Pay as You Earn information⁷ which provides immediate, regular data on employee earnings (Tarr and Finn, 2012). The information is collected directly from employers who since 2014 have been required to provide HMRC with income details immediately after each payment they make to employees. HMRC developed the new RTI collection system in order to respond to the changing employment landscape, particularly given the greater frequency with which employees change jobs and the increased likelihood of people having concurrent jobs (i.e. holding multiple jobs at the same time) (HMRC, 2014).

DWP is a key user of the RTI information as it is a cornerstone of the Universal Credit system. Since DWP is now receiving scans of RTI it has been possible for analysts to match this month-to-month earnings information to a subset of Work Programme and Work Choice participant records, and it is these matched records which the thesis utilises. The dataset contains a record of gross earnings received by each participant within each month. The first month record is the month of programme attachment. The

⁷ Pay-as-you-earn tax (PAYE) is a withholding tax on income payments to employees where income tax, national insurance contributions and other deductions are made from payments to employees and pensioners.

data then provides gross earning information in pounds for each of 15 consecutive months following attachment to the programme. The RTI dataset contains the encrypted NINO for each case, enabling connection to other administrative data.

RTI is one of – if not *the* – most exciting administrative dataset in the UK for both analytical and policy purposes. RTI remains strongly guarded and minimally shared by HMRC, yet this project’s unique academic access highlights its rich potential. The quality and granularity of this data in particular represents a key advance in relation to the patchy p45/46 data which has historically been used. Within DWP it is underused beyond routine analysis of Universal Credit and the analysis offered in Chapters 7 and 8 is at the forefront of innovative exploratory research using this data source.

While this RTI data is understood to be the best quality and most detailed available information in relation to the earnings trajectories of welfare-to-work programme participants ever available for the UK there remain three limitations to the data. Firstly, the RTI data collects information from employers each time they pay an employee as part of payroll arrangements. Self-employed individuals are presently not included within this data collection exercise and so self-employed individuals are necessarily excluded from the earnings analysis. It is possible for Work Programme (and Work Choice) participants to trigger Job Outcome payments for providers by moving into self-employment and hence it is important to note that some programme participants may have triggered job outcomes while being absent from the RTI data. For a participant to be without earnings information within the RTI is therefore not a definitive indicator of a failure to enter employment.

Secondly, the available data provides a monthly earnings figure (£) but does not provide information on the number of hours (or days) worked within the same period. While it is possible to source information on hours worked from the RTI extract, data on ‘hours’ is understood to be considerably less robust than the total earnings data and is generally not used within DWP. This means that a participant who has monthly earnings that look to be equivalent to working full time at national minimum wage may in fact be earning at a considerably higher hourly rate but be in employment for a short number of hours/days. In most cases this is unlikely but it remains possible.

Relatedly, it is not possible to identify within each month at what point employment started, ceased, or whether there was a break in employment *within* the month. Hence, periods of short employment instability – i.e. spells of employment punctuated by unemployment – can only be detected if the period of unemployment extends beyond a full calendar month. This means that shorter periods of unemployment (occurring within a calendar month) will not be reported and therefore the degree of instability detected is likely to be an underestimate of the true extent of ‘revolving door’ employment experienced amongst participants, even if considerable progress is made from previous data on this issue of ‘churn’.

4.3.6 Additional external datasets to capture local economic context

In addition to the administrative datasets, variables relating to the local labour market context and other local contextual information were prepared externally and merged onto the administrative data using a look up table connecting programme participant postcodes to higher geographies (Lower Layer Super Output Area and Local Authority). Previous research suggests that local labour market conditions are an important predictor for the performance of welfare-to-work programmes (Turok and Webster, 1998; Sunley et al., 2006; Peck, 1998; Peck and Theodore, 2000) but local economic indicators are not routinely connected to the administrative data within DWP analyses.

Firstly, a local authority level indicator “Local authority JSA claimant rate” is introduced. Notably, JSA claimant rates have fluctuated considerably across the Work Programme implementation period and some authority areas moved out of the economic downturn more quickly than others. The Institute for Fiscal studies note that 2011 was a particularly challenging year for the UK economy (Emmerson et al., 2012) but macro-economic conditions were more buoyant by 2014/15 (Emmerson et al., 2016). There are also important and long-standing regional and sub-regional variations in labour market conditions (Martin et al., 2016). In order to capture this variation a bespoke local authority JSA rate is calculated for each participant. This is calculated as the mean of the home authority JSA claimant rate for the month of programme attachment and subsequent 11 months. Thus the figure is tailored both spatially (it relates to the home local authority of the programme participant) and temporally (it relates to the economic conditions for the year following their attachment to the programme). Other studies of employment support programmes have similarly used ‘local authority’ as a spatial container for local labour market conditions (Sunley et al., 2006; Davies and Raikes, 2014). Challenges with this definition of labour markets are discussed more fully in Chapter 6.

Detail on the degree of deprivation in a participant’s neighbourhood is also introduced as a potential contextual variable. Although it is notoriously difficult to isolate area effects from compositional effects (Griggs et al., 2008) several studies have shown that living in a deprived neighbourhood may create additional social and economic challenges for residents (Atkinson and Kintrea, 2001; Buck, 2001; Green and Owen, 2006; van Ham et al., 2012) and the analysis therefore seeks to account for this in understanding variations in Work Programme performance. In the UK, each of the four constituent countries measure deprivation using their own distinct index of multiple deprivation (IMD) designed to facilitate targeting of policies within that particular country. Payne and Abel (2012) have undertaken an exercise to produce a Great Britain-wide standardised set of indicators (described fully in Abel et al., 2016). The underlying data was sourced through the CRESR research team at Sheffield Hallam University who kindly provided the underlying figures for Lower Layer Super Output Areas (and in Scotland the equivalent, but slightly smaller, Data Zones).

4.4 Bringing the datasets together

The datasets are combined as illustrated in Figure 4.2 and a full description of the underlying variables including cleaning processes and the treatment of missing data is outlined in the Appendix.

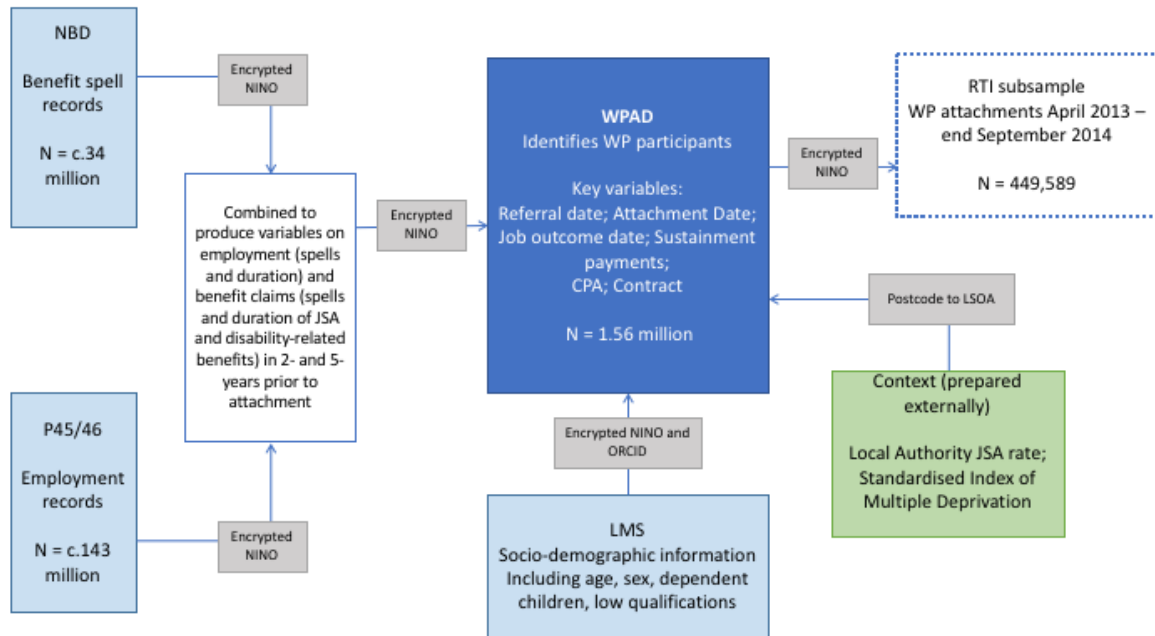


Figure 4.2 Connections across multiple administrative and externally held datasets

The result is that this thesis, unlike any previous independent analysis of the Work Programme, uses the richest possible set of explanatory variables and official programme outcome information sourced from administrative data.

The cleaning and merging of these datasets results in a master file containing 1,563,874 Work Programme participants. This equates to 99.64 percent of the official attachments within the period June 2011 to June 2014 inclusive. The final analysis was conducted in late 2016-17. Using the June 2014 cut-off date for programme referrals ensures that all participants included in the analysis have received the maximum 2-year extent of programme support and have had the full opportunity to achieve job outcomes.

Each of the empirical chapters uses a different quantitative method and draws on a slightly different portion of the administrative data in order to respond appropriately to the research question at hand and maximise the data used in its analyses. This is briefly summarised in Figure 4.3. The following section of the chapter discusses the methods used across the empirical chapters.

Chapter	Data used	N	Method
5	Full Work Programme participant dataset	1,563,874	Binary logistic regression with dependent variable as achievement of Work Programme job outcome
6	Full Work Programme participant dataset but with cases removed where home postcode does not nest within CPA	1,515,473	Visualisation of job outcome rates at different geographies; bivariate regression for local authority performance and binary logistic regression for job outcomes. Note 0.3% of cases from Chapter 5 are removed as most recent home postcode does not nest within allocated contract – i.e. participants have moved out of the contractual provision area
7	Work Programme RTI sample	449,589	Sequence analysis and optimal matching using month-to-month labour market status
8	Work Programme RTI sample and comparable dataset for Work Choice participants	WP 449,589 and WC 28,018	Quasi-experimental impact evaluation using propensity score matching. Additional sequence analysis to explore variation in employment and earnings trajectories across matched Work Choice and Work Programme cohorts

Figure 4.3 Data sources, sample size and methods used in each of the empirical chapters

4.5 Investigating the promise of differential payments: methods used in Chapter 5

Chapter 5 investigates whether the differential payment structure has calibrated provider incentives within the private power market such that all programme participants have the same likelihood of entering and sustaining paid work, regardless of their characteristics and circumstances.

In order to assess whether the Work Programme’s differential payment system has supported its stated objective to overcome the ‘gap’ in employment outcomes between ‘easier-’ and ‘harder-to-help’ participants the analysis outlines a range of characteristics and circumstances which previous research has shown to be associated with employment outcomes (or reduced employment outcomes). The combined administrative datasets provide a range of variables which function as explanatory variables, that is the variables capture characteristics and experiences which would be expected to predict transitions into sustained employment – including age, long-term health conditions and disabilities, caring responsibilities, and previous employment. Previous research (Bryson et al., 2002; Dorsett, 2004; Rees et al., 2014; Carter and Whitworth, 2015) was supplemented by model testing to identify the most parsimonious set of variables (set out in Figure 5.7).

The key outcome variable used in Chapters 5 and 6 is the achievement of the official programme ‘job outcome’, which is defined as six-months of sustained (or cumulative) employment for the JSA 18 – 24 and JSA 25 plus payment groups, and as three months of sustained employment for all other payment groups. Although Work Programme providers are also able to receive sustainment payments (considered more fully in Chapter 7) the job outcome is the most sizeable single payment point and is the core metric around which contractualised performance levels are framed. As argued previously in Carter and Whitworth, 2015, job outcome payments therefore make for a sensible basis on which to assess any gap in employment *outcomes* experienced between participants. The Work Programme administrative dataset contains a record of the job outcome for each participant and in the analysis this is operationalised as a

binary variable where ‘1’ indicates that a Work Programme defined job outcome was achieved (and paid for) and ‘0’ denotes that a job outcome has not been achieved by that particular participant.

The dates of job outcomes are not used, in part because of concerns over disparities between ‘achievement’ and authentication but principally because according to the Work Programme’s own logic some participants may require a longer duration of pre-employment support before being ready to enter work (the justification for the two-year programme support period). The ‘gap’ to be closed is articulated as one of employment *outcomes* and these may indeed be expected to take varied amounts of time to achieve, thus time to job outcomes is not key to the programme logic or these analyses.

Descriptive statistics are used initially to convey the degree to which the successful achievement of job outcomes varies according to key participant characteristics. The analysis then uses multivariate logistic regression to estimate Work Programme participants’ predicted probability of transition into sustained work, given their characteristics. Informed by the employability literature and model testing the explanatory variables are grouped into six overarching themes: benefit history; health and disability characteristics; other individual-level characteristics; household context and local area context.

Logistic regression is the appropriate approach due to the binary nature of the job outcome variable: it is either achieved and paid for by DWP (1) or not (0). Other forms of model specification for binary outcomes are available (e.g. probit) but logistic regression is the most widely used approach and previous research suggests that the functional form of the model does not make a great deal of difference to model performance (Bryson and Kasparova, 2003). In previous work using similar data to predict binary outcomes logit models have proved marginally preferable (Bryson and Kasparova, 2003). Informed by the employability literature and model testing, the set of variables included in the final model are shown in Figure 5.7.

The multivariate binary logistic regression models estimate, for the full available sample of 1,563,874 programme participants, the predicted probability of achieving a job outcome *within the Work Programme*, given their characteristics.

4.6 Investigating the geography of incentives and performance: methods used in Chapter 6

Chapter 6 investigates variation in Work Programme performance between different geographic areas and at different geographic scales. The key outcome variable used in this Chapter is the Work Programme’s own job outcome metric. The chapter questions Work Programme’s contractual obsession with large regional CPAs. In this chapter job outcome rates – that is the proportion of attached programme participants who ultimately achieve a successful job outcome – are explored at a range of geographical scales and for alternate Prime provider contracts. Spatial variation in job outcome rates

is then studied using descriptive statistics and mapping using ArcMap software. Bivariate regression analysis is used to assess the relationship between Local Authority job outcome rates and indicators of local labour market context.

The final stage of the empirical work extends from the individual-level binary logistic regression model advanced in Chapter 5 which is used as the basis for the development of a spatialised ‘contextual value added’ approach to performance. This provides a preliminary estimate of Prime contract performance which is sensitive to context.

4.7 Exploring the employment and earning trajectories of Work Programme participants: methods used in Chapter 7

Given the prevalence of the ‘low-pay no-pay’ cycle experienced by those leaving working- age-benefits and on the lower rungs of the UK employment ladder (Goulden, 2010; Shildrick et al., 2012; Thompson, 2015) it is crucial to investigate Work Programme participants’ transitions into (and, potentially, back out of) employment using a method which is sensitive longitudinally to experiences which do not follow stable, linear employment trajectories.

Complex, and often circular movements through multiple employment statuses and earning levels cannot be easily conveyed through conventional – even if complex multivariate – statistical approaches which typically summarise outcomes at a point in time (for example, the unemployment rate), over a specified period (e.g. amount of time spent in employment in the previous year), or the likely time taken for a particular event to occur (e.g. hazard-rate models for job entry) (Halpin, 2012, 2017; Dorsett and Lucchino, 2014).

In order to capture the full richness of participants’ employment and earnings experiences Chapter 7 adopts an innovative longitudinal classificatory statistical technique – sequence analysis with Optimal Matching and cluster analysis – in order to assess in subtle detail the nature of different types of longitudinal employment dynamics in order to compare individual’s experiences and uncover key patterns.

Relatively underutilised, the method of sequence analysis refers to “the holistic treatment of lifecourse trajectories by calculating similarities or distances between pairs of trajectories, viewed as whole units” (Halpin, 2012, p. 1). This explorative data-driven method was developed initially by biologists seeking to find similar patterns within DNA and is understood to have been first introduced to the social sciences by Abbott and Forrest (1986). Increasingly the approach has been applied across a range of social science contexts and the method has been used particularly fruitfully in labour market research where Optimal Matching has been applied to life courses and career patterns (Halpin and Cban, 1998; Pollock et al., 2002) and to the exploration of youth transitions into the labour market (Brzinsky-Fay, 2007; Dorsett and Lucchino, 2014; McVicar and Anyadike-Danes, 2002; Quintini and Manfredi, 2009; Scherer, 2001).

A clear advantage of the sequence analysis method over more conventional analyses is that it enables one to compare individual unemployment-earning trajectories longitudinally and classify them into one of several transition pathways. By doing so, it unlocks a summary and comparison of the entire longitudinal employment and earning trajectory experienced by each Work Programme participant following their attachment to the programme – including the nature of each month’s labour market status and the ordering of spells within the wider transition pattern –rather than a more typical approach that may focus instead on a significantly informationally reduced binary summary outcomes of those full, rich longitudinal trajectories.

There are three key steps within the implementation of sequence analysis (Quintini and Manfredi, 2009):

- 1) **Construction and coding** of individual trajectories, with decisions made as to how to identify distinct events and/or phases;
- 2) **Distance between trajectories is measured** using Optimal Matching techniques to produce a measure of dissimilarity between each pair of sequences;
- 3) **The most similar trajectories are grouped together** by applying cluster analysis to the derived measures of dissimilarity.

Since sequence analysis through Optimal Matching is a relatively unusual technique in the social sciences a more detailed discussion of each of these steps is outlined below.

4.7.1 Construction and coding of earnings trajectories for the sequence analysis

Unemployment-to-earnings transitions for Work Programme participants are investigated by developing a typology of participant earnings histories, where these earnings trajectories are referred to as ‘sequences’. The ‘sequence’ for each Work Programme participant is a period of 15 consecutive months following their month of attachment to the programme (month 0), where each month is an ‘element’ which can take a certain earnings status or value (Figure 4.4).

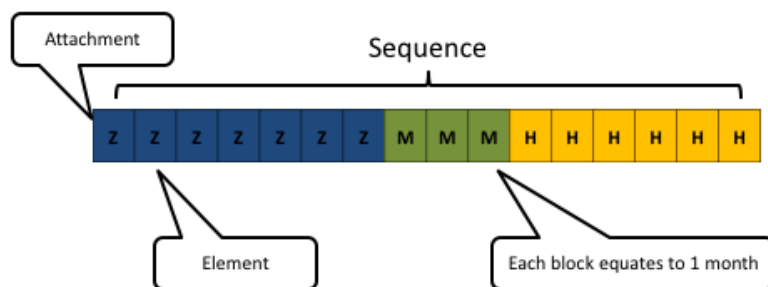


Figure 4.4 Earnings trajectories as a sequence of distinct elements

Previous studies have developed employment trajectories with four (for example, Quintini and Manfredi, 2009) or five (Brzinsky-Fay, 2007) distinct monthly labour market status types. In a review of the application of sequence analysis in the social sciences Abbott and Tsay (2000) note that a number of works have accepted organisationally determined or advised codes. For example, Carpenter (2006 cited in Abbott and Tsay, 2000) uses official job titles and other studies are heavily influenced by ILO classification of employment status to enable international comparisons. Since any coding and separation of earnings levels is to some degree arbitrary the coding adopted in this chapter is informed by boundaries, distinctions and hierarchies that are prioritised by policymakers in DWP.

One key boundary is the ‘Conditionality Earnings threshold’ under Universal Credit which is effectively the earnings level at which recipients will escape the conditionality regime in this revised benefit system. Although this is individually tailored in practice this will commonly be calculated by multiplying the National Minimum Wage (NMW) by the number of hours people are expected to work, normally 35 hours per week (Simmons, 2011). A monthly record of earning above this level (35 hours per week at NMW) is then used to identify a ‘higher earning month’.

Another key policy threshold for DWP is the ‘sixteen hour rule’ where individuals remain eligible for key out of work benefits such as JSA and ESA if they work for fewer than 16 hours a week (DWP, 2011b). This then serves to demarcate a ‘low earning’ element, where individuals have received some income from employment in a month but where earnings fall below the level of 16 hours per week at NMW.

In order for the element categories to be exhaustive and non-overlapping a code for ‘zero earning’ monthly elements and a ‘middle earning’ element are added alongside a classification of elements for individuals who do not have RTI records. Figure 4.5 shows the final resulting codes utilised in the sequence analysis that represent the status options in each monthly element of participants’ trajectories.

Element	Definition
---------	------------

Z	No earnings recorded for this month in RTI
L	'Low earning month' earnings are equivalent to earning less than 16hr per week at NMW
M	'Middle earning month' earnings are equivalent to earning more than 16hr per week at NMW but less than 35hr per week at NMW
H	'Higher earning month' earnings are equivalent to earning more than 35hr per week at NMW
U	No RTI data available

Figure 4.5 Definition of element types used in sequence analysis

For the purpose of comparing trajectories which started at different time points (from April 2013 – September 2014), and where hourly earnings levels for those participants who do enter work are likely to be at or close to NMW, a process of standardisation was necessary. This ensures that those referred later within the timeframe are not assumed to be on a 'higher earning' trajectory through the artefact of slight incremental increases in NMW over time. The boundaries for each 'element' are adjusted in relation to the prevailing rate of NMW for each month.

The element codes are therefore justified through their policy relevance. However, it is pertinent to note that in the only available study investigating the impact of variation in coding, particularly in terms of the level of detail, the results were remarkably stable across different degrees of subtlety (Abbott and Tsay, 2000). An inevitable limitation with this approach is that the allocation of a single status to each month neglects information in terms of other activities that may be undertaken in parallel, for example, individuals who are 'without earnings' in any given month may be undertaking a variety of activities in support of future work transitions. Similarly those with earnings may be undertaking in-work training, which might facilitate future earnings increases, but which again would not be attributed as an activity taking place within the trajectory. However, this is a descriptive rather than explanatory exercise at this stage.

Temporality is a common challenge in applying sequence analysis to social science data, particularly where sequences are of different lengths. In this study, since each trajectory is the same length the sequences are simply aligned by vertical stacking rather than using a 'padding' of unknown earning status to align according calendar month.⁸ The result of the coding is that each participant has a series of 16 (status at month of attachment and 15 consecutive months) month-to-month distinct, non-overlapping earning 'elements' forming a sequence. The next stage is to compare the

⁸ This means that Christmas will effectively fall at different time points within each of the sequences, but overcomes the issue of length differences in the sequences which can result in high dissimilarity values for sequences which might otherwise look similar in composition.

holistic similarity and difference of these longitudinal sequences to one another across all Work Programme participants analysed.

4.7.2 Measuring the distance between earning trajectories in the sequence analysis
 The process of Optimal Matching is used to construct a measure of dissimilarity between each pair of participant sequences (Halpin, 2012). The Optimal Matching algorithm derives a measure of dissimilarity between two sequences as a function of the number and type of operations on the elements that are necessary to transform one sequence into an exact replica of the other. Quintini and Manfredi (2009, p. 17) note that the distance is “roughly speaking, the number of steps one must perform in order to make both sequences equal”. This process is referred to as ‘alignment’ and is achieved when the two sequences ‘read’ the same from left to right – i.e. each of the sequences contain the same elements which occur in the same order. Such an alignment can be achieved by substituting one element type for another, as shown in the left hand portion of Figure 4.6, where to convert sequence B to sequence A, L is substituted with Z; Z is substituted with M and M is replaced with H. The use of substitutions only (referred to as the Hamming distance by Halpin (2012)) retains the timing across sequences and measures dissimilarity as the number of elements that need to be altered.

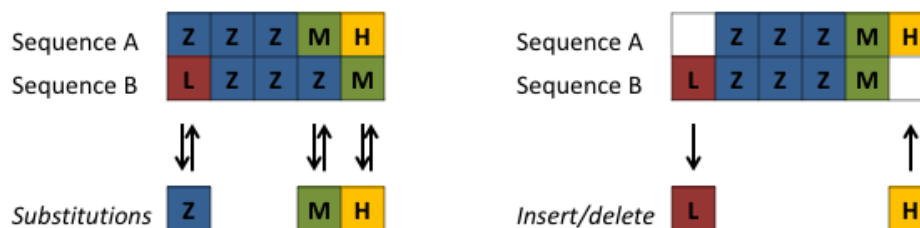


Figure 4.6 Optimal Matching operations

In the right hand portion of Figure 4.6, insertion and deletion functions (referred to as indel) are used to align the common portion of the two sequences such that sequence B is converted to sequence A. The use of indel in the algorithm serves to emphasise the presence of common subsections and the ordering of elements. Importantly, however, “as elements are deleted or inserted, any relationship of contemporaneity across sequences may be broken” (Dorsett and Lucchino, 2012, p. 5). Because this potential ‘time warping’ effect may not be appropriate when the exact timing of turning points or transitions is of key interest, some studies, for example (Dorsett and Lucchino, 2014) use only the Hamming distance.

Between these two extremes – the use of only substitution or only indel functions – alternative measures of dissimilarity can be constructed using both types of operation in combination by assigning each operation a specified ‘cost’. The relative ‘cost’ applied to different operations will have implications for how each pair of sequences is brought into alignment and costs effectively determine how dissimilarity is defined

in the context under study. The decision regarding the quantification of these costs is left to the researcher and must be theoretically justified (Brzinsky-Fay, 2007). A number of the most recent studies of youth transitions from school to the labour market – which represent the most similar application of Optimal Matching to the present study – opt to weight substitution costs at twice the value of indel costs (Brzinsky-Fay, 2007; Quintini and Manfredi, 2009).

Setting the cost of substitution to twice that of indel can be seen as something of a ‘default’ and this convention (in Quintini and Manfredi’s (2009) terms) makes intuitive sense since a substitution in operational terms can be understood as a combination of one deletion and one insertion so that both sequences have the same activity in the corresponding element cell. This cost framework requires no a-priori economic assumptions (Quintini and Manfredi, 2009) and does not preference indels over substitutions or vice versa. For the present study, the exact timing of transitions to work is not the focus of interest. Rather, the analysis seeks to identify similar unemployment-to-earnings experiences and so indels should be enabled. This study adopts the conventional 1:2 weighting approach. Where studies operating with a 1:2 indel-to-substitution ratio have tested different variants of substitution and indel costs for sensitivity (Brzinsky-Fay, 2007), results have showed only slight differences.

4.7.2.1 Selecting socioeconomic proximity costs

In defining costs it is also possible to define the socioeconomic proximity of different elements through the costs tied to different pairs of substitutions. Brzinsky-Fay et al. (2006) suggest that in general, substitution costs should decline as elements become more similar. This principle is implemented by Anyadike-Danes and McVicar (2005) who consider self-employment to be ‘closer’ to employment in socioeconomic terms than to inactivity and so a substitution of self-employment to employment is given a lower cost to that of a shift from inactivity to employment. Informed by the policy rationale for the Work Programme it is appropriate to design a set of socioeconomic costings to reflect the implicit hierarchy across labour market element types. This cost matrix is designed such that a substitution from nil earnings to low-level earnings is valued at twice that of the exchange between different positive earnings levels.

The socioeconomic cost matrix is shown in Figure 4.7 and offers a reflection of policy preferences: high earning is ‘best’ but any earning is ‘better’ than nil earnings. Moreover, this weighting structure also responds to the suggestion that socioeconomic weightings should reflect the data structure such that less frequent events (e.g. transitions from non-earning to earning elements) receive higher cost values than more frequent events (Brzinsky-Fay et al., 2006).

	Z	L	M	H	U
Z		2	3	4	.
L	2		1	2	.
M	3	1		1	.
H	4	2	1		.
U	

Figure 4.7 Weighting socioeconomic proximity. The ‘distance’ from no earnings (Z) to any earning level is given a higher cost than transitions between the earning states

Although the weighting structure adopted is well justified it is also reassuring to note that where alternate cost structures have been investigated “any reasonable set of costs” seems able to distinguish between distinct trajectory types “quite simply because these distinctions are there in the data” (Anyadike-Danes and McVicar, 2005, p. 515).

4.7.2.2 Running the Optimal Matching algorithm

With the costs set the Optimal Matching algorithm then serves to align the sequences, count the number of indels and substitutions required, weight the operations in terms of their respective costs, and add them up – “which heuristically defines the Levenshtein distance” (Brzinsky-Fay et al., 2006, p. 449).

The final challenge is to resolve the process of alignment to arrive at a single value for the distance or dissimilarity between two sequences, when in practice there are several possible ways of bringing alignment. The conventional approach is to select the alignment with the minimum distance between the two sequences and this is found via the Needleman-Wunsch algorithm (Brzinsky-Fay et al., 2006). To conduct Optimal Matching across the earning trajectories the Stata command ‘sqom full’ within the SQ-Ados (Brzinsky-Fay et al., 2006) is used. This command requests that every possible comparison between all participant sequences be calculated, identifies the minimum distance and produces a distance matrix on which cluster analysis can then be conducted.

Running the Optimal Matching ‘full’ command is a highly computationally intensive process and is restricted in terms of the number of sequences that can be compared by Stata's limitation on matrix size (personal correspondence with Halpin). While alternative sequence analysis programmes are available (for example, Halpin, 2017) these rely on C plugins which were not compatible with security systems in place at DWP. This meant that the analysis was restricted to a comparison across a random sample of 20,000 Work Programme participants who feature in the RTI dataset.

4.7.3 Grouping similar trajectories together

The key output from the Optimal Matching process is a distance matrix: a symmetric matrix where both rows (i) and columns (j) represent the individual trajectories in the sample and each cell a_{ij} contains the distance between the sequence of individual i and of individual j. Once this matrix has been produced, cluster analysis is the most common first-stage analytic strategy used to group similar sequences together (Abbott and Tsay, 2000). The objective is to assess whether the sequences can be summarised meaningfully where clusters of sequences are brought together in such a way that the resultant clusters *maximise the similarity* of sequences *within* each cluster and *maximise the differences* or dissimilarities *between* different clusters whilst producing a reasonable number of clusters. Importantly, Everitt et al., (2001, p. 4) note that in general a classification “should largely be judged on its usefulness, rather than in terms of whether it is ‘true’ or ‘false’”.

Selecting the clustering algorithm and the number of clusters requires careful consideration. Several of the most recent applications of sequence analysis in the social sciences have applied Ward’s hierarchical method (Köppe, 2017; Quintini and Manfredi, 2009; Brzinsky-Fay, 2007) which groups sequences in such a way as to minimise the variance within each cluster. Alternative hierarchical clustering methods were tested but Ward’s linkage provided the most viable cluster solutions. This algorithm was run for k target groups ranging from 2 to 15.

Judgment and discretion must be applied by the researcher in the selection of the appropriate number of clusters since whilst statistical analyses can aid cluster selection there is no formal statistical method for identifying definitively the ‘best’ solution. Cluster analysis stopping rules have been developed to support this task and the commonly applied Caliński-Harabasz pseudo-F index (Caliński and Harabasz, 1974) is produced for each solution. The final choice for the number of clusters was guided principally by comparison of statistical indices of fit and also by the desire to explain key patterns in the data with sufficient granularity.

The final step of the analysis involves the production of sequence index plots for each of the clusters in order to convey visually the complexity of the data analysis and longitudinal patterns seen. This representation plots entire individual earning trajectories horizontally, that is each trajectory can be ‘read’ left-to-right as a representation of the monthly activity status experienced by programme participants.

4.8 Comparing employment and earnings outcomes for participants in Work Programme and Work Choice: methods used in Chapter 8

Chapter 8 of the thesis explores whether the configuration of the quasi-market affects programme outcomes via the creative use of quasi-experimental statistical techniques to exploit the partial overlap of cohorts between the Work Programme and a smaller

parallel scheme – Work Choice – which is focussed exclusively on those with health conditions and disabilities.

On paper, the referral processes for Work Programme and Work Choice offer distinct pathways and are targeted at groups with different characteristics and complexities in terms of their likely progression into the labour market. Given that individuals with different characteristics would be expected to respond differently to a single policy intervention, the heterogeneity across the participant populations would rule out a simplistic approach to estimating the relative programme effect by comparing the mean employment outcomes of Work Programme participants with those of Work Choice participants.

In practice ambiguities and tensions in the referral processes for the two programmes mean that participants in each scheme are much more similar in terms of their disability-related employment challenges than when the programmes were initially conceived. Many Work Programme participants were/are facing multiple and complex barriers to employment, including significant disability-related support needs (Newton et al., 2012; Purvis et al., 2013). Many Work Choice participants were in receipt of the ‘mainstream’ JSA benefit rather than the disability-specific ESA, which programme designers had expected to dominate. The official Work Choice evaluation team were asked to introduce an additional research theme to compare the participants referred and support offered to people with disabilities through Work Choice with that offered via the Work Programme and this work identified “a number of situations where disabled people with complex support needs, who might have been suitable for Work Choice support, were being referred to the Work Programme” (Purvis et al., 2013, p. 145).

Those providers delivering interventions within both programmes suggested that some Work Programme participants “were as much in need of intensive [disability-related] support as some of the disabled people being referred to their Work Choice provision” (Purvis et al., 2013, p. 149). And that there were cases where individuals were referred to the Work Programme but who providers considered would be more suitable for Work Choice provision (Purvis et al., 2013). Swift referrals to Work Programme were understood to follow some Work Capability Assessments (particularly for those with mental health conditions) or through limited access to Disability Employment Advisors (DEAs, who are the referral agents for Work Choice) (Purvis et al., 2013).

In parallel, there has been slippage in the degree to which Work Choice participation has been targeted at those with the most profound disability-related employment support needs. Suitable candidates for Work Choice are defined in guidance documents as those who: “experience complex work-related support needs arising primarily from disability” and who “following Module One, expect to be able to work for a minimum of 16 hours per week” (DWP, 2010d, p. 2). The programme evaluation suggests that there is a tension between selecting those with the most complex needs and identifying participants who are likely to enter employment within 6-months. Coupled with caps on the number of referrals DEAs are able to make and the rejection

of some participants who providers felt to be unsuitable the programme has ultimately reduced access for people with the highest disability-related support needs.

The degree to which each of these programmes has experienced stretch and tension in relation to referral routes and eligibility criteria leads to a situation where there is marked overlap in terms of the characteristics expressed by participants attending each of the schemes. The high degree to which Work Programme participants express significant disability-related employment support needs means that in practice there is a group of people who are claiming out of work benefits, have disability-related barriers to entering the labour market, but who could have been referred to either scheme. It is this partial overlap of characteristics, at the intersection of the two programmes that unlocks the potential for well-considered quasi-experimental analysis of the sort presented in Chapter 8.

A further concern when devising an approach to provide a comparison of the ‘effect’ across the two programmes is the potential lack of comparability in terms of a standardised outcome metric. Although both programmes position ‘job outcomes’ as the key unit of success these are defined differently across and within the programmes. Nevertheless, for both Work Choice and Work Programme the overarching objective is to secure stable employment for participants that endures for the long-term and this objective occurs in a wider policy context which aspires to shift the UK to a “higher wage, lower welfare economy” (Oakley, 2015, p. 4). Hence paid employment, which is both more stable and higher paying is the desirable, universal policy outcome. Both the number of months with earnings, (where a greater number of months with recorded earnings is the more preferable outcome); and earnings value, (where greater income levels are seen as superior), serve as appropriate metrics to capture programme ‘effect’ in a comparable manner across the two schemes.

4.8.1 Identifying the counterfactual through propensity score matching

Hence, although there is likely to be an overlap in participant characteristics between the two schemes – in the absence of genuine experimental work – the key challenge is to identify and implement a rigorous non-experimental evaluation method to compare experiences across Work Programme and Work Choice.

In seeking to do so the thesis steps into the realm of the classic ‘evaluation problem’. As with all experiments, there is the issue that in order to truly know the effect of participation in one programme (e.g. Work Choice) compared to the other (Work Programme) we must compare the observed outcome with the outcome that would have happened had that person participated in the alternate programme. However, only one outcome is observed in practice. Bryson et al. (2002, p. 3) concisely summarise this essential difficulty of programme evaluation as ‘one of missing data’: “all approaches to evaluation attempt to provide an estimate of the counterfactual and use this to identify the programme effect”.

When considering the parallel implementation of these two distinct programmes as akin to a quasi-experiment, with Work Choice participation serving as the ‘treatment’ there are two immediate concerns: that of constructing a synthetic ‘counterfactual’ (from a broad Work Programme ‘control’ group) and identifying an appropriate ‘outcome’ metric on which to compare intervention effects. Turning to counterfactual construction, there was no ‘random assignment’ of participants between the two programmes which would have offered the gold standard evaluation methodology. There is no clear, sharp assignment rule distinguishing participants from the ‘control’ group so a regression discontinuity design is not viable. The approach is therefore reliant on the partial – but messy – over of programme participant eligibility and referral.

Propensity Score Matching offers a key method in this regard as it uses information from a pool of non-participants (i.e. those attached to the Work Programme) to identify what would have happened to the treatment group (here Work Choice) in the absence of the intervention. As Heinrich et al. note, “[B]y comparing how outcomes differ for participants relative to observationally similar nonparticipants, it is possible to estimate the effects of the intervention” (Heinrich et al., 2010, p. 3).

This method has been used widely in the evaluation of British active labour market policies (Bryson et al., 2002), has an intuitive appeal arising from the way it mimics random assignment through the construction of a control group post hoc (ibid), and offers key advantages over conventional regression based approaches (discussed further below). Since the method is highly specific and is less commonly used in the social sciences than standard regression analyses the next section briefly summarises its key intentions and operation. The discussion of the method is structured across four key sections: describing the basic mechanics of the approach; outlining assumptions; considering the benefits of matching compared to other approaches and outlining the data requirements and administrative data available to the present study.

4.8.1.1 Basic mechanics of the approach

Propensity score matching begins from the acknowledgement that, in the absence of random assignment, the allocation of participants to treatment is typically not random: people receiving and not-receiving an intervention may differ not only in their treatment status but also in other characteristics that affect both participation and the outcome of interest (Heinrich et al., 2010). To overcome this potential bias, the approach matches individuals exposed to the treatment to individuals not exposed to the intervention but who are otherwise ‘similar’ across a set of modelled characteristics and goes on to estimate the mean intervention impact as the difference in outcomes between the two matched groups (Bryson et al., 2002; Heinrich et al., 2010; Pattie et al., 2015).

Matching each treatment group individual with an untreated individual who is similar in terms of a single variable (e.g. age), or even two variables (e.g. age-sex), is fairly intuitive. But as the definition of ‘similarity’ is expanded to include multiple covariates

the idea of ‘closeness’ becomes more complex both to understand and to capture. Rosenbaum and Rubin (1983) made an important contribution in overcoming this ‘problem of dimensionality’ in demonstrating that matching on a single index reflecting the probability of participation could achieve estimates consistent with simultaneous matching on all covariates. This index is referred to as the propensity score such that “the probability of participation summarizes all the relevant information contained in the X variables” (Heinrich et al., 2010, p. 21).

The estimation of propensity scores then forms the foundation for the remainder of the method, which is described as a four-step process and is discussed more fully in Chapter 8:

- 1) Estimating programme participation and construction of propensity scores;
- 2) Choosing a matching algorithm and performing the match;
- 3) Assessing the performance of the match through the balance of characteristics across the treatment and matched comparison groups and fulfilment of common support assumptions;
- 4) Estimating the impact of the intervention with the matched sample (and calculation of standard errors).

4.8.1.2 Assumptions with Propensity Score Matching

In common with all non-experimental evaluation techniques, propensity score matching depends on a series of assumptions about the nature of the process by which individuals participate in a programme, achieve outcomes, and the data available to the researcher. A series of core assumptions are discussed below in turn alongside a consideration of the plausibility of these positions in relation to the present study, noting that with fewer and more credible assumptions “the more likely it is that estimated effects will approximate real programme effects” (Bryson et al., 2002, p. 7).

Conditional Independence Assumption

“...there is a set X of covariates, observable to the researcher, such that after controlling for these covariates, the potential outcomes are independent of the treatment status...”

Heinrich et al., 2010, p. 16

To be credible in suggesting that any measured difference in outcome between treatment and control groups is attributable to the programme the researcher must be confident that all variables affecting both participation and outcome are observed and available as explanatory variables. Inevitably, as with any study, there are limits to the observed data accessible to this research, where explanatory variables are drawn from administrative datasets and have not been commissioned specifically to facilitate the matching.

In propensity score matching it is assumed that any selection on the basis of unobservable characteristics is trivial and that these ‘unobservables’ are not correlated with outcomes. The research here faces issues in relation to unobservable attitudinal and behavioural characteristics, such as motivation, confidence, and application, which are notoriously difficult to capture. As a voluntary programme, it is likely that motivation and cooperation play a part in the process of referral and attachment to the Work Choice programme and are also likely to be correlated with employment outcomes. Some Work Programme participants matched will similarly have participated voluntarily, but not all. This is not considered to be a prohibitive flaw in the approach however as previous persuasive studies have proceeded on the basis of administrative datasets which similarly lack explanatory variables directly relating to motivation and confidence (e.g. Dorsett, 2004). Additionally, Bryson et al., (2002, p. 5) suggest that the “judicious use of observable characteristics can go some way towards minimising the bias associated with unobservables”. Observable features which are highly correlated with motivation, in this example, the proportion of time spent in employment in the period preceding programme participation, are capable of capturing (and hence controlling for) some of the motivation effect.

Administrator discretion may systematically bias impact estimates. If a service is targeted at those who are closer/further from the labour market then this may bias estimates of programme effects up/down. Though there is a rationing of Work Choice places and DEAs have discretion it is not clear whether referrals are systematically made for ‘easier’ or ‘harder-to-help’ individuals. As a result, it is not possible to conclude whether advisor influence is likely to be contributing systematically to the under- or over- statement of programme effect. Since the direction of bias is not clear, there is no mandate for adjustment.

The variables used in the analysis are outlined in chapter 8. Overall, although the available observed variables may not perfectly and comprehensively account for the differences between the treated and untreated groups any remaining selection on unobservables is considered relatively trivial. The selection bias has been accounted for to such a degree that the matching process is analogous to creating an experimental dataset, such that, after controlling for observed characteristics, treatment assignment is “as good as random” (Heinrich et al., 2010, p. 16).

Common Support or ‘overlap’ condition

“...persons with the same X values have a positive probability of being both participants and non-participants”

Heckman et al., 1999; in Caliendo and Kopeinig, 2005, p. 4

Common support means that the propensity scores of the matched treatment and control groups should fall over the same range. Where they do not then this suggests

that those cases are not well matched. In this study this is not a cause for significant concern as only 53 cases are affected.

The most usual approach for dealing with any common support problem is to identify and exclude those participants who are poorly matched – i.e. who lack ‘common support’ – and then omit them from the estimation of the treatment effect (Bryson et al., 2002). This means that for treated individuals for whom there is no support in the control pool population these individuals are dropped from the analysis. Consequently the estimation of the treatment effect is redefined as “the mean treatment effect for those treated falling within the common support” (Bryson et al., 2002, p. 12). The processes of dropping cases to meet the common support assumption can result in the loss of a substantial portion of the treated population (Bryson et al., 2002). In situations where a sizeable proportion of the treatment group is dropped this may have implications for the policy relevance of results. In these analyses only 53 cases lacked common support and these cases were dropped from the analysis. This is less than 0.01% of Work Choice cases and hence is not a cause for concern.

Stable Unit Treatment Value Assumption

As with most other impact evaluation methods, propensity score matching assumes that there is no “interference between units” (Rubin, 1986, p. 961) and this is identified as the “stable unit treatment value assumption” by Rubin (1980, p. 591). This means the method assumes that the impact of the programme on one participant does not depend on who else, or on how many others, are in the programme or what treatment the other participants receive (Bryson et al., 2002; Rubin, 1986).

As with most other impact evaluation methods, propensity score matching assumes that there is no “interference between units” (Rubin, 1986, p. 961) and this is identified as the “stable unit treatment value assumption” by Rubin (1980, p. 591). This means the method assumes that the impact of the programme on one participant does not depend on who else, or on how many others, are in the programme or what treatment the other participants receive (Bryson et al., 2002; Rubin, 1986). This assumption seems reasonable in the context of the current analysis since programme designers and those overseeing policy implementation equally make the same assumption. Within both Work Choice and Work Programme prices and targets have not been flexed in response to changes in referral volumes or with adjustments to the profile of participants over time. The entire welfare-to-work landscape in the UK operates on the assumption that interventions are scalable and that there is fidelity of service regardless of cohort size and composition.

Relatedly, the method ignores the impact that the intervention may have on the outcomes and behaviour of *non*-participants. Here again the assumption mirrors that of all UK activating labour market interventions which tend to operate on legacy assumptions flowing from Layard’s work (1991) which suggests that substitution effects do not really exist (Bryson et al., 2002, p. 5). Because this assumption pervades the wider policy landscape it seems reasonable to also maintain the assumption here.

4.8.2 Benefits of Propensity Score Matching compared to other available methods
Matching offers important advantages when compared to conventional regression-based approaches. As noted, the common support problem ensures that programme effects are not extrapolated beyond the common support area. An additional advantage of matching over conventional regression is that matching is non-parametric and so is not dependent on any functional form assumptions or specifications.

One additional impact evaluation method would be to combine matching with difference-in-difference (a before-to-after trend analysis of outcomes) but, as noted by Dorsett (2004), this requires observations both before and after participation and is not possible in this analysis because of the structure of the available data. In any case, participants within both the Work Programme and Work Choice principally become eligible for the schemes through their experience of long-term unemployment so the ‘trend’ in employment and earning outcomes before programme enrolment is both similar and incorporated into the matching covariates. Other potential methods, such as those using an instrumental variable to introduce an element of randomness, were also considered. A fundamental limitation with the instrumental variables approach is that it requires at least one independent variable that determines programme participation but which is wholly independent of programme outcomes. In this study, as in many others (Bryson et al., 2002), it has not been possible to isolate a variable capable of performing this role.

4.8.3 Propensity Score Matching in Summary

The limitations associated with alternative methods and the specific features of the interventions under consideration here, particularly in terms of the fuzzy and ambiguous referral routes, leave propensity score matching as the most appropriate method for investigating the relative employment effects of Work Choice and Work Programme. Although there are a number of assumptions bound up with propensity score matching, the maintenance of these is both viable and evidenced empirically in Chapter 8.

4.9 Research considerations and positionality

Whilst the chapter thus far has focused on the more technical aspects of the research’s data and methods, this final section discusses wider considerations around the broader motivations and positionality of the researcher within the study.

In selecting the research design, Creswell (2009) and others (Creswell and Plano Clark, 2011; deMarras and Lapan, 2004; Shannon-Baker, 2016) suggest that although philosophical and epistemological ideas remain largely hidden in research (Slife and Williams, 1995) they still influence the practice of research and need to be identified. It can be challenging for an early career researcher to identify and espouse their own philosophical ‘worldview,’ particularly when research handbooks and methods text books tend to simplify and categorise these into a series of discrete text boxes: *what if I identify with multiple statements across these? Do I even have a coherent worldview?*

This author's views are shaped by frustrations over remnants of the "paradigm wars" between quantitative and qualitative methodologies (Schutz et al., 2004) and galvanised by those calling for a more democratic approach (Ricucci, 2010) who promote our ability to embrace, utilise, and learn from the multiple research traditions active in human geography and social policy. Nevertheless, the underlying epistemology that this research most comfortably aligns with is the form of postpositivism described by Ricucci, (2010) which suggests that all observation is imperfect or fallible and subject to inaccuracies: "research can strive to reach reality, but that goal can never be realized. Postpositivists accept the existence of error and consider research results or findings as probable until falsified" (Ricucci, 2010, p. 305). The research approach is necessarily informed by the author's background and training in quantitative research methods, particularly the use of regression-based analysis.

The study aspires not only to deliver conceptual and empirical advances in our scholarly understanding of UK and wider international welfare-to-work but also to be relevant for the future decision-making of policymakers.

Guides for those striving to produce research used by policymakers emphasise the need for researchers to produce 'high-quality' and 'trustworthy' studies (deMarrais and Lapan, 2004). What is perhaps less well discussed is the balance that applied academic research strikes between extending critical analysis *beyond* the insights that will be achieved through in-house or commissioned research, (which has been specified by the government department from which the policy originates and which rarely, if ever, questions the underlying logics, justifications or approaches embodied within any particular programme or policy) whilst anchoring the research within a frame that will be both intellectually-accessible and politically-palatable (i.e. neutral) to policymakers.

The desire to walk this tightrope of pragmatism however flows more from the author's ambitions and preoccupations as a researcher rather than from any pressure or scrutiny applied by the DWP to the analysis or reporting of findings. Once inside DWP with a completed set of data awareness training there was no formal reporting in terms of detailed analysis plans and no oversight or management by DWP of the research. Findings were shared with policymakers on an ad-hoc basis. Discussions with DWP colleagues on the research findings were driven principally through conversations between the researcher and policy leads responsible for the design of future employment support programmes, particularly in relation to the payment mechanism.

This flexibility of data access and high degree of researcher discretion sits in contrast to preliminary expectations which would have suggested that a relatively inexperienced researcher, situated within a large central government department which has well-known collective nervousness around data sharing and external scrutiny, may be likely to experience (and potentially succumb to) pressures to report findings which

are more lenient toward the policy ambition and programme design. This was not the experience and hence the responsibility for research framing, implementation of data preparation, analytical approach and interpretation (and any limitations or shortcomings) reside entirely with the researcher. The ability to publish findings was secured through the Memorandum of Understanding between the DWP and the University of Sheffield and there is no pre-specified approach for securing sign-off on academic publication by the DWP.

Having outlined the data, methodological and wider research foundations of the project, the thesis turns now to the discussion of its empirical findings, beginning in Chapter 5 with interrogation of the ability of Work Programme's differential payment system to support its stated desire to narrow performance gaps between 'easier to help' and 'harder to help' participants.

5 Equality through inequality: pursuing differentiated universalism through the Work Programme payment structure

5.1 Chapter summary

The Work Programme's configuration as an exemplar of the private power market leaves particular vulnerabilities and risks in relation to the degree to which politically and administratively specified programme objectives are subaltern to provider roles and interests. A perennial fear in fully outsourced 'black box' welfare-to-work models with payments based overwhelmingly on job outcome results is providers' incentives to 'cream' and 'park' claimants. The DWP has sought to mitigate such provider behaviours through Work Programme design, particularly via the use of claimant groups and differential pricing. The empirical work across this chapter considers whether this design approach has succeeded.

This chapter considers the implications of the Work Programme quasi-market design particularities for the degree to which the programme is capable of delivering against its stated ambition to "narrow the gap" between employment rates for "disadvantaged groups and everyone else" (DWP, 2010a, p. 4). To do so the analysis investigates the patterning of job outcomes *subsequent to the application of the differential payment schedule* and uses logistic regression to assess whether variation in employment outcomes is associated with participants' characteristics and previous labour market experiences.

5.2 Reconciling Work Programme objectives: the four E's of the differentiated universalism tightrope

The Work Programme quasi-market has been constructed in such a way that provider flexibilities and discretion are dominant. As discussed in Chapter 3, guided by the influence of Freud (2007) the Work Programme operates as a private power market, key elements of which include: direct contracting with few, large Prime providers over geographically extensive contract areas; strong provider discretion through 'black box' commissioning with no mandatory service components; steering through marketised forms with pricing in the form of 'payment-by-results' as the principal means through which to incentivise behaviour; and light-touch protections for programme participants where minimum service guarantees (MSGs) are (often vaguely) defined by providers themselves and there is minimal scrutiny of service activity or provision quality.

This private power market configuration has been pursued in DWP's belief that such an approach will effectively unlock both service innovation and cost efficiencies, leading to better performance alongside enhanced value-for-money. Yet this provider-directed formulation leaves open significant vulnerabilities and risks in relation to the

degree to which politically and administratively specified programme imperatives are subaltern to the providers' powerful independent roles and pursuit of profit. In such contexts where there are incomplete contracts and service producers have more control, Hart et al., (1997) argue that private providers may demonstrate an ability to pursue cost-cutting (and hence profit enhancing) innovations but may do so without necessarily responding to the desires of the state or users of services. As discussed below, creaming and parking become endemic risks and concerns.

It is within this quasi-market context that policymakers have outlined the central objectives for the Work Programme. The Department's Outline Business Case, as reported in NAO (2014), sets out the key aims:

“The Work Programme aims to increase employment, and reduce the time that people spend on benefits. In particular it aims to improve support for those who are harder-to-help. The Department expected to achieve these aims for a lower cost per referral than previous welfare-to-work initiatives.”

NAO, 2014

Chapter 3 (specifically Figure 3.7) sets out these specific policy objectives and the bold set of design promises aligned to the pursuit of each of these. The programmatic aims can therefore be considered as ‘the four Es’: efficiency; efficacy; economy; and equality.

Efficiency The first intention for the programme is to “move more participants into work” (DWP, 2010a; NAO, 2014, p. 2). This ‘efficiency’ argument is framed in terms of commissioner spending, where the ambition is to purchase a greater number of employment outcomes at the same (or lower) cost than previous schemes. This ambition for efficiency can be captured through the outcome to cost ratio. Good programme delivery against this measure would equate to a reduced average unit cost, i.e. in the present the spend on an individual is reduced.

Efficacy The second ambition can be framed in terms of efficacy since the programme commits to move participants into *sustained* employment, explicitly pursuing an increase in the average time in employment for participants, compared to previous programmes (DWP, 2010a). Here it is helpful to consider programme efficacy as a rectangular space where the horizontal axis is the duration of employment and the vertical axis is the proportion of participants who have entered employment (and their hours per week in that employment). The intention then is to drive up the overall (shaded) area of the rectangle through improved job outcome rates, the speed with which participants enter employment and the degree to which this is then sustained in the medium term.

Economy A key rationale for the scaling of employment support provision in the Work Programme is that HM Treasury will make net financial gains from Work Programme. This is expected to arise primarily from a direct reduction in DWP

spending on out-of-work benefits and through increased income tax and national insurance contributions for HMRC via sustained employment transitions. Through this lens ‘good’ programme performance relates to a reduction in future benefit spending and net savings to the Treasury.

Equality The Work Programme sets out an explicit intention to ensure that those participants who are ‘harder to help’ receive appropriate and effective support. The fourth programmatic aim is for “less ‘parking’ of harder-to-help groups” which is to be assessed through a reduction in the gap between employment rates for disadvantaged groups compared to other participants (NAO, 2014, p. 2)

Although this suite of objectives has been presented uncritically in previous analysis (CESI, 2013; Finn, 2011a; NAO, 2012, 2014) there are multiple important tensions and contradictions across the achievement of these collective objectives (Rees et al., 2014). International evidence suggests that the essential challenge faced by current activation policies is the “tightrope walk” (Heidenreich and Graziano, 2014, p. 3) between cost-effectiveness and equitable, universal support. A particular challenge, and the core analytical focus of this empirical chapter, is therefore the degree to which the Work Programme has been capable of reconciling the tensions wrought through the programme’s ambitions to simultaneously deliver ‘efficiency’ – that is the achievement of the maximum number of employment outcomes for any given spend – and equality – i.e. to “narrow the gap ... in employment rates for disadvantaged groups and everyone else” (DWP, 2010a, p. 4).

The simultaneous achievement of these aims presents a fundamental allocative challenge, particularly given the heterogeneity of support needs within Work Programme’s large and diverse programme cohort. The challenge is to find the ‘sweet spot’ of type, intensity and cost of effectively tailored personalised support across its diverse set of participants. Importantly, the programme’s commitment to the pursuit of equality is a particular form of equality. The Work Programme’s objective is framed explicitly as equality in the space of *outcomes* for participants, rather than equality in the conventional ‘procedural’ sense of standardized ‘equal’ services and supports across all participants. The intent is to secure equal potential for achieving employment for each participant, irrespective of their starting point.

Given the inevitable and substantial heterogeneity in participant characteristics within such a large and by design deliberately diverse cohort of unemployed participants, each individual is likely to require qualitatively and quantitatively distinct supports in the quest for their sustained employment outcomes. Taken with the parallel programmatic commitment to efficiency, the pursuit of outcomes equality cannot be pursued by giving each participant access either to the maximum possible support (which for many participants would in large part be unnecessary) or equal (i.e. identical) supports, given their differing types and intensities of support needs.

Rather, in seeking to reconcile tensions across the efficiency and equality dimensions the concept of ‘differentiated universalism’ enables a clearer unpacking of the policy challenge. The objective, as articulated previously by this author (Rees et al., 2014, p. 226 emphasis in original), is “to treat different claimants *differently* dependent upon their distance to the labour market and barriers to work, in order that all claimants receive the amount and type of support so as to *equalize* opportunities to move into employment”. This policy objective is tautologous to Lister’s concept of ‘differentiated universalism’ (1997). Although Lister’s term was developed as an attempt to combine the core of universal citizenship with a policy of difference (Lister, 1997, 1998; van Ewijk, 2009) here it helpfully captures the simultaneous programme ambitions of working ‘efficiently’ across a highly diverse cohort of unemployed participants and seeking equality of employment outcomes whilst – indeed *through* – responding to difference.

Differentiated universalism therefore acts as the lodestar by which to steer between efficiency and equity. This is “allocative efficiency” in van Berkel and van der Aa’s (2005, p. 338) terms, where tailored lighter touch or more substantial and longer lasting individualized supports respond to participant needs, ambitions and characteristics. This *difference* in support must be calibrated such that each participant is equal in the domain of outcomes – that is, so each has an *equal probability* of achieving sustained employment, the key payable outcome for the programme. It is this *inequality* in service treatment that is required to give *equity* in support, which is then understood as the facilitator of both equality and equity in the space of employment outcomes.

Within the Work Programme’s private power market the task for programme designers then is to construct accountability tools such that they coax providers to the pursuit of differentiated universalism, utilising the flexibilities afforded to providers. Importantly, the objective is for equality ‘up’ – i.e. bringing the employment prospects for those furthest from the labour market up to the level of the most immediately employable members of the cohort – not equality ‘down’, which would clearly run counter to the ambition of programme efficacy. And as discussed in Chapter 3 (above) it is the differentiated payment system of the market accountability lever that is *the* key accountability approach to achieve this, both in Work Programme’s design on paper and, given weaknesses in other accountability dimensions, even more so in practice.

It is worth being explicit here in terms of what this achievement of differentiated universalism would mean in statistical terms since this is the barometer for assessing the degree to which programme design has been capable of responding to and reconciling tensions across the policy objectives in this chapter’s empirical analyses: differentiated universalism requires for each participant that the likelihood of achieving a successful job outcome is brought upwards to the same level, regardless of their initial position in relation to the labour market.

5.3 A troubling pursuit? International experience in relation to the attainment of differentiated universalism in quasi-marketised welfare-to-work systems

With the programmatic goal of differentiated universalism now set it is pertinent to reflect briefly on evidence relating to the ease with which this objective may be accomplished. Although this specific concept has been newly introduced to the welfare-to-work field by the author (Rees et al., 2014), its principles are not new to this policy field and its challenges have been widely discussed in a burgeoning international literature.

Under these particular quasi-marketised configurations it is expected, and increasingly detected, that economically ‘rational’ providers will respond to financial pressures and incentives by ‘creaming’ off easier to serve participants whilst ‘parking’ harder to support individuals. By the process of ‘creaming’ the literature refers to providers ‘skimming off’ participants who are closer to the labour market and targeting services on them in the expectation that they are more likely to trigger an outcome payment (and that services required to facilitate this will be relatively low-cost). Conversely, ‘parking’, is experienced by those participants who are deemed to be unlikely and/or relatively expensive to generate an outcome payment and who are therefore de-prioritised, receiving the minimum possible service (Rees et al., 2014). Processes of creaming and parking are understood to be most likely in situations where profit-motivated providers dominate provision, where there are tight performance targets and cost pressures, where client needs are diverse and where there are light-touch minimum standards and monitoring: all of which are present in the Work Programme (Carter and Whitworth, 2015; Rees et al., 2014).

International evidence highlights the extent of creaming and parking in practice (Bredgaard and Larsen, 2008; de Graaf and Sirovátka, 2012; Finn, 2012; van Berkel et al., 2012b). Assessments within countries at the leading edge of outsourced, quasi-marketised welfare-to-work services each provide evidence where providers have enacted processes of creaming and parking despite designs intended to limit such behaviours (Finn, 2011b, 2011a). The US (Finn, 2011a; Heckman et al., 2002; McConnell et al., 2003), Australia (Dockery and Stromback, 2001; Finn, 2011b; OECD, 2001; Struyven and Steurs, 2005) and the Netherlands (van Berkel and van der Aa, 2005) provide pertinent experience of such practices. In Australia, which has a long pedigree in outsourced employment support provision, the process of parking has proved to be a particular challenge, leading an OECD evaluation to conclude that “few providers appeared to be offering effective services to address the underlying barriers to employment” of the hardest to place (OECD, 2001, p. 19).

The accomplishment of the differentiated universalism puzzle is therefore an ongoing, considerable and potentially intractable policy challenge of international employment support programmes. Work Programme’s combination of open recognition of the challenge, commitment to overcome it, and particular private power market

configuration at such scale and variability of cohort renders it a leading case study in the attempt to square the differentiated universalism circle.

These risks of creaming and parking ‘matter’ both for service users and commissioners because such processes clearly cut against the achievement of differentiated universalism: the achievement of outcomes is not equalised across participants regardless of their characteristics. For commissioners this has important financial implications. As this author has evidenced and argued previously (Carter and Whitworth, 2015), creaming and parking effectively translates to *overpayment* by government for any given employment outcomes secured. Payment models on paper are pegged off averages of full cohort distributions when in practice providers – if they cream and park – are targeting only a self-selected ‘easier’ slice of the participant population who have both increased likelihoods of job starts and/or cheaper support needs, thus simultaneously eroding the achievement of value for money and equality of outcomes. As this author has separately evidenced elsewhere, for service users who are denied meaningful support there are serious psychological implications since unemployment, when coupled with lack of meaningful support or progression and low perceptions of control over future life course, is associated with statistically significant reductions in well-being whilst on programme (Carter and Whitworth, 2017).

In seeking to militate against practices of creaming and parking the international literature makes clear that the specificities of programme design and payment structures can play a key role in either facilitating or buttressing against such provider behaviours (Considine, 2000; Considine et al., 2011; Finn, 2011a, 2012; Struyven and Steurs, 2005; van Berkel and van der Aa, 2005). In outsourced programmes there is a clear trade-off to be made between promoting innovation in the services adopted by providers and the extent of specification and monitoring of processes to assess whether or not services are being provided as desired (Dockery and Stromback, 2001; Finn, 2011a). Clearly the provider discretion at play within the Work Programme’s black box is considerable: within this unconstrained realm where government absconds from specifying services there is considerable scope both for innovation but also for neglectful (and undetectably so) practices. The challenge then for Work Programme is to design away expected provider cost-cutting through negative aspects of creaming and parking in a context where – within its logic of the private power market – the mix of available accountability tools is deliberately weighted heavily towards market levers, providers are granted considerable discretion, and cohorts are highly diverse.

5.4 Why would the Work Programme be any different? Expectations of programme design for the pursuit of differentiated universalism

As discussed in Chapter 3, to facilitate differentiated universalism within this programmatic context Work Programme policy designers have developed what, at least superficially, appears to be a complex and nuanced payment-by-results structure with distinct payment groups and differential payment levels across these. The DWP offers minimum service guarantees and differential pricing as two formal mechanisms

to defend against creaming and parking (Carter and Whitworth, 2015). In practice, weaknesses in the substance, consistency and potential enforceability of providers’ minimum service guarantees render these a relatively weak and unreliable protection (Finn, 2012; NAO, 2014). There is therefore a reliance within the Work Programme on the effectiveness of the market accountability lever via payment groups and the differential pricing structure to defend against creaming and parking and facilitate differentiated universalism. It is necessary to first understand the payment group structure to position this chapter’s later empirical analyses.

The differential payment structure across nine payment groups is shown in Figure 5.1.

Claimant Group number	Description	Max Payment Year 1 starters (£)	% of Work Programme population
1	Jobseeker’s Allowance aged 18 to 24	3,810	20.09
2	Jobseeker’s Allowance aged 25 and over	4,395	44.94
3	Jobseeker’s Allowance early entrant	6,600	24.65
4	Jobseeker’s Allowance ex-Incapacity Benefit	6,600	0.71
5	Employment and Support Allowance volunteer	3,700	1.76
6	New Employment and Support Allowance claimant	6,500	5.92
7	Employment and Support Allowance ex-Incapacity Benefit	13,720	1.03
8	Incapacity Benefit and Income Support (England only)	3,285	0.23
9	Jobseeker’s Allowance prison leavers	5,500	0.67

Figure 5.1 Work Programme Payment Groups and payment levels. Reproduction of previous work published in Carter and Whitworth, 2015

The differential payment model is seen as an innovative marketised steering tool in recent international perspectives (Finn, 2010a), yet the payment groups and differential payment model adopted within the Work Programme are widely considered to be crude and rather simplistic (WPSC, 2013, 2015, 2011). Nonetheless, the approach has been consistently defended by Ministers and senior civil servants. Chris Grayling, then Minister for Employment within the DWP explained, “we needed to find something that was simple to administer that was likely to be reasonably reflective” and that “was simple, easy to understand, where there was no scope for debate and discussion... There will be variations within each [payment] group, that is inevitable but we think as a *broad average* it gives the providers a sensible basis to work with” (Grayling in WPSC, 2011, p. 28 emphasis added). Robert Devereux, then Permanent Secretary for the DWP, augments this sentiment, suggesting that the Work Programme payment groups and differential pricing structure “begins to move us towards trying to reflect some of the average difficulty [of moving participants into sustained employment]. Everything we have done here takes us a really long way forward compared to where we were” (PAC, 2012, p. 26).

Commentators and the academic community have been considerably more sceptical in terms of the likely capacity for the payment scheme to mitigate risks around creaming and parking. It has been repeatedly acknowledged that the current segmentation and payment structure based on benefit type is an overly crude basis on which to proxy distance to the labour market and, as a consequence, on which to calibrate groups and differential payment levels (Lane et al., 2013; WPSC, 2013, 2011). As outlined previously, there is growing and consistent evidence from within DWP's official evaluation (Lane et al., 2013; Meager et al., 2014; Newton et al., 2012), from parliamentary select committees (PAC, 2012, 2013, WPSC, 2013, 2011), and from this author's academic research (Carter and Whitworth, 2015; Rees et al., 2013a, 2014) that the differential payment regime is not holding up empirically in its ambition to guard against creaming and parking processes. The first phase of the DWP commissioned qualitative programme evaluation is as frank as any government commissioned evaluation is ever likely to be: "the available evidence to date suggests that providers are engaging in creaming and parking, *despite* the differential payments regime" (Newton et al., 2012, p. 124, emphasis added).

There is therefore considerable doubt as to whether the Work Programme – through its differential payment model – has been capable of facilitating differentiated universalism and the equalization of job outcome probabilities. Whilst these almost universally qualitative studies have provided strong evidence of the limitations of the payment system in this regard, the present chapter goes further than any previous study in its quantitative exploration of the limitations of the payment system. Previous rich qualitative research is highly valuable in its insights but is inevitably limited in its systematic comprehensiveness and generalizability and – rightly or wrongly – has proven unable to persuade UK policy makers of key programme design issues in this area.

The author's unique academic access to the DWP's data enables the most detailed and comprehensive ever quantitative investigation of the Work Programme's differential payment system using DWP's own authoritative datasets. The full achievement of differentiated universalism can only be detected by considering at the level of individual participants whether the likelihood of achieving a successful job outcome has been equalised, regardless of participant characteristics and contexts. To detect this the analysis requires, for each individual participant, a suitable and standardised indicator of employment outcomes and a suitable set of covariates on which basis we might expect participant distance to sustained outcomes to vary. Both are possible within the DWP data accessible to this research:

- **A Standardised indicator of employment outcome.** The analysis in this chapter uses the programme's official binary 'job outcome' metric. The empirical work draws on a comprehensive linked administrative dataset containing 1,563,874 Work Programme participants who were referred to the programme between June 2011 and June 2014 and who therefore at the time of analysis had each experienced the maximum 2-year extent of Work Programme provision, and who had therefore

had sufficient time in which to achieve an official 'job outcome'. This outcomes data is sourced from the DWP payment system.

- **A rich set of independent variables.** Informed by the employability literature (Bryson and Kasparova, 2003; Dorsett, 2004; Goulden, 2010; Matty, 2013; Payne and Payne, 2000) suggested covariates include, for example, the duration and stability of any previous employment experience, age, the presence of caring responsibilities, long-term health conditions, qualification levels.⁹ The set of independent variables utilised in the analysis is the richest possible array of covariates available within the administrative datasets and is achieved through a complex process of cleaning and linkage across several of DWP's administrative datasets.

5.5 Findings: Differential payments but *still* differential outcomes across payment groups

To get an immediate sense of the nature and scale of the differential achievement of outcomes within the Work Programme, Figure 5.2 simply plots the job outcome rates for each of the DWP defined payment groups.

It is pertinent to begin by considering what we would expect to see in terms of the patterning of programme outcomes had the Work Programme achieved 'differentiated universalism'. Success on this objective would be the equalisation of job outcome likelihoods through well calibrated variation in supports and, as a result, the consequent elimination of variation in outcome performance. As this author has argued previously, "*If the differentiated payment system is effectively calibrating providers' incentives between the Work Programme's claimant groups in terms of some idea of the 'average claimant' within each of these groups, then one would, on average, expect the job outcome rates to be relatively evenly balanced between the various payment groups*" (Rees et al., 2014, p. 229, emphasis in original).

⁹ Such characteristics have been identified in previous qualitative research as the basis on which Primes and other providers, under economically rational behaviour, have developed profiling tools through which to 'triage' their caseloads to identify those 'rated 'green', focusing energies and resources on those easiest and most likely to move into work, whilst parking claimants rated 'red' who are considered to need more time and resource to support back into work' (Rees et al., 2014: 228).

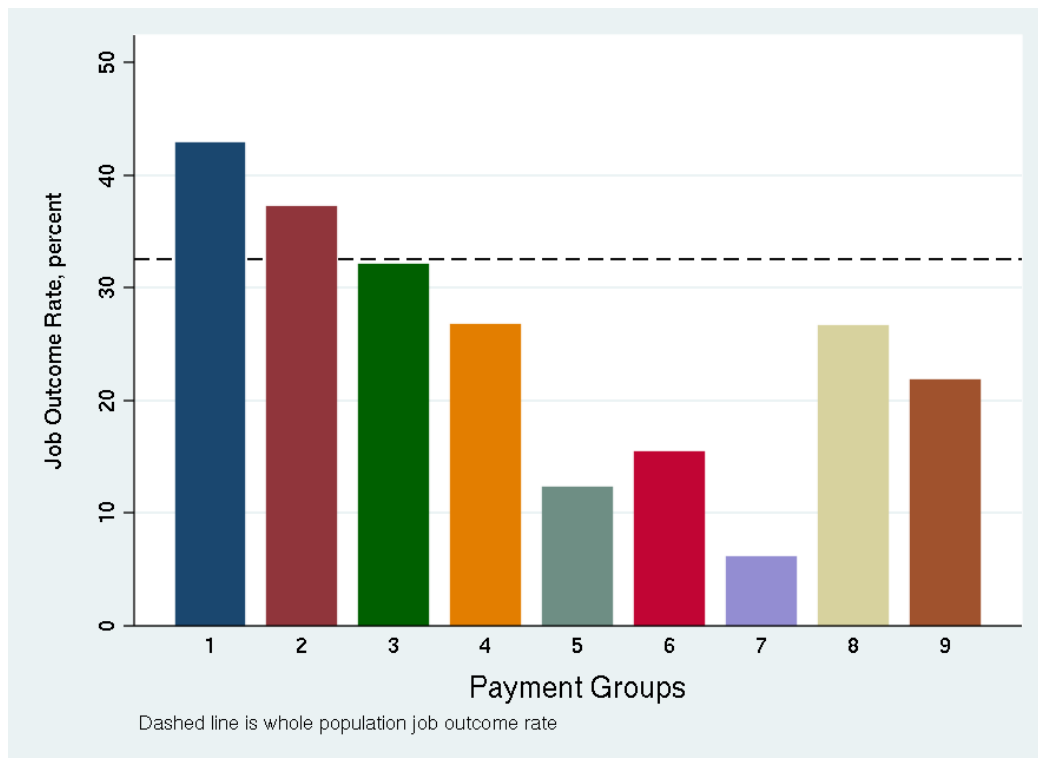


Figure 5.2 Different job outcome rates between Work Programme payment groups

Figure 5.2 contrasts with this expectation, however, showing considerable imbalances in job outcome rates across the payment groups. Two groups are performing markedly better than the cohort average job outcome rate (indicated by the horizontal line at 32.6%) while a number of payment groups are doing notably less well. Strikingly, there is a gap of 37 percentage points in the job outcome rate between the best and worst performing payment groups (JSA 18 – 25 and ESA ex-IB, respectively).

The divergence in job outcome rate between payment groups 3, 4 (and 6) is particularly notable, given that the maximum payment to providers for the achievement of sustained job outcomes for each participant within these three groups was pegged at the same level (in year one of implementation, Figure 5.1). Furthermore, the payment group associated with the highest potential per-participant outcome fees (payment group 7, ESA ex-IB) is the group with the lowest job outcome rate. Under the DWP's logic of outcome payments being scaled to the *average* difficulty of participants in each group, these analyses flag real questions around the ability of the Work Programme's payment mechanism to deliver the objective of differentiated universalism.

Figure 5.3 plots the job outcome rate for each payment group for each month's cohort of participants and by monthly attachment date. Such trend analysis for final job outcomes across monthly cohorts has not been possible previously because of the aggregation methods used in the presentation of publicly available statistics through Stat-Xplore.

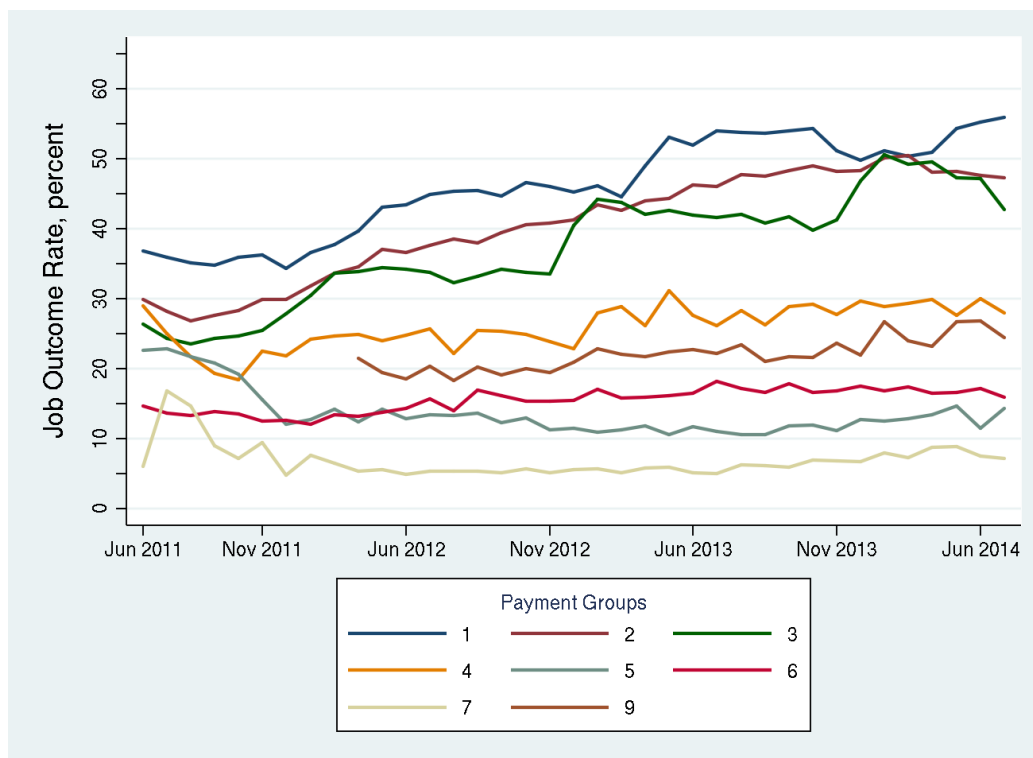


Figure 5.3 Divergent job outcome rates between Work Programme payment groups across the implementation period. Source: DWP Administrative data. Note Payment group 8 is not plotted due to insufficient population size in month-to-month cohorts for the production of percentage figures

Here again the difference in job outcome rates across the payment groups dominates the graph. Not only does the gulf in outcome rates across different payment groups persist over time but it actually worsens: better performing groups show some degree of improvement in job outcome rate from early 2012 whilst the worst performing groups show no indication of performance improvement across the implementation period. It should be noted that the eligibility for participation in the Work Programme through payment group 6 (a key health and disability cohort that has been subject to continued scrutiny during Work Programme due to its persistent poor performance) has been widened over time such that the group has tended to become more challenging to support. The performance of payment group 6 did improve in the final period of programme implementation, after this analysis period, under concerted attention from DWP performance managers following sustained media and select committee critique of poor performance (DWP, 2017a).

The large and systemic gaps in the job outcome rate between the payment groups (Figure 5.2 and Figure 5.3) indicate that the payment structure – which is the key accountability mechanism through which to calibrate provider incentives – does not seem to be successfully equalising the prospects of sustained employment for the average participant within each group. Indeed, even in relation to the relatively modest ambition for the payment groups to function “as a sensible basis to work with” (Graying in WPSC, 2011, p. 28) in capturing “some of the *average* difficulty” (PAC, 2012, p. 26, emphasis added) the payment groups are floundering. Overlaying this failure to reflect even the average difficulty of participants within each payment group is the more substantial issue (as will be seen from the modelling work below) that a

participant's payment group membership is a weak predictor of job outcome likelihood.

This brings us to a more fundamental consideration of the extreme and undue coarseness of the Work Programme payment groups, which principally use age and the type of benefit received as a proxy for the level of participants' perceived support needs (PAC, 2012, 2013, WPSC, 2013, 2011). It is therefore crucial to consider Work Programme's ability to deliver differentiated universalism *within* as well as *across* the DWP defined payment groups. This research's unique access to DWP's individual-level participant data brings the opportunity for the first time to investigate comprehensively *within* payment group variation in order to explore the hypothesis that this is where the bulk of the variation lies, though missed entirely by the Work Programme differential payment design.

An initial and particularly obvious route to consider the existence of within group variation in the achievement of job outcomes is to look *inside* payment group 3 – the JSA early access group – which comprises diverse participants routed to the programme through three distinct referral pathways: mandatory entry of 18-year-olds not in education, employment or training ('NEETs'); mandatory entry of JSA 'repeaters' (those receiving JSA for 22 of the previous 24 months); and voluntary early entry for pre-identified 'vulnerable' JSA claimants (DWP, 2013).

Each of these groups have distinctive characteristics and it is not clear why each sub group would be expected to have an identical job outcome likelihood (on average), which is the implication of rolling the three together within the same payment group and fee structure. Figure 5.4 sets out the job outcome rate for each of these sub-groups within payment group 3. There is indeed considerable variation in job outcome rate within this single payment group, yet all participants carry the same financial rewards to providers.

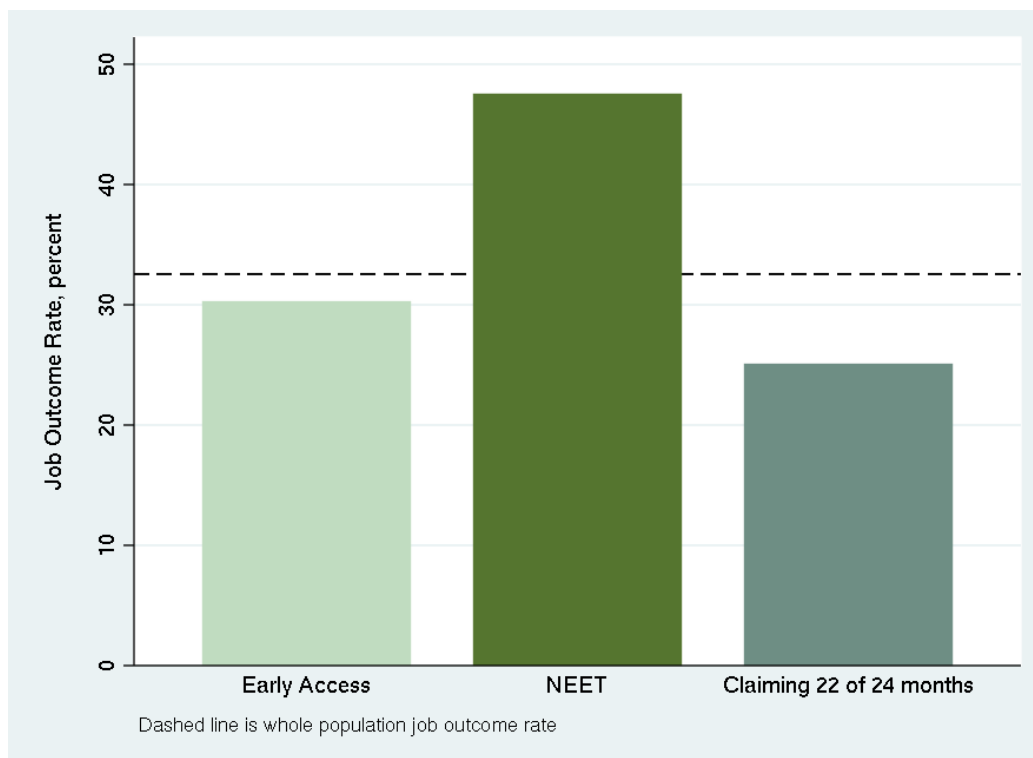


Figure 5.4 Variation in job outcome rates for sub-groups within payment group 3. Source: DWP administrative data

5.6 Moving beyond official ‘payment groups’

Moving beyond payment group 3, there are a number of other characteristics that cut across all payment groups and against which there is a gradient in the achievement of job outcomes. Much of the previous research which has raised concerns about the Work Programme’s ability to ‘work for all user groups’ (Davies and Raikes, 2014; Riley et al., 2014) has focused on the low performance of particular payment groups, with specific concern framed around the performance of payment group 6 for ESA claimants where payment group membership is attained by demonstration of a disability or long-term health condition through DWP’s Work Capability Assessment. Disability emerges as a particularly striking characteristic against which the Work Programme is failing to deliver equitable employment outcomes (Quirke, 2015; Rees et al., 2014; Riley et al., 2014).

In practice, however, disability and longstanding health conditions are experienced by participants across all payment groups far beyond DWP’s core confines of the Work Capability Assessment for the purposes of determining eligibility to ESA and, hence, far beyond payment group 6. There are a substantial proportion of participants within each of the nine payment groups who self-identify as having a disability (Figure 5.5, upper right pane). A quarter (25.7%) of participants within the JSA 25 plus payment group (payment group 2) report themselves as having a disability or long-term health condition for instance. As seen in the upper left pane of Figure 5.5, those who identify themselves as having a disability have a job outcome rate around half that of those who do not. This echoes previous qualitative analysis which notes some degree of provider surprise in terms of the substantial mental and physical health concerns of

participants in ‘easier to help’ JSA payment groups (Rees et al., 2014). Even within the better performing JSA payment groups those participants with a self-reported disability exhibit a job outcome rate that sits well below the aggregate programme average (as seen in the lower pane of Figure 5.5).

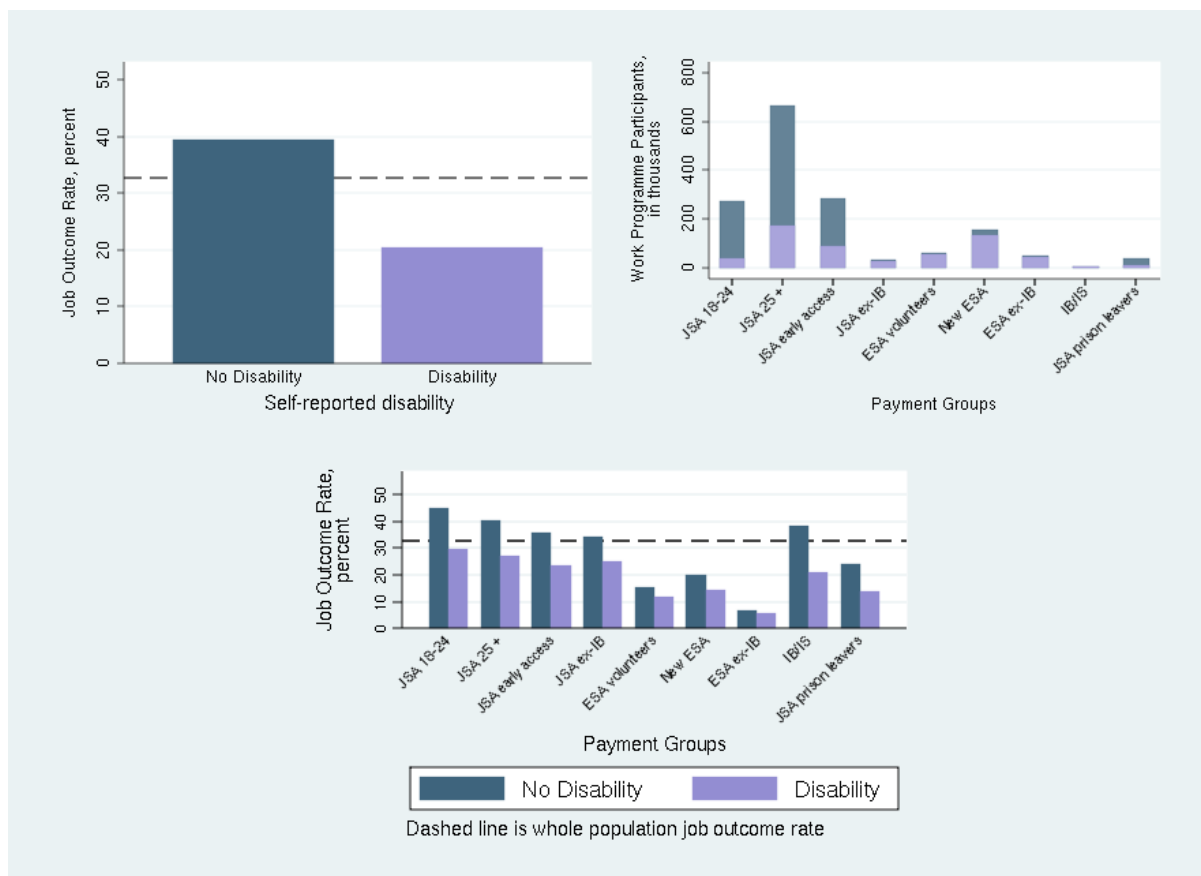


Figure 5.5 Participants with disabilities and/or health conditions within the Work Programme payment groups: Source: DWP administrative data

5.7 Concealed characteristics: extending the analysis of differentiated universalism beyond payment groups

In this context, therefore, it is crucial to move beyond the types of aggregate-level and between-payment-group analyses that dominate the current evidence base around Work Programme and instead to develop more detailed individual-level and within-payment-group analyses.

The ability of previous research to investigate those individual characteristics that are effectively concealed *within* the DWP payment groups has however been severely constrained in two key ways by the form of the publicly available programme data. Firstly, outcome breakdowns are only available in aggregated form over relatively few characteristics beyond the official payment groups: age band, disability indicator, ethnicity, gender, lone parent status, and broad classes of medical condition. Secondly, the tabulation tools in Stat-Xplore only permit a limited number of cross-tabulations to be performed simultaneously such that more subtle multivariate analyses are precluded.

Figure 5.6 instead uses the DWP administrative data to neatly summarise the patterning of job outcomes across those characteristics that are publicly available whilst also serving as a marker for the maximum, but limited, extent of analysis currently viable through the publicly available aggregate data.

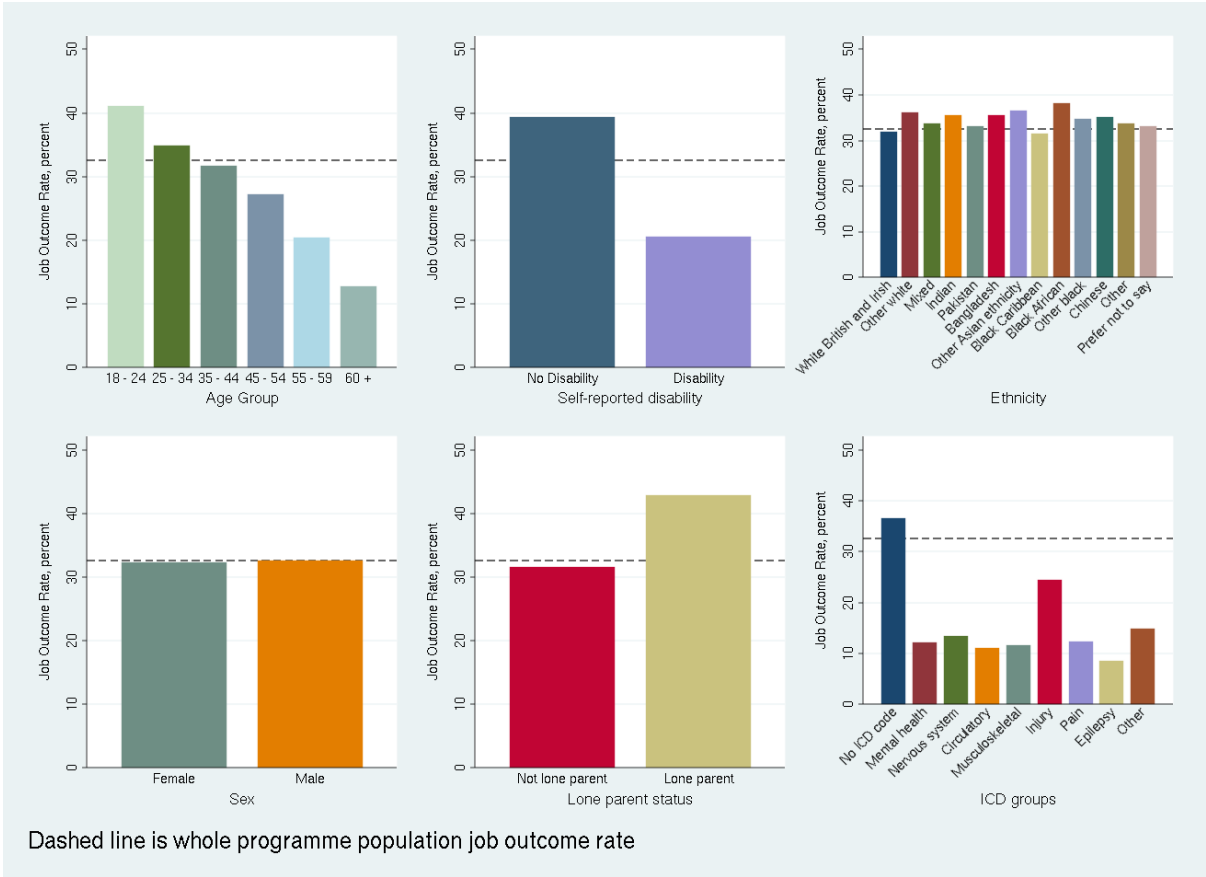


Figure 5.6 Variation in job outcome rate across observable participant characteristics in publicly available data. Source: DWP administrative data

5.8 Getting personal

In exploring issues related to differentiated universalism the availability of DWP administrative data to the present study remedies the limitations of publicly available data on two fronts. Firstly, via the complex cleaning and linkage of different DWP administrative datasets it has been possible at the level of the individual to connect the outcomes data with a considerably larger and broader array of ‘employability’ related explanatory variables than ever previously studied. Secondly, the granularity of the data enables far richer multivariate analyses at the necessary individual level. Access to these administrative datasets, alongside the patience and capacity to link them together, therefore unlocks for the first time the most powerful and appropriate analytical foundations to assess the achievement of differentiated universalism within Work Programme according to the DWP’s own data.

The multivariate binary logistic regression models are estimated for the full available sample of 1,563,874 programme participants. Model testing identified the variables shown in Figure 5.7 as the most parsimonious.

Benefit history in period prior to Work Programme	Health and disability	Demographics
Benefit type and history	Self-reported disability	Age
JSA only (ref)	Yes (1) No (0)	18-24 (ref)
JSA ex-IB	Primary health condition	25-34
Contribution-based ESA	Participants without any diagnosis code (ref)	35-44
Income-related ESA	Mental health (excluding depression, anxiety, and stress, which are reported separately)	45-54
ESA ex-IB	Depression	55-59
IB/IS	Anxiety	60+
Number of days claiming ESA within period 2-years prior to Work Programme attachment	Stress	Disadvantage due to drug use, alcoholism or ex-offender status
(Ratio)	Circulatory or respiratory	None (ref)
Has spent any time claiming Incapacity Benefit in period 5-years prior to Work Programme attachment	Musculoskeletal (excluding back pain)	Ex-offender
Yes (1) No (0)	Back pain	Misuser of drugs
Number of distinct spells claiming ESA (or equivalent) in period 5-years prior to Work Programme attachment	Injury	Ex-offender and misuser of drugs
(Ratio)	Pain (excluding back pain)	Misuser of alcohol
	Alcoholism	Ex-offender and misuser of alcohol
	Drug abuse	Misuser of drugs and alcohol
	Epilepsy	Ex-offender and misuser of drugs and alcohol
Employment history in 5 years prior to Work Programme attachment	Arthritis	Qualification level
Number of days in employment	Other ICD class	No record available (0)
(Ratio)	Cumulative impact of disability on every-day tasks	Low qualification (1)
Any period in employment	Incremental score where higher value is greater number of challenges	English language challenges
Yes (1) No (0)		None (0)
Distinct employment spells		Written or oral English language difficulties (1)
(Ratio)	Household characteristics	Homeless
	Gender	Yes (1) No (0)
	Male (1) Female (0)	Ethnicity
Local context	Has dependent children *	White British or Irish (ref)
Local authority JSA claimant rate	Yes (1) No (0)	Other white
(Ratio)	Woman with youngest child aged under 5	Mixed
Index of Multiple Deprivation deciles	Yes (1) No (0)	Indian
Reference is 10% least deprived LSOAs	Lone parent aged under 21	Pakistan
	Yes (1) No (0)	Bangladesh
	Lone parent aged over 21	Other Asian ethnicity
	Yes (1) No (0)	Black Caribbean
		Black African
		Other black
		Chinese
		Other ethnicity
		Prefer not to say
		Refugee or other humanitarian protection
		Yes (1) No (0)

Figure 5.7 Explanatory variables used in logistic regression analysis

The explanatory variables shown across Figure 5.8 to Figure 5.12 were included within a single logistic regression model but are presented below within separate figures simply for ease of visualisation and discussion. These explanatory variables have been grouped into five blocks covering: benefit history; health and disability; socio-

demographics; employment and household characteristics; and local context. These are described in turn below, but to clarify: their effects were estimated together within a single model.

In reading the regression output it is pertinent to recall what it would mean in statistical terms if differentiated universalism *had* been achieved in practice. Importantly, the analysis here has been conducted on real programme participant and outcomes data, that is, any job outcomes achieved or unachieved by participants have been delivered *within* the Work Programme's differential payment system. If the differential payment system had been successful in ensuring that '*all programme participants have the same likelihood of entering and sustaining paid work, regardless of their characteristics and circumstances*' then the odds ratios associated with each explanatory characteristic should all be equal to, or very close to, one.

To clarify interpretation, an odds ratio of 1 for categorical variables means that the odds of one group achieving the outcome of interest is the same as the odds of outcome achievement for the reference category. Each of the Figures from 5.8 to 5.12 are bifurcated with a vertical line where the odds ratio is equal to 1 to aide with interpretation. If job outcome likelihoods are equal regardless of characteristics then theoretically these indicators should no longer function as statistically significant explanatory variables. In practice the very large sample size available to this study means that even marginal variations according to characteristics are likely to emerge as 'statistically significant' (indeed *all of the independent variables included in the model and set out in the figures below are statistically significant predictors of job outcomes, where $p < 0.001$*) hence, it is arguably more appropriate to focus on the substantive *size* of effects (as informed by Ziliak and McCloskey, 2008).

5.8.1 Benefit history in period prior to Work Programme

The first block of variables relates to benefit claiming history. Those who are referred to the programme whilst claiming disability-related ESA are significantly less likely to trigger a job outcome than those claiming mainstream JSA unemployment benefits, *ceteris paribus*. This negative effect is larger for those who are in receipt of income-related ESA than contribution-based. Participants who formerly claimed Incapacity Benefit and who have been transferred to ESA are less than half as likely to achieve a job outcome, compared to those in receipt of JSA, controlling for other factors. An increased duration of time claiming ESA is associated with increased distance from the labour market, since each additional day a participant has been in receipt of ESA over the 2-year period preceding attachment decreases their likelihood of entering stable employment, all else equal. Similarly, having spent any time claiming incapacity-related benefit in the five-years prior to programme participation and additional spells of ESA or IB receipt are each associated with a reduced likelihood of achieving sustained employment.

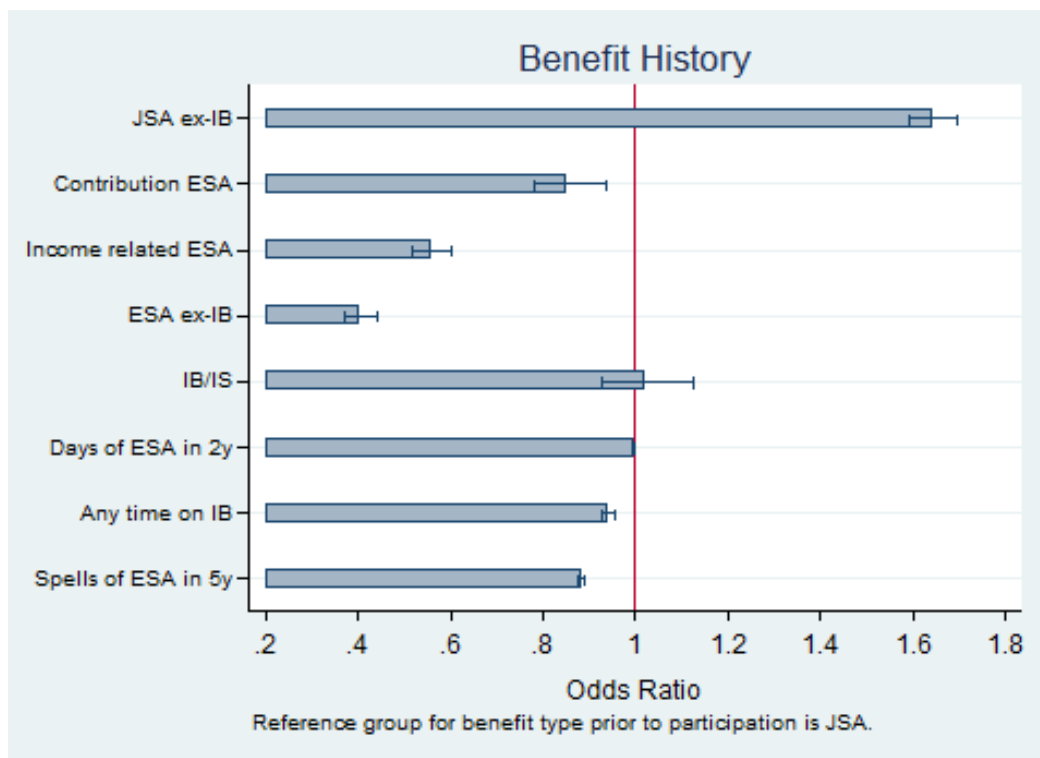


Figure 5.8 Predicting the achievement of job outcomes through administrative data: benefit history. Source: DWP administrative data

5.8.2 Health and disability

The next set of explanatory variables relate to health and disability-related characteristics. These are known to have been important factors in structuring access to Work Programme support services (WPSC, 2013) and have been an area of concern in relation to poor performance during the programme (Quirke, 2015; Riley et al., 2014; WPSC, 2013).

The administrative data contain information related to ‘self-reported disability’ where this has been disclosed by a participant to their Jobcentre Plus work coach prior to attending the Work Programme. The data also includes departmental records of primary health conditions according to the International Classification of Diseases (ICD) – the standard diagnostic tool for epidemiology and health management. The medical conditions recorded within claim forms are related to evidence submitted by participants at the outset of their claims and through the ‘Work Capability Assessment’ and therefore are only generally available for those claiming ESA.

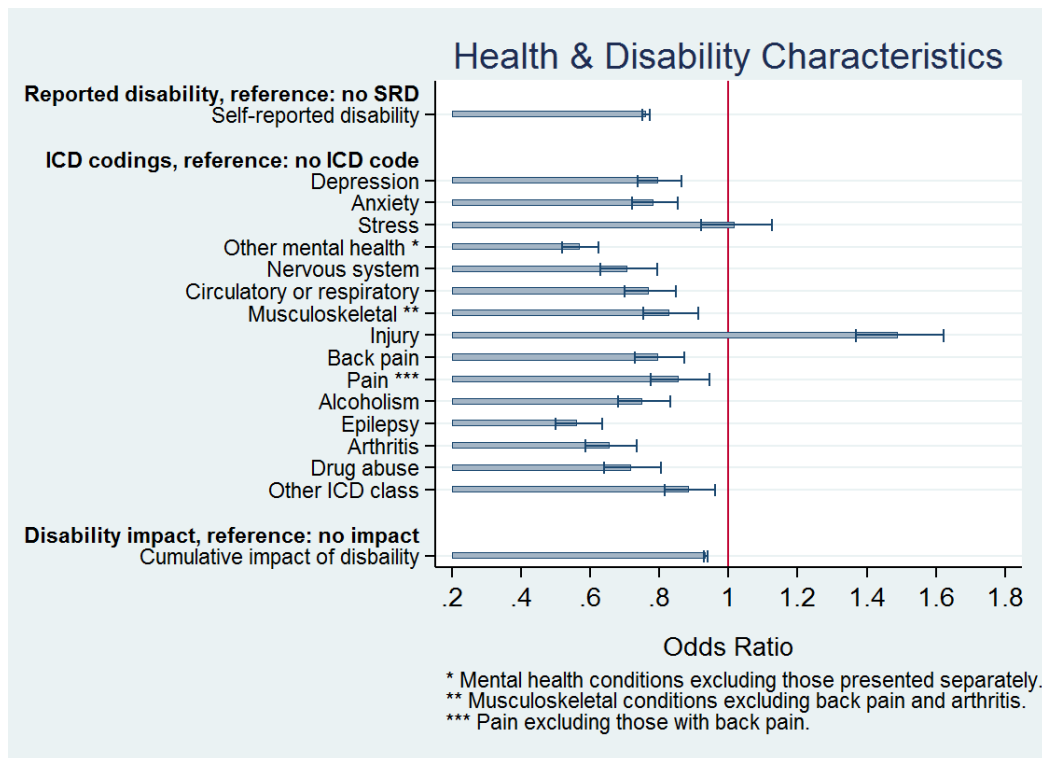


Figure 5.9 Predicting the achievement of job outcomes through administrative data: health and disability characteristics. Source: DWP administrative data

Importantly these two sets of information – self-reported disabilities and, for ESA claimants, ICD records – do not neatly connect and relate to one another. Close to one-hundred percent of participants in payment groups where eligibility results from the receipt of disability-related benefits (i.e. ESA or IB) (payment groups 5, 6, 7 and 8) have an ICD record. A substantial proportion of payment group 9 participants also have a formal ICD record. For the JSA-related payment groups, in large part the DWP does not have a formal record of ICD even where participants self-identify as having a disability. Overall nearly 60 percent (59.56 %) of participants who report as having a health condition or disability *do not* have an ICD record. Because of the lack of coherence across these disability-related variables, both self-reported and ICD records are included as independent variables and retain significance when included alongside one another.

Controlling for other factors (including official ICD codes) those participants who self-report as having a disability are around three-quarters as likely (OR = 0.76) to trigger a job outcome as those who do not identify themselves as disabled. Most of the ICD code categories are associated with quite considerable effect sizes compared to those without a diagnosis record and each (aside from injury) serves to reduce the likelihood of an individual achieving a job outcome, *ceteris paribus*. A large proportion of the ICD code categories have an odds ratio of between .7 and .8 meaning that participants with these conditions have odds of achieving sustained employment between 20% and 30% less than those without an ICD marker, all else equal. Those with a mental health condition which is not captured within the separate groups for depression, anxiety, stress and addictions are less likely than other ICD groups to enter sustainable employment, holding other factors constant. This suggests that the Work Programme

may be working particularly poorly in terms of effective support for those with more severe and less common mental health conditions such as schizophrenia and bipolar disorder.

The cumulative disability variable has been constructed by summing the number of daily activities which are adversely impacted by the presence of an individual's health condition or disability: a score of zero reflects no impact on day-to-day tasks whilst a score of ten reflects a maximum that all ten daily tasks and usual activities are negatively impacted by a person's disability. This measure is treated as a continuous variable and each additional 'challenge' is associated with a reduced likelihood of achieving sustained employment, *ceteris paribus*.

The effect sizes of both self-reported disability and the ICD codes is notable as it suggests that the formal ICD records of primary disability held by DWP do not sufficiently capture any 'disability effect' on a participant's employment prospects. The size of the population reporting a disability within the JSA groups (25.6 % of those in JSA-related payment groups report themselves as having a disability) and the sizeable effect of self-reported disability (as noted, OR = 0.76) highlights the failings of DWP data to capture relevant information on these participants' health conditions. Improved information on participants' health conditions and disabilities across benefit types – not just for those undertaking the Work Capability Assessment for the purposes of ESA eligibility – is likely to be particularly important in ensuring that participants are appropriately served both within Jobcentre Plus and beyond.

5.8.3 Demographics

Key demographic factors are the focus of Figure 5.10. Controlling for other factors in the model, being in the 60 plus age bracket is associated with a very low probability of achieving a job outcome. Those aged over 60 are around one-fifth as likely to achieve a job outcome as those aged 18 – 24, all else constant. This finding tallies with the performance achieved for older programme participants in previous schemes such as the New Deal and Flexible New Deal (Foster et al., 2014). Mainstream employment support programmes have consistently failed to achieve comparable employment outcomes for older and younger participants and it is therefore surprising that within the Work Programme's differential payment model age was only included by separating out those JSA claimants aged under 25. By grouping all participants aged over 25 within the same payment categories the Work Programme payment schedule neglects the often-seen age-related differential in job outcome rates.

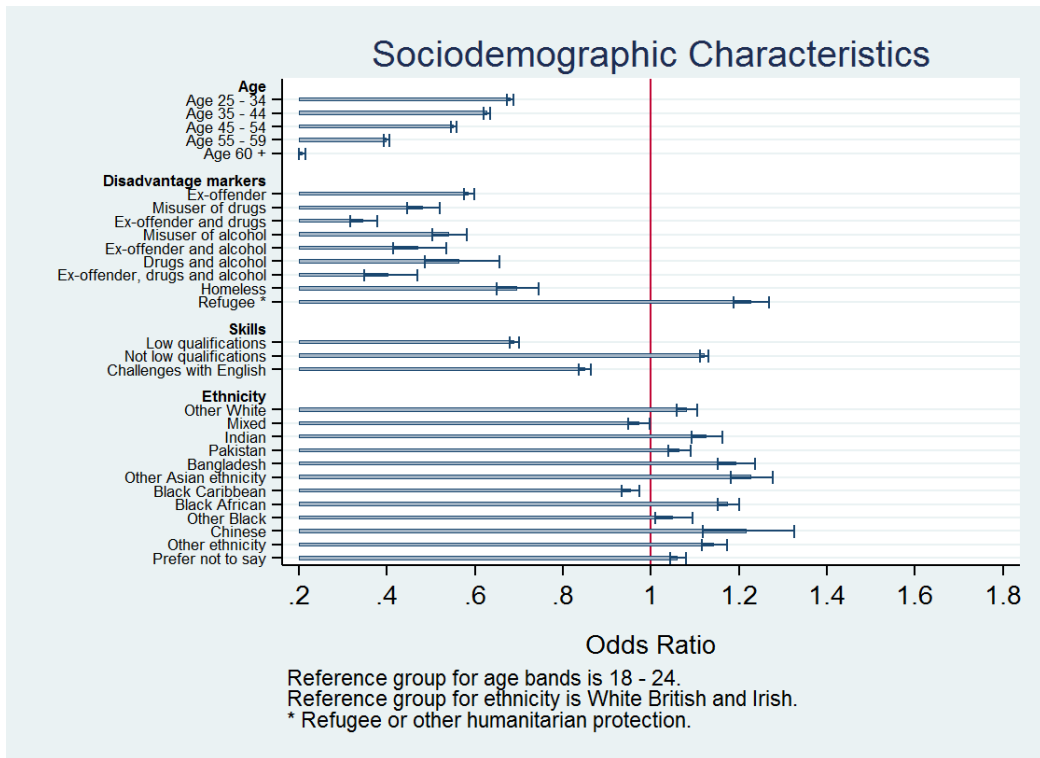


Figure 5.10 Predicting the characteristics of job outcomes through administrative data: demographic characteristics. Source: DWP administrative data

5.8.4 Employment history

The next block of variables assessed relates to employment history (upper portion of Figure 5.11) and household characteristics (lower portion of Figure 5.11). Although the effect associated with the number of days in employment seems small, this should be considered alongside the unit of this continuous variable (i.e. the odds ratio of 1.000629 is for a single additional day in employment within the five year period). Were a participant to have spent an additional 6-months in employment within the five-year period prior to Work Programme referral then their odds of entering sustained work would be 1.15 times larger than another participant, without the additional 6-months in work, all else equal. Those participants who do not have a record of employment within the period 5-years prior to programme participation have odds of achieving sustained employment that are only around three-quarters as large as the odds for someone who does have an employment record, ceteris paribus. Each additional employment spell is associated with an increased likelihood of achieving a Work Programme job outcome.

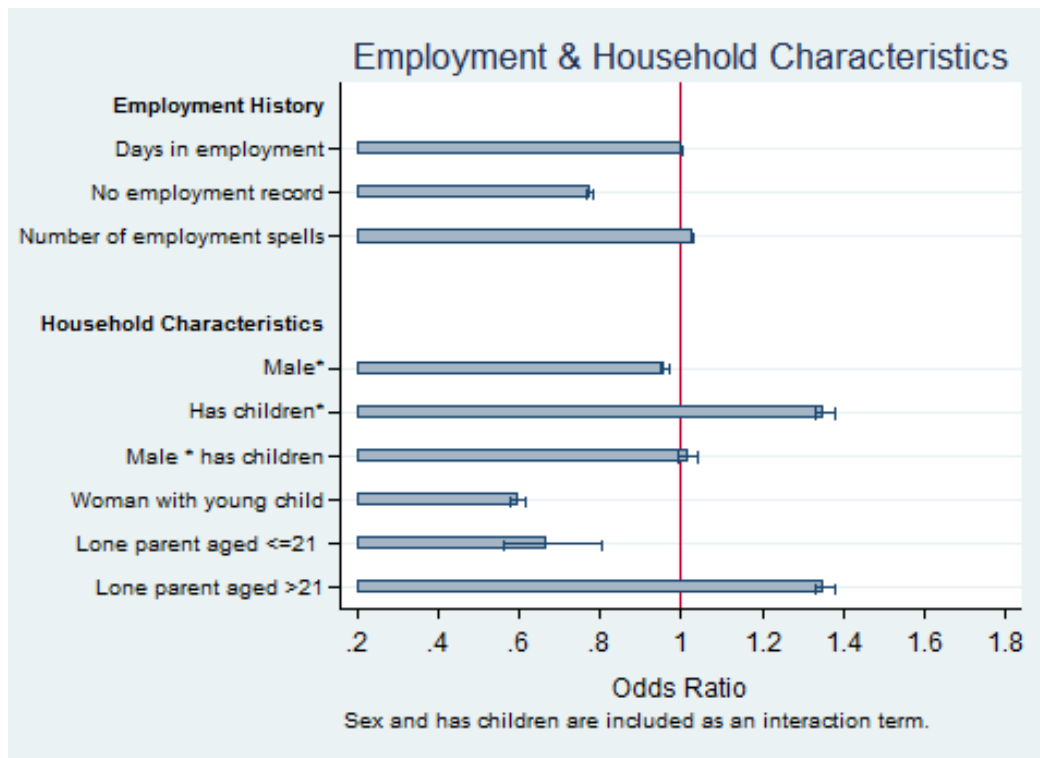


Figure 5.11 Predicting the achievement of job outcomes through administrative data: Employment history and household characteristics. Source: DWP administrative data

5.8.5 Household Characteristics

In the lower portion of Figure 5.11 gender and the presence of dependent children (aged 16 or under) are included as an interaction term, thus allowing the model to account for differences between men and women with respect to the effect of dependent children on employment outcomes. There is no significant difference in the likely achievement of job outcomes between men who have dependent children and women who do not have dependent children. Compared to women with dependent children, men without dependent children in the household are significantly less likely to achieve job outcomes. All else equal, women with dependent children are 1.3 times more likely to achieve a job outcome than women without dependents although this effect does not hold when only very young children are considered. Where children are of school age the presence of dependents within the household is consistently associated with increased likelihood of entering work, controlling for other factors.

As with previous research, younger lone parents fare consistently less well than non-lone parents in terms of their job outcome rates (Rees et al., 2014). Contrastingly, all else equal older lone parents are 1.35 times more likely to achieve a job outcome than non-lone parents. This may be because older lone-parents tend to have stronger human capital and fuller employment histories than younger lone parents (Coleman and Lanceley, 2011) and this elevated human capital may not be fully detected through the limited variables on qualifications and work histories available in the administrative data.

5.8.6 Local geographical context

The final set of explanatory variables relate to local geographical context and are shown in Figure 5.12. Geography is a generally neglected feature of what is an overwhelmingly a-spatial Work Programme, both in its policy design and empirical research analyses (discussed further in Chapter 6). Yet local authority JSA claimant rate has a negative effect on an individual's prospects of achieving a sustained employment outcome: for each additional percentage point on the local working age population claiming JSA, the odds of achieving an employment outcome are 0.95 times as large, all else equal. The degree of deprivation in a participant's home LSOA (akin to a much smaller 'neighbourhood') is also a substantive predictor of job outcomes. For participants with otherwise identical characteristics, those living in the most deprived 10% of LSOAs have odds of achieving job outcomes that are only two-thirds as large as the odds of sustained employment for those in the least deprived areas.

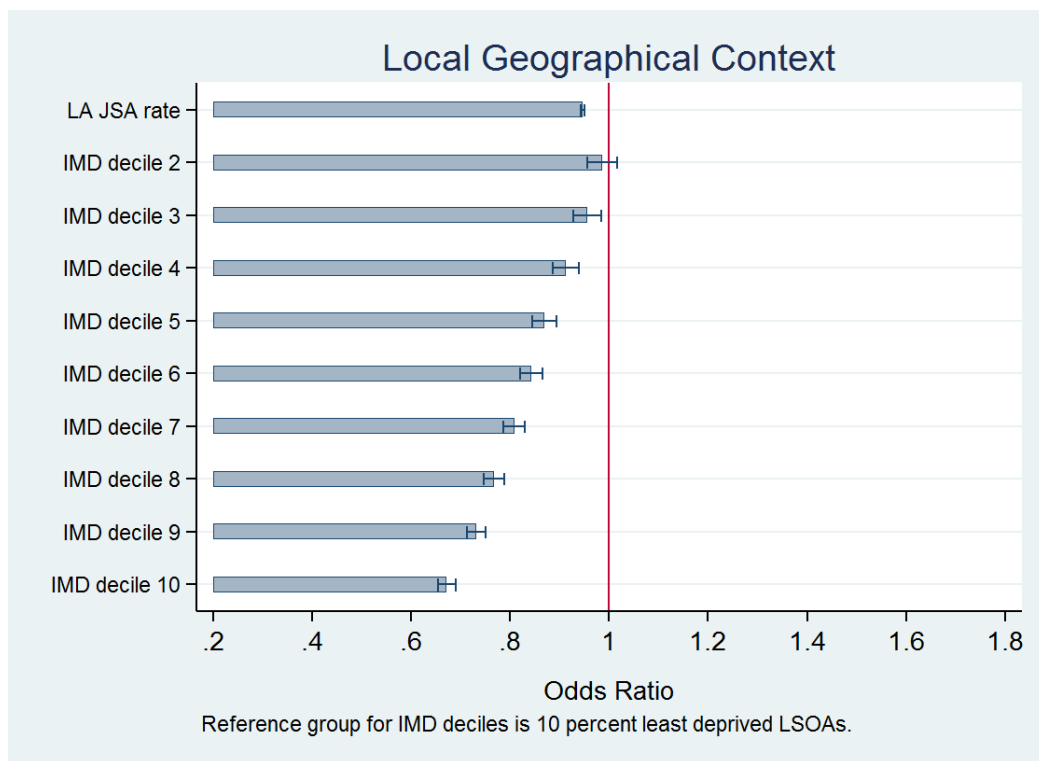


Figure 5.12 Predicting the achievement of job outcomes through administrative data: local geographical context. Source: DWP administrative data coupled with externally prepared claimant rate and IMD figures

5.8.7 Summary of model findings and implications for differentiated universalism

If functioning as intended, the payment groups themselves – coherently and internally – would perfectly capture all variation in participant characteristics associated with labour market outcomes. The difference in payment levels across the groups would then ensure that those who, at the outset had the lowest chance of finding work, ultimately have the same job outcome rates as the most immediately employable members of the cohort. If differentiated universalism had been achieved via the Work Programme's differentiated payment system then there are two important and related signifiers that we would expect to observe within the multivariate modelling process.

The first signifier in this context would be that *post* the application of the differential payments none of the usual markers of labour market disadvantage would display non-null effect sizes (i.e. odds ratios markedly different from one in binary logit models such as these) since any variation in job likelihoods would have been fully captured and compensated for within the payment structure. Secondly, and relatedly, subsequent to the application of the differential payments all programme participants would have similar predicted probabilities of achieving job outcomes.

Importantly, the analysis shown here is conducted *after the application of the differential payment structure* and yet Figure 5.8 – Figure 5.12 indicate that this first indicator of differentiated universalism is not met. Indeed, what we see is quite the opposite: across a large and relatively diverse set of independent variables there are statistically significant and (more importantly, given the very large sample) sizeable effects. Instead of neutralising the effect of these explanatory variables, in the aftermath of Work Programme’s payment-by-results system there remain persistent, sizeable and significant effects across a range of well-known indicators of labour market disadvantage. This inevitably has implications for the likely achievement of job outcomes for individuals with different personal characteristics and contexts.

5.9 Considering the differentiated payment system as a ‘model’: possibilities for enhanced profiling in the UK context

As noted above, the second signifier for the achievement of differentiated universalism is that the payment system should ensure that all participants would have very similar – if not identical – predicted probabilities of achieving a job outcome.

The blue bars in Figure 5.13 plot the predicted probability of each participant achieving a job outcome (according to the full administrative model) and these probabilities range from close to 0, where a participant’s characteristics mean that they appear incredibly unlikely to achieve a job outcome, to 0.85, where a participant expresses characteristics that mean they have a strong likelihood of entering sustained employment. Participant predicted probabilities vary widely: who you are does matter in relation to your likely employment outcomes under the Work Programme, even after application of the differential payment model. Differentiated universalism has not been achieved, despite this being a key policy objective.

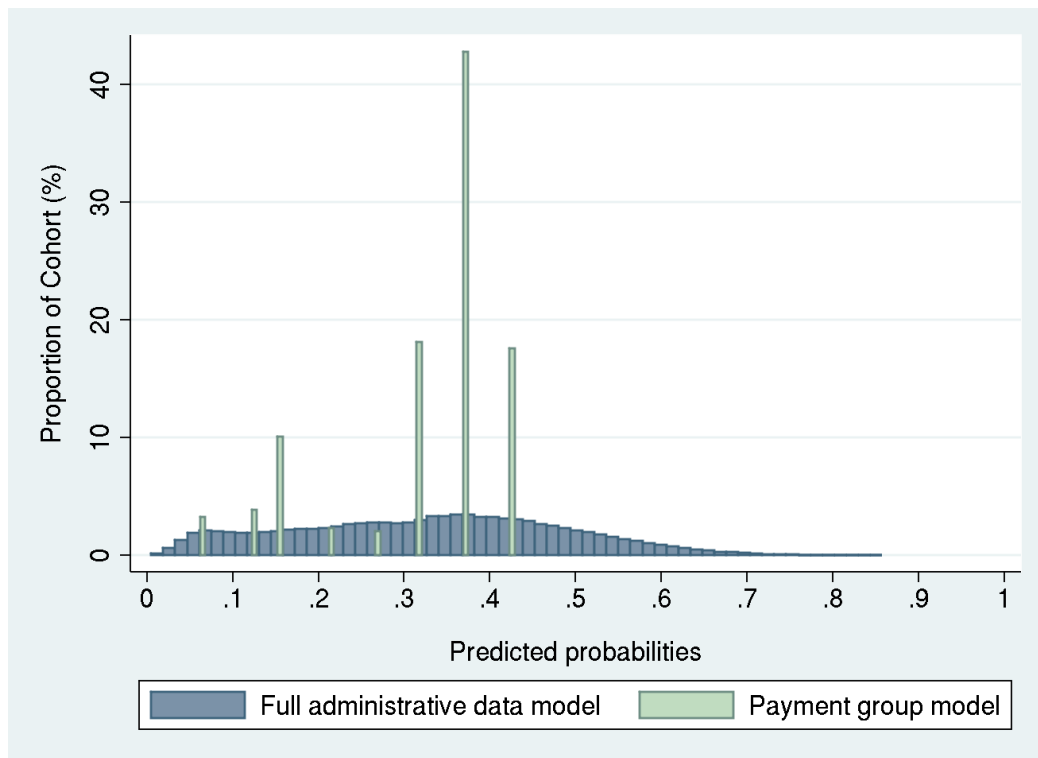


Figure 5.13 Exploring the achievement of job outcomes through administrative data: distribution of participant predicted probabilities. Source: DWP administrative data

The key assumption within the DWP differential payment model is that the payment groups act as effective proxies for participants’ distance to the labour market and that by attaching different prices to these groups any differences in job outcome likelihood can be neutralised. Having now established that Work Programme’s payment system has failed to deliver differentiated universalism as desired, it is possible to next consider the reasons, implications and possibilities for improvement.

There are two principle routes through which the payment system can have failed in this regard. Firstly, there may be an issue only in the realm of relative price setting. Here, the payment groups may perfectly capture the variation in participant characteristics – that is each payment group contains within it participants who are ‘similar’ in their likelihood of achieving job outcomes – but the differential in payment levels (i.e. the steepness of the incentive gradient across the groups) may be inappropriately calibrated, such that ‘harder’ groups perform worse in terms of their job outcomes. Secondly, and more fundamentally, the payment groups may themselves be internally heterogeneous in terms of participant characteristics such that the official groups fail to adequately capture differences in participants’ likelihoods of achieving job outcomes. Given the availability of individual-level administrative data here it is possible to fully explore these issues for the first time.

Though it usually is not thought of as such, the DWP differential payment system can be considered as a simple ‘model’ – that is a bivariate logistic regression model where job outcome is the dependent variable and payment group membership is the only predictor. The green bars in Figure 5.13 illustrate the resultant predicted probabilities for participants under the payment group ‘model’. Thinking in this way can facilitate

an initial baseline understanding of the breakdown of the payment group system in its pursuit of differentiated universalism. The green bars in Figure 5.13 illustrate the resultant predicted probabilities for participants under the payment group ‘model’. By exploring the predictive power of the DWP payment groups in this manner a comparison of model ‘fit’ between the payment group model and the alternative, richer, administrative data model may then be made.

	Explanatory variables	Pseudo-R²	Correctly predicted 0s	Correctly predicted 1s
DWP payment group model	Work Programme payment group membership	4.04	100.00	0.00
Full administrative data model constructed by author	Full suite of employability-related explanatory variables available in administrative data, as outlined in Figure 5.7	9.58	90.71	24.43

Figure 5.14 Predicting the achievement of individuals’ job outcomes: a comparison of alternative models. Outcome variable is the achievement of the official programme job outcome

There is no agreement in the literature as to how best to evaluate the fit of binary logistic models such as these. Figure 5.14 offers three commonly used model fit statistics – pseudo-R² and correctly predicted groups (which indicates the proportion of job outcomes which are successfully predicted). Whilst commonly used with binary logit models these measures of model fit are known to be problematic – particularly the use of the default threshold of 0.5 to separate predicted positives and negatives – these are presented here for initial ease of illustration and discussion given that these outcomes are rare events. Although these metrics should be handled with care, they do however begin to give an indication of the relative predictive power of the full administrative model compared to the DWP payment group design.

The full administrative dataset model detailed above has a larger, although still modest, pseudo-R² figure, explaining just under 10% (0.096) of the variation in job outcomes. This is nevertheless a notably stronger predictive tool than the payment-group-only model. What is notable is that predicting job outcomes is a challenging statistical task and in part this is likely to be driven by the ‘rareness’ of job outcome achievement amongst the Work Programme participant cohort. This serves as a reminder that programme participants are people who are long-term unemployed and less than one-third of participants achieve job outcomes. This means that it is considerably easier to predict ‘0s’ than ‘1s’: simply ‘predicting’ that no participant will achieve a job outcome can on some measures be considered a successful model in a technical sense, though this is clearly of limited practical use for policy.

The success of any statistical profiling depends inevitably on the quality of data available to predict outcomes. Where good explanatory data exist then such predictive modelling is shown to be reasonably powerful such that it can be used to aid policy in the targeting of services. Though such models will never be perfect and are often revised over time, Ireland, Czech Republic, Kentucky (US) and Australia, have each successfully instituted statistical profiling to aide with the targeting or streaming of people who are unemployed for the receipt of particular support packages (Black et

al., 2003; European Commission, 2015; Lipp, 2005; Loxha and Morgandi, 2014; O’Connell et al., 2010; Soukup, 2011). The limited set of potential predictor variables held within DWP administrative datasets has in the past been understood as a key impediment to the application of statistical profiling in the UK (Bryson and Kasparova, 2003; Matty, 2013) and this remains the case today.

In order to reconnect this model development work to the policy task at hand it is pertinent to consider what these models would need to facilitate in order to move the Work Programme closer to the achievement of differentiated universalism through a reformed superior payment model. Importantly, to more effectively enable marketised accountability levers to move towards differentiated universalism statistical models do not necessarily need to perfectly predict or fully account for variation in job outcome likelihoods. Rather, “the policy imperative for the Work Programme in terms of profiling and differential payments is the particular need to differentiate between the *relative* likelihood of unemployed individuals moving into sustained employment far more than it is to fully account for those employment transitions; the two are related, but they are not identical” (Carter and Whitworth, 2015, p. 288, emphasis in original).

In assessing the models’ ability to distinguish the *relative* likelihood of achieving employment outcomes this study follows previous work within the DWP (Matty, 2013) and in this author’s previous research (Carter and Whitworth, 2015) by adopting segmentation analysis, which is recommended as a more conceptually and technically meaningful approach to evaluate model power and utility in these contexts. In this segmentation approach programme participants’ predicted probabilities are broken into ten equally sized deciles with the 10 percent of cases with the lowest predicted probabilities on the left-most bar through to the 10 percent of cases with the highest probabilities on the right. For each segmented decile in terms of their predicted outcome probabilities, the vertical axis shows the proportion of each group who did successfully achieve an official programme job outcome. Within Figure 5.15 the left-hand pane relays information from the DWP payment group ‘model’ and the right-hand pane is for the full administrative data model.

The key features of the distribution that emerges from the DWP payment group model is both its irregular bumpiness – the deciles do not incrementally progress evenly from left to right – and the shallower gradient across the deciles when compared to the administrative model. The uneven way in which the decile job outcome rates progress, for example, with decile 4 having a job outcome rate markedly higher than both deciles 5, 6 and 7, means that the model is not successful in ‘ordering’ participants’ relative likelihoods of moving into sustained work. This would suggest that for the achievement of differentiated universalism there is an issue not only in the realm of relative price setting (since those with lower predicted probabilities tend to have lower job outcome rates) but that the payment groups themselves do not coherently, internally capture variation in participant’s distance to labour market. The administrative model is better able to effectively ‘line up’ participants according to their relative likelihoods of achieving sustained employment outcomes, and its steeper

gradient and wider reach at the top and bottom of the distribution highlight that it is better able to differentiate between participants who are the very closest and furthest from the labour market.

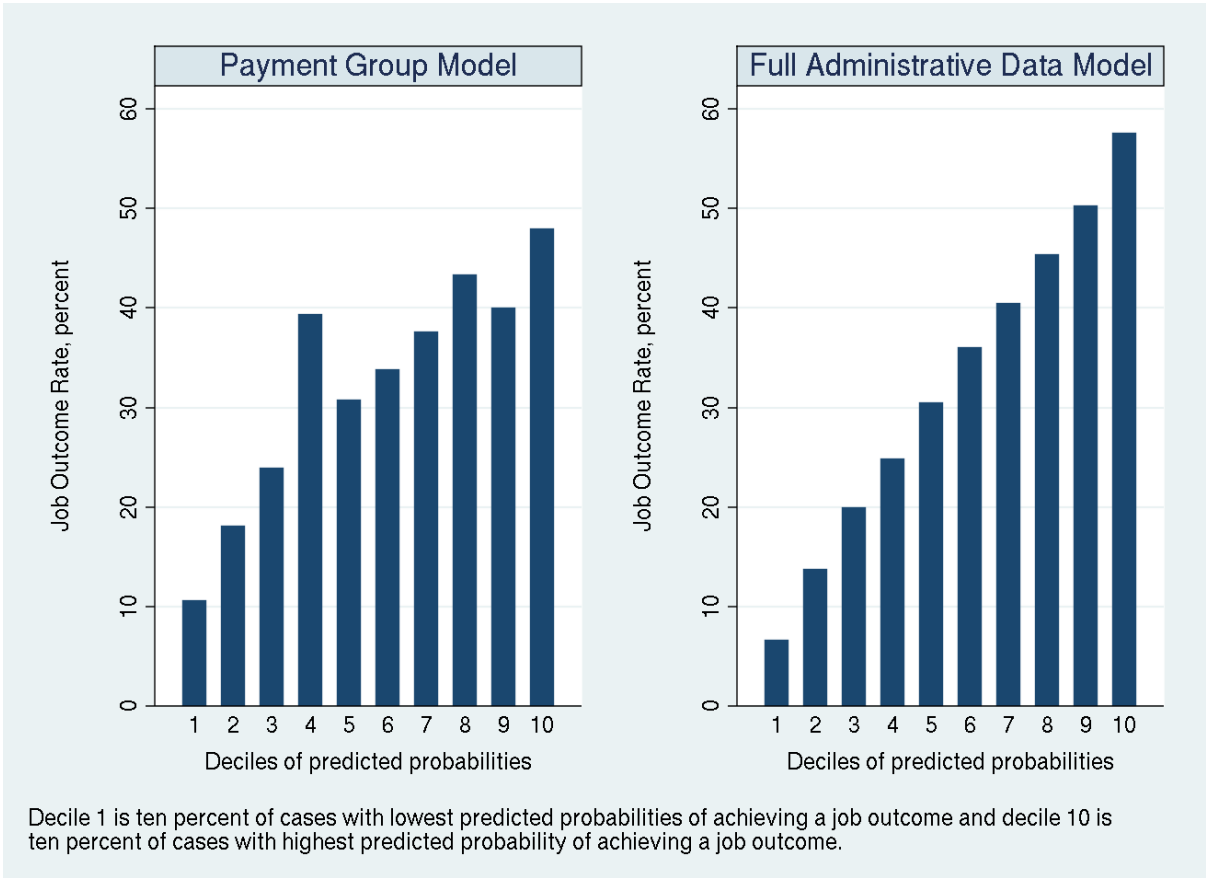


Figure 5.15 Claimant segmentation and predictive accuracy

These findings have two important implications. Firstly, assuming economic rationality on the part of profit motivated providers, given that all participants within each payment group share the same outcome fee level those individuals with the lowest predicted probabilities of moving into sustained work *within* each payment group seem to be at considerable risk of being parked, while those individuals with the highest predicted probabilities seem most liable to be creamed. Qualitative evidence indicates that this is precisely the behaviour that can be observed on the part of providers, who ignore the official payment group structure and triage caseloads extensively according to detailed in-house profiling tools, targeting efforts on those participants deemed most likely to enter work (Meager et al., 2014; Rees et al., 2014). As indicated by this author previously (Carter and Whitworth, 2015), the Work Programme’s approach to profiling and differential pricing seems to be designing *in*, rather than designing *out*, provider incentives to cream and park.

The second implication is that with improved approaches to understanding the likely achievement of employment outcomes it *is* possible to embark on the design of an alternate system to better calibrate payment levels and facilitate differentiated

universalism. There is considerable interest within DWP for such progress and it is to the exploration of this possibility that the chapter turns in its final section.

5.10 An alternative payment system? – moving beyond crude payment groups

The key finding from the model development work and comparison of model performance outlined above – that utilising additional explanatory variables, even when these are a relatively constrained set of administratively held variables, facilitates an improvement in the prediction of job outcomes and a more successful ‘ordering’ of participants’ relative likelihoods of moving into sustained work – immediately implies that it is possible to do a better job of profiling and segmenting participants than has been seen with the Work Programme’s differential payment system.

The use of modelled participant predicted probabilities emerges as a key tool through which to facilitate differentiated universalism. Figure 5.16 visualises these predicted probabilities and offers an alternative payment structure grounded in these relative likelihoods of employment success. The left-hand portion of Figure 5.16 makes explicit the weaknesses with the current Work Programme payment group structure. The navy points plot the individual predicted probabilities of moving into sustained work from a random 1 percent sample of the full Work Programme administrative data, based on the full administrative data model outlined above. The payment groups (on the horizontal axis) are a poor basis for profiling and price setting given that virtually all of the variation shown is *within* rather than *between* payment groups. This context is ripe for creaming and parking by Work Programme providers and over-payment by the DWP who calculated outcome fees on the basis of ‘average’ participants rather than the easiest to help within each group.

One obvious way to facilitate differentiated universalism is to calibrate payment levels more closely with modelled likelihoods of moving into sustained employment (Carter and Whitworth, 2015). This approach conditions the value of job outcome payments on the individualised probabilities of achieving sustained employment outcomes and is demonstrated in the right-hand pane of Figure 5.16. Here individual participant predicted probabilities are used to segment the cohort into three provisional payment groups or programme streams. The left-most stream captures those participants with predicted probabilities more than 1 standard deviation below the mean predicted probability for the cohort, this group therefore includes individuals with markedly lower than average probabilities of moving into employment. To achieve differentiated universalism this is the group who would be expected to require more intensive and costly tailored support from providers in order to elevate their job outcome probabilities and under the revised payment proposal would therefore attract a ‘high fee’. Correspondingly, the middle stream contains those participants whose predicted probability of achieving a job outcome sits within 1 standard deviation of the mean. It is suggested that these participants with middling probabilities become associated with a mid-level job outcome fee. Finally, the right-most stream holds those participants

who have the highest probabilities of entering sustained work and who would therefore form a low-fee stream within the revised payment structure.

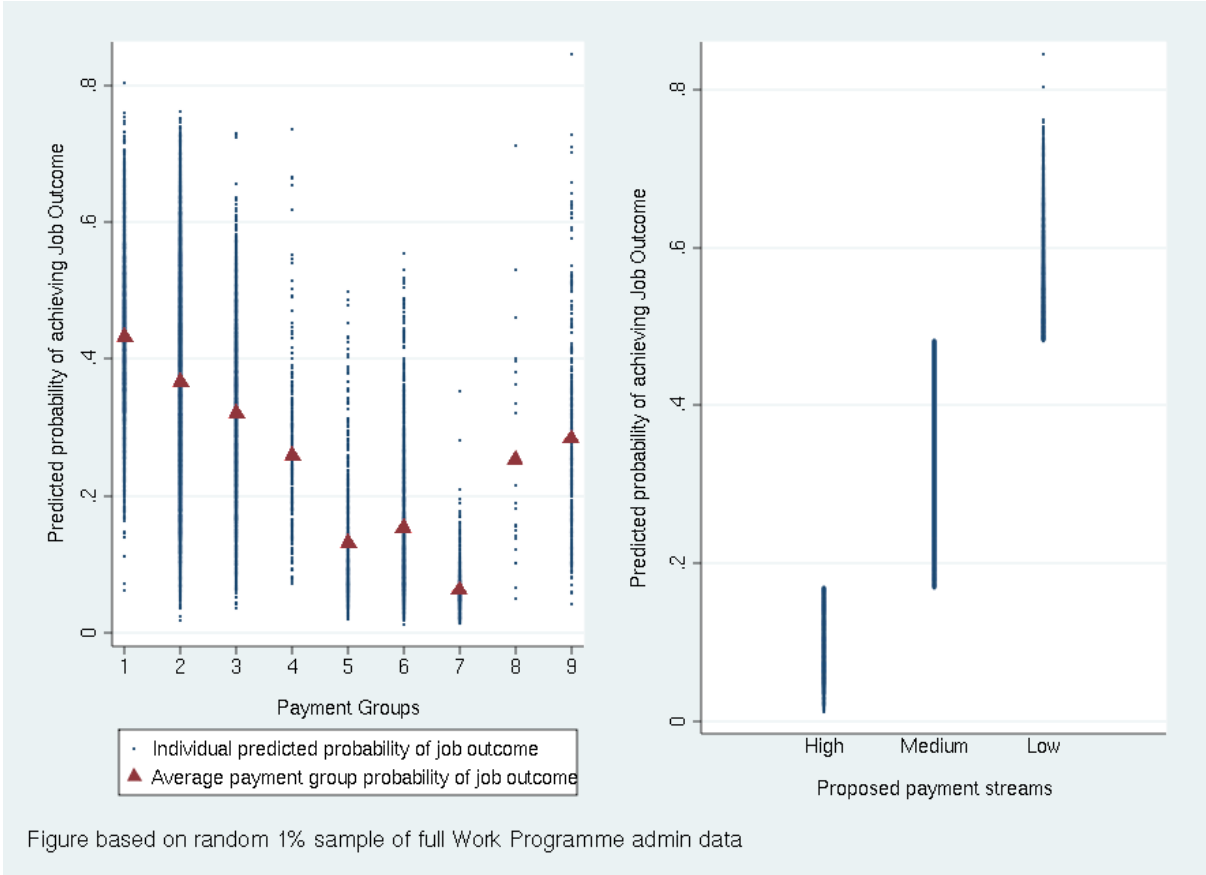


Figure 5.16 From predicted probabilities to an improved differential payments system. Source: DWP administrative data

It is possible to add a more granular scheme within this overall approach by using different thresholds to identify any number of distinct payment groups. Crucially, the design of this payment system seeks directly to act as a richer and stronger counterweight to known variation in the more richly evidenced likelihoods of participants entering paid work. Nevertheless, while this proposed revised payment structure is preferable to Work Programme’s official payment model in terms of its ability to deliver differentiated universalism, there are three important caveats.

Firstly, the participant predicted probabilities derived from the administrative data model, and hence, underpinning the revised payment levels, are far from perfect. No profiling approach will ever be perfect since there will always be unobservable as well as purely random aspects of whether and when a person will (re)enter employment. As argued by this author elsewhere (Carter and Whitworth, 2015), and demonstrated internationally (O’Connell et al., 2010), the key question is not whether richer statistically based approaches to profiling and payment model design are in an absolute sense perfect – they are not, and never can be – but instead whether they are in relative-terms better than alternative approaches and, as result, better able to help policy

makers achieve their desires around differentiated universalism through market-like means.

Secondly, where service providers are capable of more subtle profiling than commissioners – which is always viable given their ability to dedicate time in the early portion of the participant support process on more detailed data collection and triage than is typically practicable for commissioners – even a statistically enhanced payment structure will never fully eliminate risks of provider gaming. A more subtle statistically informed pricing structure will never on its own fully mitigate against risks around the delivery of differentiated universalism, and nor is it a reasonable expectation for it to be able to do so. Yet statistically enhanced approaches to profiling and segmentation can more effectively deliver a (partial) role in future UK provision to *better support* differentiated universalism, alongside other suitably designed accountability levers to buttress remaining (albeit lesser) vulnerabilities.

To improve the ability of a statistically-led profiling approach to facilitate equality in the achievement of job outcomes for participants a broader and richer set of explanatory variables will need to be collected for all unemployed programme participants, prior to enrolment. The detailed profiling approaches adopted by providers themselves – through for example, thorough questionnaires on employment history, skills, qualifications and caring commitments – offer clues for a more sensitive calibration of the payment system, as do the best practice international examples pursued in Ireland, Czech Republic, Kentucky and Australia.

Finally, the analysis outlined above is limited in that it does not explicitly set the precise cash values of job outcome payments within the revised payment system. The approach offered is a schematic through which to more appropriately set *relative* prices, but does not offer a methodology for allocating final values of payments. A simplistic approach might work from within the existing spending envelope and simply redistribute outcome payments in proportion to predicted probabilities. This simplistic linear inversion between probabilities and cash values for outcomes – while intuitive – seems unlikely to be the most appropriate function for translating predicted probabilities to outcome fees. Given the known high support costs for those individuals most distant from the labour market (Morphy et al., 2012) and known underperformance even within high fee paying groups would suggest that both higher overall outcome fees and a greater assurance (for example moving away from such heavy weighting of payment on outcomes and introducing alternative non-marketised accountability levers, as discussed in Chapter 8) may be required to move the British employment support system closer to differentiated universalism.

5.11 Chapter summary

Irrespective of its initial apparent subtlety compared to previous employment programmes, this chapter's original analyses – using an unprecedented empirical approach in terms of data richness and comprehensiveness – reveal that the Work Programme's differentiated payment model has failed woefully to deliver its

objectives around differentiated universalism. Instead of neutralising the effect of well-known indicators of labour market disadvantage, in the aftermath of the Work Programme's payment-by-results system there remain persistent, sizeable and significant effects remaining across a range of participant characteristics and contexts.

Advancing UK policy design around richer statistically informed profiling and payment design remains a priority in the coming decade. Further data than is currently available within DWP's administrative records will be a key part of this, as will enhanced design of other non-market accountability levers. Yet even within the constraints of current administrative data the enhanced analysis, considerations, and comparative assessment set out in this chapter indicate that it is viable for UK policy makers to do better and offers a proposed way forwards.

6 Jeopardising geography? The spatial patterning of Work Programme performance

6.1 Chapter summary

This chapter investigates the geography of incentives and performance in the Work Programme, assessing potential tensions between the spatially extensive Contract Package Areas (CPAs) and far smaller local labour market contexts within which providers and service users exist and act. The chapter asks: *Is the geographical container of regional Contract Package Areas an appropriate spatial scale at which to build key elements of welfare-to-work programmes including upholding minimum performance levels and incentivising [competition between] providers?*

The analysis investigates variation in Work Programme performance between different geographic areas and at different geographic scales. One of the key criticisms levelled at previous work-first employment support schemes is that they entrench pre-existing spatial inequalities (Peck, 1998; Turok and Webster, 1998; Sunley et al., 2006). Indeed, the title for this chapter is adapted from Turok and Webster's (1998, p. 309) paper which asks whether Labour's New Deal would be "jeopardised by the geography of unemployment?", a concern that turned out empirically to be well founded (Sunley et al., 2001, 2006). The analysis in this chapter therefore assesses the nature and extent of spatial variation in programme performance as well as the relationship between local labour market context (considered at a range of geographic scales) and programme performance. This is investigated using descriptive statistics, mapping, and regression based approaches to deliver a spatially sensitive 'contextual value added' performance metric for Prime contractor performance taking into account other factors.

6.2 Reprising the spatial logic of the Work Programme's private power market

Under the Work Programme's private power market the intention and belief on the part of policy-makers is that the competitive forces within quasi-markets will keep prices low, ensure that the best, most efficient practices are adopted and that innovation will emerge and be swiftly mainstreamed (OECD, 2014). An inevitable and persistent challenge within such marketised systems, however, is the effective management of principal-agent relationships (Bredgaard and Larsen, 2008; Struyven and Steurs, 2005). Increasingly, the challenge of steering provider behaviour has been pursued through the application of payment-by-results where providers are largely (and as per the later years of the Work Programme contract, *only*) paid following the delivery of specified, measurable social outcomes. This forms a conceptually and operationally straightforward backbone for programmes to direct and measure the performance and financial rewards of outsourced providers. Rhetorically the payment-by-results mechanism is frequently presented by politicians as a simplistic resolution to the principal-agent problem since "taxpayers only pay for results" (Duncan Smith in DWP, 2015).

Payment-by-results (particularly within a private power market) has been pursued more widely and positioned by UK government at the core of the public-service reform agenda (Cabinet Office, 2011). Central government protagonists have, however, explicitly noted that the approach must be pursued in a way that “align[s] incentives correctly between the provider and the public interest”, with a parallel acknowledgement that additional design complexity is required in order to prevent gaming (Cabinet Office, 2011, p. 33). As described in the previous chapters, the straightforward application of payment for results may not do justice to the full suite of programmatic objectives (as captured in Figure 3.7). Policy-makers, in their role as market-stewards, therefore recognise (to greater or lesser degrees) that it is necessary to add subtlety to the incentive structure through additional accountability levers and design tools. The implication then is that payment-by-results in a private power market is not a single mechanistic ‘fix’ which in and of itself will resolve issues of steering and incentives: there are other design prerogatives which need to be brought to bear in order to allow the market to operate without the commissioner being ‘played’ by providers.

As noted above, the cost-minimising strategies which are expected to be pursued by profit-seeking providers in private power market arrangement are linked to particular risks. A considerable body of international literature has consistently raised fears that in lightly regulated, outsourced payment-by-results schemes, providers will respond to financial pressures and incentives by ‘creaming’ off easier to help participants whilst ‘parking’ those who are harder to help and hence less likely to trigger paid for outcomes (Carter and Whitworth, 2015; Considine et al., 2011; Struyven and Steurs, 2005). Indeed, the challenge of unlocking the policy ambition of differentiated universalism sits at the heart of the preceding empirical chapter.

Of particular relevance to this chapter, these concerns around ‘gaming’ through creaming and parking processes have however been highly individualised. Critics have asked whether payment-by-results incentive structures are appropriately calibrated to ensure that appropriate support is provided regardless of the level, complexity and cost of an *individual’s* labour market barriers (WPSC, 2013; Davies and Raikes, 2014; Holmes, 2014). Calibration of incentives at the level of *individual* programme participants has been a central preoccupation, both in terms of programme design (with hubristic claims made for the differential payment structure from programme designers (Carter and Whitworth, 2015)) and from those scrutinising the programme (PAC, 2012; WPSC, 2013).

Considerably less well discussed however is the equivalent – but neglected – spatial dimension to these risks and needs. The geographical dimension of Work Programme functioning *has* received some limited coverage in think tank research (Davies and Raikes, 2014) but no detailed scrutiny from either government select committees or the academic community. The absence of geographical considerations in programme design and analysis is particularly surprising given the emphasis placed by academics on the need for a broader conception of employability that incorporates situation and

context (McQuaid et al., 2005; McQuaid and Lindsay, 2005) and the findings of previous research which suggest that the effectiveness of welfare-to-work policies – and in particular work first employment support programmes – differ significantly across local labour markets (Sunley et al., 2006; Theodore, 2007; Turok and Webster, 1998). This chapter responds to the overlooked spatial dimension in extant Work Programme analysis by considering the role of geography in mediating the effectiveness of programme incentives across space.

The remainder of the chapter proceeds firstly by introducing key design features of the programme including the ‘geographical’ structure in terms of the crude, large-scale spatial containers within which provision is organised. The empirical material then explores the extent of spatial variation in Work Programme performance for cohorts receiving 2-full years of service provision (for participants who joined the programme from June 2011 – June 2014 inclusive) considering local contextual factors associated with performance.

6.3 The geography of Work Programme incentives: contracts to tempt international providers versus meaningful local labour markets

Before turning to an analysis of the geographical variations in Work Programme performance it is appropriate to unpick the *potential* for (and protections against) unevenness in programme performance across space. Sunley et al. (2006) suggest that the geographical outcomes of activation programmes can be understood as being driven by two sets of processes. Firstly, outcomes reflect programme governance: the ways in which central and local agencies interact and allocate resources, functions and responsibilities. The second set of processes is informed by “the interactions between the intended and unintended effects of the programme and local labour market conditions, which shape local outcomes” (Sunley et al., 2006, p. 19). This two part framing informs the discussion below, which begins with the geography of programme governance *on paper* whilst the empirics seek to reveal *in practice* the ways in which the anticipated spatial neutrality interacts with spatially uneven local labour market realities.

The Work Programme operates from within an a-spatial understanding of the unemployment challenge. Its design is rooted in supply-side orthodoxy which presents unemployment as the result of individual’s ‘employability’ deficit (Crisp and Powell, 2017). The onus of worklessness is placed squarely upon individuals and has been repeatedly framed by politicians as a cultural deficit (Grover, 2007; Slater, 2012; Wiggan, 2012) despite detailed evidence rebutting the presence of ‘dependency culture’ (Harkness et al., 2012; Macdonald et al., 2014). Structural aspects of persistent unemployment and poverty are marginalised in favour of a behavioural explanation (Slater, 2012; Wiggan, 2012).

Despite the Work Programme’s a-spatial understandings of unemployment the governance of the programme *does* have a geography, of sorts: within the programme Britain is divided into 18 regional Contract Package Areas (CPAs) at which level

contracts are awarded to Prime providers. The CPAs are geographically large: all of Scotland sits as a single CPA for example (see Figure 6.1).

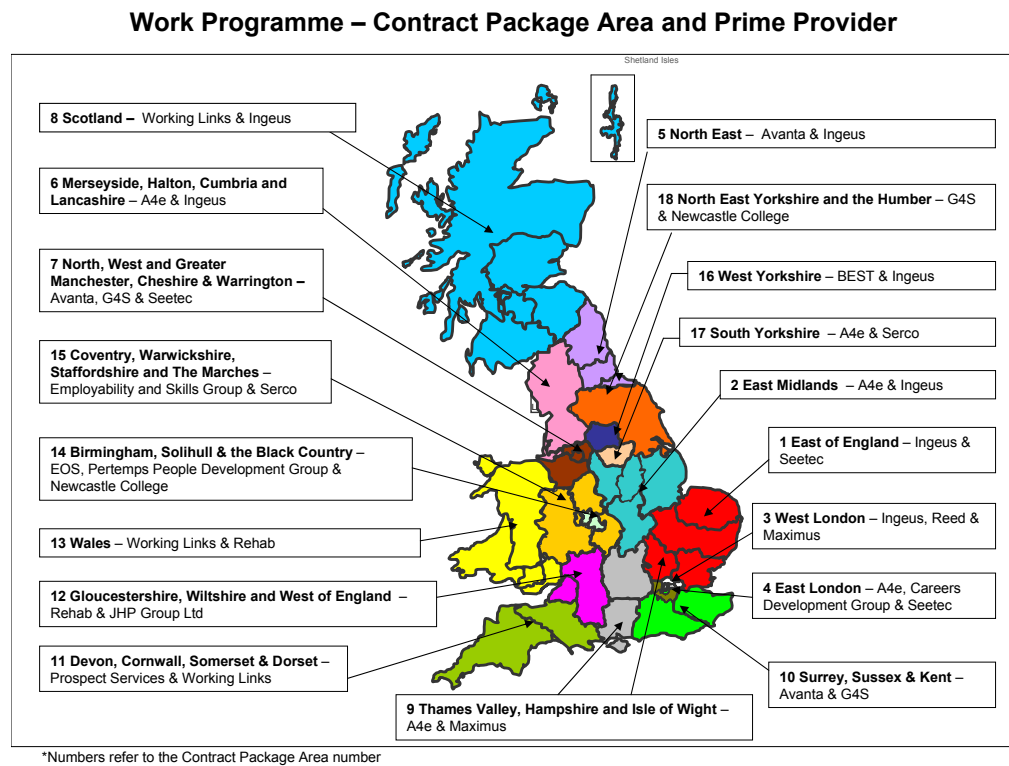


Figure 6.1 Work Programme Contract Package Areas and Prime providers.

Source:

http://webarchive.nationalarchives.gov.uk/20130507013604/https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/186922/cpa-preferred-bidders.pdf

These CPAs represent the scale at which the tender contest for providers originally took place with two or three Prime providers contracted to operate within each CPA. For Freud – as noted above, a leading figure in UK activation redesign over the past decade due principally to his authoring of the key Freud Report (2007) and role from 2010 as Minister for Welfare Reform – the large geographic extent of these contracts was intended to offer the large financial scale “appropriate to attract major players from around the world” who would have sufficient scale to arrange private finance and tap into investment from the banking community (2007, p. 63).

The justification for these large scale contracts is fully aligned with the underlying programmatic logic of the private power market that to unlock stronger innovation and higher performance levels than has been seen in previous schemes the programme must unleash the entrepreneurialism of private providers. Through the Prime provider model the DWP sought to reduce the administrative complexity and costs of its contractual arrangements (in 2008 the department had 1,419 welfare-to-work contracts with over 580 contractors (WPSC, 2009)). The Department considered that the risks and cash flow implications inherent within its preferred *outcome-based* contracts

would best be carried by large ‘top-tier’ providers who would then manage “healthy, high performing supply chains” and where longer contracts would offer sufficient length (subject to performance) to enable long-term planning and investment (DWP, 2008, p. 20). Crucially, this large regional geography was designed with an eye to attracting capitalised providers rather than because CPAs are the relevant, internally coherent and meaningful economic geography of Britain.

At the level of countries and CPA regions the Work Programme bidding process built in the potential for Prime providers to anticipate *cross-regional* variation in the ease of achieving employment outcomes. Primes were able to offer discounts on the value of job outcome payments in their bid proposals. The understanding was that Primes would offer more sizeable discounts in regions with more buoyant labour markets where it was expected to be cheaper (in aggregate) to support programme participants into work (Holmes, 2014). Importantly, this discounting occurred at the level of the CPA only and there is no variation in pricing below these extensive contract areas.

Arguably, on paper the differential payment structure exerted across individual programme participants (the payment groups discussed in the preceding chapter) may accommodate some variation in the *compositional* characteristics of people who are long-term unemployed from place to place. Crucially, however, and as evidenced in Chapter 5, the differential payment system is known to be highly imperfect as a tool for calibrating appropriate support for participants with more complex or compound support needs. Further, the payment-by-results system takes no account of *sub*-CPA contextual demand-side factors. Spatial variability and parking are in this context just as much a risk – even if a largely neglected risk in current policy analysis – as individual variation and parking. Indeed, the interaction of ‘hard to help’ people’ with ‘hard to help’ places represents a potential ‘double whammy’ of currently unexplored risk in such activation programmes.

Looking back across the academic discussion surrounding the rise of outsourced welfare-to-work interventions first articulated in the 1990s it is perhaps surprising that the geography of labour markets and the spatial variation of programme successes have not been more fully considered in the latest wave of programme design. Peck, largely informed by the US experience (though helpfully so, given the rate of trans-Atlantic policy transfer that has occurred), concluded as early as 1998 that the successes of and indeed variability of work-first activation models is “predicated on a very particular set of local conditions” (Peck, 1998, p. 553). A comprehensive UK investigation into the geography of New Labour’s New Deal for Young People similarly found that the performance for this suite of interventions varied significantly between different local labour market areas across Britain and in particular had been markedly less effective in inner-urban and depressed industrial labour markets (Sunley et al., 2001, 2006). A key conclusion from this work as well as official policy evaluations (Beale et al., 2008) is that the issue of *local* labour demand – both in terms of the volume and range of jobs available – is vital within such schemes. Absence of demand locally cannot be assumed away as self-correcting.

Given this previous experience of spatial variation in performance, both internationally and in the British context, it is helpful to reflect on the ways that local labour market conditions were discussed by officials and ministers in the early phases of Work Programme design, commissioning and delivery. Post-2010 there has been a partial acknowledgement in the UK policy context that there *is* sub-regional variation in local labour market conditions and that these may intersect with welfare-to-work programming: “typically every single one of these regions [CPAs] has more variation in unemployment inside it than there is between regions” (Devereux, 2012 in PAC, 2012: Ev19). Yet whilst variation across labour markets is acknowledged, there is no explicit attempt to calibrate pricing or performance metrics and processes to reflect the relative differences in difficulty of transitioning to employment in different local contexts, nor are any other accountability mechanisms brought to bear on mitigating the potential for sub-regional performance variation.

The absence of design features to account for sub-CPA labour market variation has been justified along two lines. Firstly, Robert Devereux, Permanent Secretary for the DWP (from January 2011 – January 2018) has stressed that: “...everybody [Primes] is having to manage areas that have got hotspots and very difficult labour markets” (Devereux, 2012 in PAC, 2012: Ev19). In essence, CPAs – at an aggregate level – are seen by DWP as ‘equal’ in the sense that each contains some difficult labour markets and, therefore, on the basis of CPA-level performance monitoring sub-regional variation is of little or no consequence. Secondly, and contrastingly, a former DWP minister has argued that locally differential payments would bring excessive complexity to the system (Grayling, 2011 in WPSC, 2011). Under this view the spatial variation in labour market context does not warrant the additional complexity that a local-corrective was perceived to entail, though clearly the variations across individuals were deemed worthy of response via the differential payments model.

Large regional CPAs therefore continue as the only contractually relevant geography to Primes since they are the scale at which performance levels are calculated and any breach of minimum performance levels is judged. It is at the CPA level that the on-going ‘competition’ between primes through market share shift to higher-performing providers takes place (Lane et al., 2013). Crucially, it is only at this aggregate CPA scale that performance ‘matters’ in the sense that it is only at the CPA level that performance is assessed by DWP and has contractual and financial ramifications.

Variation in performance levels *within* CPAs is not part of formal performance or contractual monitoring. There is no flexibility to adjust pricing to accommodate changing economic conditions or local labour market factors *within* the contracting areas. In sum, while Work Programme is nationally uniform in terms of participant eligibility, referral mechanisms and payment model there is no explicit commitment to horizontal spatial equity. In the programmatic intention to narrow the gap “in employment rates for disadvantaged groups and everyone else” (DWP, 2010a, p. 4) the experience of ‘disadvantage’ is never fully described nor explicitly acknowledged

as an experience that might vary over space. Spatial variation may however implicitly and insidiously undercut the ambition to close the gaps in employment rates. Sub-CPA spatial variation in performance is effectively ‘invisible’ – and designed to be invisible – within contractual arrangements, despite a considerable body of evidence indicating that local context matters in the achievement of job outcomes. Any patterning of performance below the CPA level is unlikely to be detected given that it is not being looked for and there are neither financial sticks nor carrots through which to mitigate such spatialised variation in performance where it occurs. Spatial variation and spatialised parking emerge as significant risks in this policy context.

In the Work Programme’s private power market, by specifying only outcomes to be achieved government commissioners absolve themselves from designing the programme’s response to need and there is no assurance that wherever a person resides uniform standards of treatment will be received. Proponents of payment-by-results may argue that this is a benefit of the approach and that locally specific interventions, tailored appropriately to respond to local context, is enabled. This flexibility is undoubtedly unlocked. But the more pertinent question is whether the failure to include geographical considerations in programme design enables (and potentially encourages) spatially variegated risks and programme performance.

The programme logics on paper may then interact in anticipated (for academics and commentators) or unanticipated (for the political sponsors of Work Programme) ways with local labour market conditions to shape outcomes. There is a clear risk that the failure to calibrate for locally varied labour market contexts will work against the programmatic ambition to reduce *gaps in performance outcomes between the easier- and harder-to-help* neighbourhoods and areas.

Not only is this spatial unevenness in outcomes of concern for those interested in differentiated universalism and equity. Spatially varying programme performance is also relevant for the Work Programme’s internal logic of ‘value for money’ (Carter and Whitworth, 2015). If the programme is experiencing variation in the spatial distribution of ‘results’ such that a greater volume of employment outcomes are being triggered and paid for in ‘easier’ local contexts then the commissioner must be systematically overpaying for these outcomes, since payment levels relate to *average* costs and outcome likelihoods. If a process of procyclical spatial performance variation is occurring then this has the potential to undermine the logic of the payment-by-results mechanism itself.

6.4 From paper to practice: is there spatial variation in performance?

If the Work Programme’s design had worked in accordance with the presumed regional labour market functioning of the scheme designers, it would be expected that the programme would show comparable levels of job outcome success in the CPAs across Britain. The aggregate job outcome rates for the four largest (by way of participant numbers) payment groups 1, 2, 3 and 6 are represented with the green triangle markers in Figure 6.2 for each of the 18 CPAs.

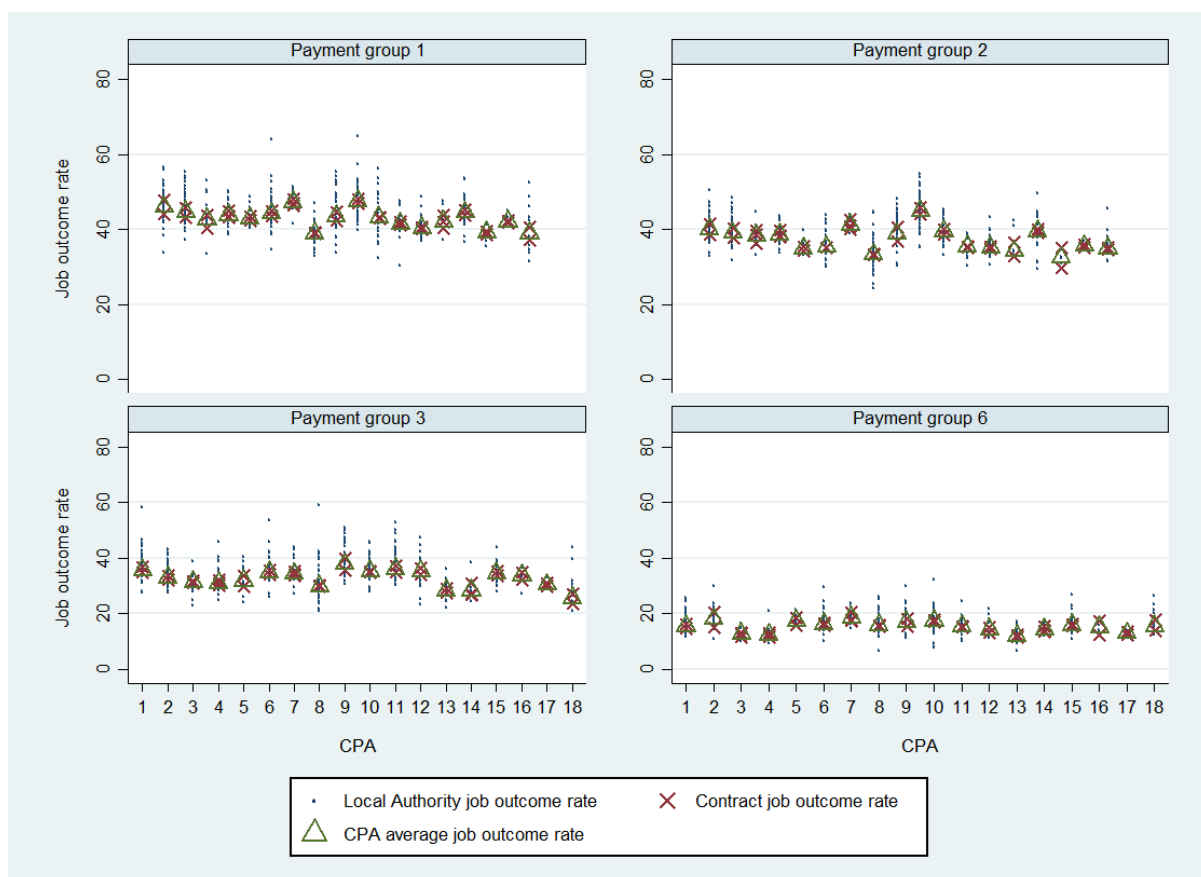


Figure 6.2 Job outcome rates within alternative spatial scales: CPA, contract and local authority

In practice then there is some variation across the regions although the standard deviation in CPA-level job outcome rate is never larger than 3.3 for any of the payment groups (Figure 6.3). There is however a notable gap between the better and worse performing contract package areas. Thames Valley and Hampshire (CPA 9), Surrey, Sussex and Kent (CPA 10) and Greater Manchester (CPA 6) are consistently amongst the better performing contract areas. By contrast, the worst performing CPAs are typically Scotland, Wales and parts of Northern England (North East Yorkshire and Humber (CPA 18) and West Yorkshire (CPA 16)). There is less variation across the CPAs in the performance of the health-related ESA payment group 6 compared to the JSA-related payment groups.

	Mean CPA job outcome rate	Standard deviation	Minimum CPA rate	Maximum CPA rate	Percentage point gap (max-min)
Payment group 1	42.98	2.57	39.02	47.58	8.57
Payment group 2	37.29	3.22	32.43	44.90	12.47
Payment group 3	32.61	3.28	25.43	37.84	12.41
Payment group 6	15.32	1.97	12.00	18.65	6.65

Figure 6.3 CPA-level job outcome rates, descriptive statistics (N = 18)

Even at this regional scale, therefore, spatial variation exists and the Work Programme does not fully break out of the persistent and entrenched regional spatial economic imbalance of the UK, where London and its surrounds have been dominant since the middle of the nineteenth century (Crafts, 2005; Gardiner et al., 2013).

Within each of the CPAs the Prime’s contract level job outcome rate is marked by the red Xs (Figure 6.2). There are statistically significant differences between Prime contractors operating in the same CPA (Dorsett and Lucchino, 2016). What is notable here is the substantively small variation in performance between the contracts inside each CPA. Figure 6.4 provides detail on the performance gap between the best and worst performing contracts inside each CPA by payment group. On average, within each contract there is only a 2 percentage point gap in job outcome rate between the best and worst performing Prime contract. This suggests that when judged at the level of CPA there is little differentiation between the performance of alternative Prime providers.

	Mean contract performance gap within CPA	Minimum gap in contract performance	Maximum gap in contract performance
Payment group 1	1.71	0.16	3.62
Payment group 2	1.92	0.01	5.43
Payment group 3	2.16	0.51	4.48
Payment group 6	2.07	0.48	5.46

Figure 6.4 Job outcome performance gap between contracts within each CPA

The contract performance figures presented here incorporate the variation in participant volumes across providers driven by the market share shift mechanism (which took effect from 5 August 2013, (DWP, 2017b)) which, if anything, would be expected to amplify the performance gap between ‘better’ and ‘worse’ performing providers. The DWP is artificially animating market competition according to the variation in contract performance within CPAs, which is actually smaller than the variation across CPAs at a national level (i.e. comparing variation in red Xs *within* a CPA to the divergence in CPA aggregate performance across Britain).

Figure 6.2 also plots the job outcome rates at local authority level within each of the CPAs, shown by the navy dot markers. Two local authorities: Isles of Scilly and City of London are removed from all of the analysis shown in this chapter due to very low cell counts (fewer than 10 participants in one or more payment groups). What is striking about the 378 local authority job outcome rates is the high degree of performance variation both nationally and *within* each of the CPAs. The percentage point gap between the best and worst performing local authorities nationally is 25.6 for payment group 6 but is over 30 for the other core payment groups (see Figure 6.5).

	Mean local authority job outcome rate	Standard deviation	Minimum local authority rate	Maximum local authority rate	Percentage point gap (max-min)
Payment group 1	44.51	5.23	30.32	64.89	34.57
Payment group 2	39.07	5.00	24.15	54.84	30.69
Payment group 3	35.16	6.26	20.82	58.82	38.00
Payment group 6	16.36	3.98	6.61	32.17	25.56

Figure 6.5 Description of Local Authority job outcome performance variation (N = 378)

Hence, spatial variation in job outcome performance rates is markedly greater at the individual local authority level than at CPA or contract scale. The standard deviation in job outcome rate for payment group 1 and 6 at local authority level is more than double that of the standard deviation across CPAs. And each of the CPAs contains enormous variation in local authority job outcome rates.

Importantly, not only is the key payment-by-results model a-spatial but the contractual Minimum Performance Levels (MPLs) and other performance metrics and processes for prime providers are specified at CPA level. These performance metrics are contractually specified for payment groups 1, 2 and 6 whilst performance expectations are specified through illustrative indicators for other payment groups. These MPLs set the number of job outcomes a provider must achieve in a year, as a proportion of the number of referrals the provider receives in that year (DWP, 2010a; Lester, 2013). The calculation of MPLs has been highly criticised (UK Statistics Authority, 2013; CESI, 2013; OECD, 2014) particularly as the metric is not comparable over time and since the numerator and denominator are not temporally aligned. Changes in referral patterns can impact the annual MPL (a pattern of declining referrals will make performance appear better, irrespective of ‘actual’ performance) (CESI, 2013; Lester, 2013). The DWP has adjusted the metrics used for performance monitoring over time and increasingly utilises a ‘transparency indicator’ which calculates the proportion of individuals in each monthly intake to reach three/six months in work during their first year (DWP, n.d.). Nevertheless, minimum expected performance levels for providers are still measured and monitored at the level of the contract and function as an indicator of unacceptably low performance in the DWP’s eyes. The DWP expects providers to significantly exceed these minimum levels and failure to achieve them is punishable by contractual action, including contract termination (DWP, 2010).

The Centre for Economic and Social Inclusion have produced minimum benchmark job outcome measures equivalent to the MPLs and these are reproduced in Figure 6.6. This provides an overall job outcome measure, capturing the average proportion of participants achieving a job outcome, for all those who have completed the programme

and so this is equivalent to the job outcome rates produced by the author across the thesis.

	Overall job outcome measure (%), equivalent minimum benchmark
PG 1 - JSA 18 to 24	38
PG 2 - JSA 25 and over	29
PG 3 - JSA early entrant	19
PG 6 - New ESA claimants	18

Figure 6.6 Equivalent minimum performance benchmarks, reproduced from CESI, 2013

Within payment groups 1, 2 and 6 there are a number of local authorities which are breaching the minimum benchmarks equivalent to the MPLs. Importantly, for payment groups 1, 2 and 3 at CPA level (and contract level) none of the Prime contracts have performance that sits below these minimum thresholds. For payment group 6 there are contracts which flout this baseline minima at CPA level and indeed the poor performance of payment group 6 has been a considerable concern throughout live running (Riley et al., 2014; Davies and Raikes, 2014; Quirke, 2015; WPSC, 2015). What is more striking is that for payment groups 1 and 2, and inevitably for payment group 6, many local authorities are experiencing job outcome rates *below* these minimum benchmarks. There are a considerable number of local authorities which, were contractual performance to be judged at this scale, would be breaching the specified contractual minima and where Primes would therefore be subject to targeted performance management intervention from DWP and under threat of contract termination. Importantly, however, because the contractual agreements are sited at the level of CPAs these local authorities with unacceptably low job outcome performance are contractually invisible since they are diluted, concealed and deemed irrelevant within the CPA aggregate performance figures.

6.5 Understanding contract performance variation or routes to ‘winning’ within the Work Programme incentive structure

As noted above, in the Work Programme key aspects of the incentive structure, both carrots (winning market share shift) and sticks (breach of contractual MPLs), are sited at the level of CPA. The risk and danger is that in conjunction with the considerable flexibilities afforded by the private power market Primes can be judged as market leaders irrespective of their spatial behaviours and outcomes, despite programmatic commitments to narrow gaps.

Reflecting specifically on the ways in which Primes might achieve high *aggregate* job outcome performance at CPA level, this might be achieved by enhancing job outcome rates in what are already better performing local areas, enhancing performance in more challenging local labour market contexts or, indeed, spatially parking those more challenging local labour market contexts and focusing scarce resources on ‘easier’

local labour market areas. All might in principle lead to upwards shifts in aggregate CPA performance, but all are clearly not equal in terms of their spatial (and, within these spaces, individual) impacts.

Figure 6.7 to Figure 6.10 plot the relationship between aggregate contract performance rates and three within-CPA performance metrics: firstly, the percentage point performance gap between the best and worst performing local authorities within the contract; secondly, the job outcome rate in the best performing local authority (i.e. performance at the ‘top end’); and, thirdly, the job outcome rate in the worst performing local authority within the contract (low end performance).

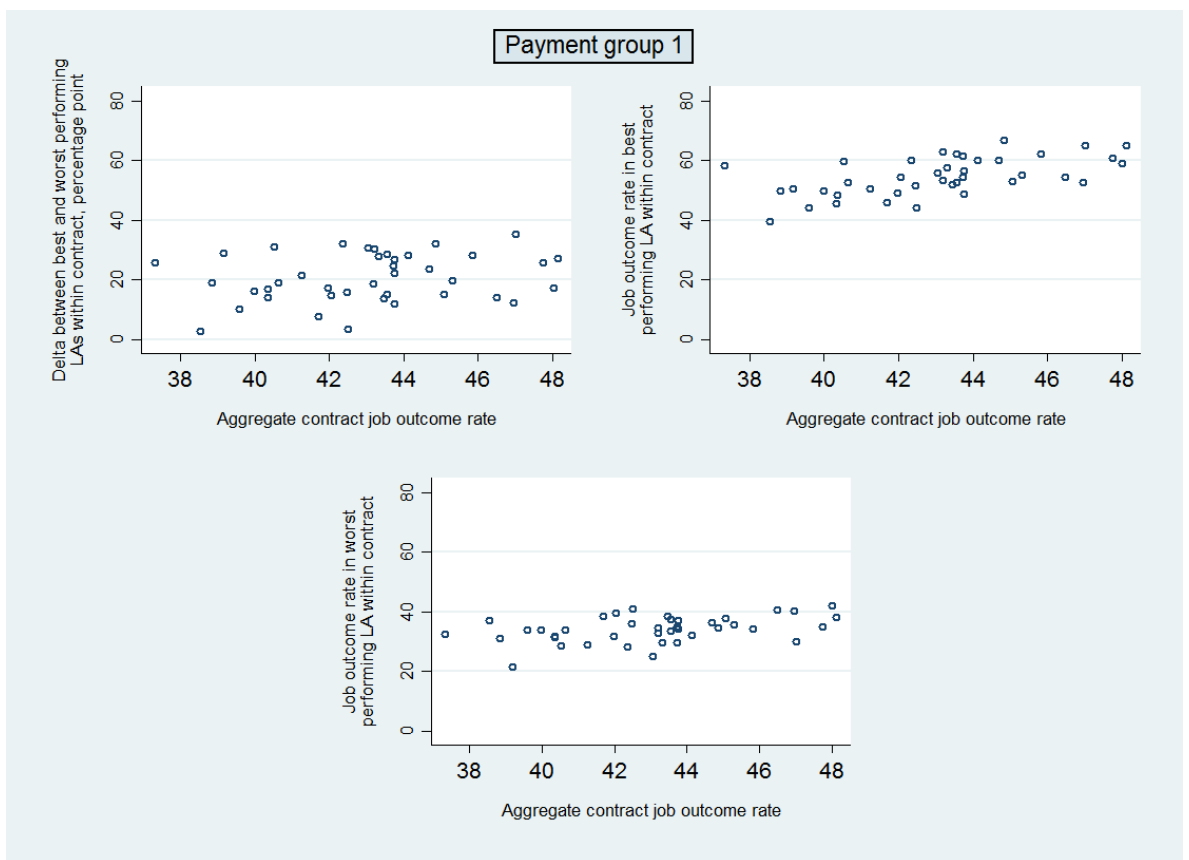


Figure 6.7 Within Prime contract local area performance variation for payment group 1

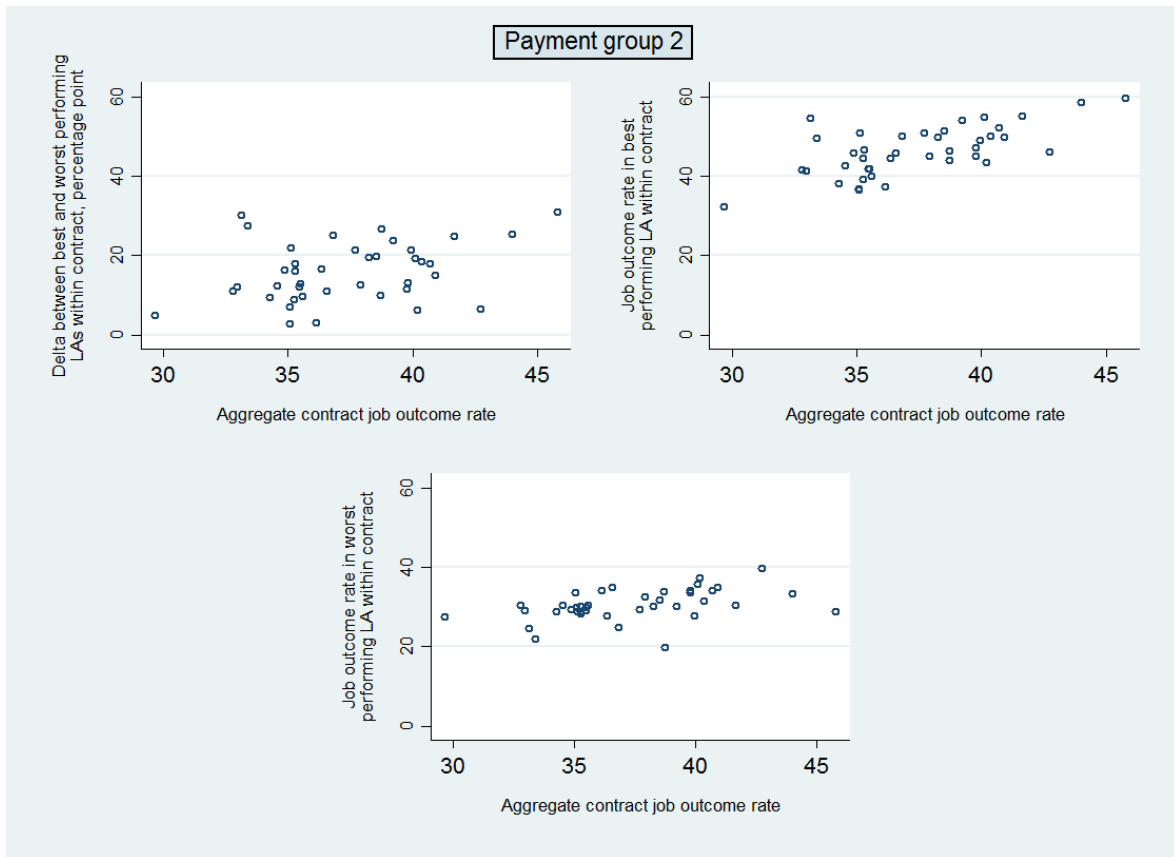


Figure 6.8 Within Prime contract local area performance variation for payment group 2

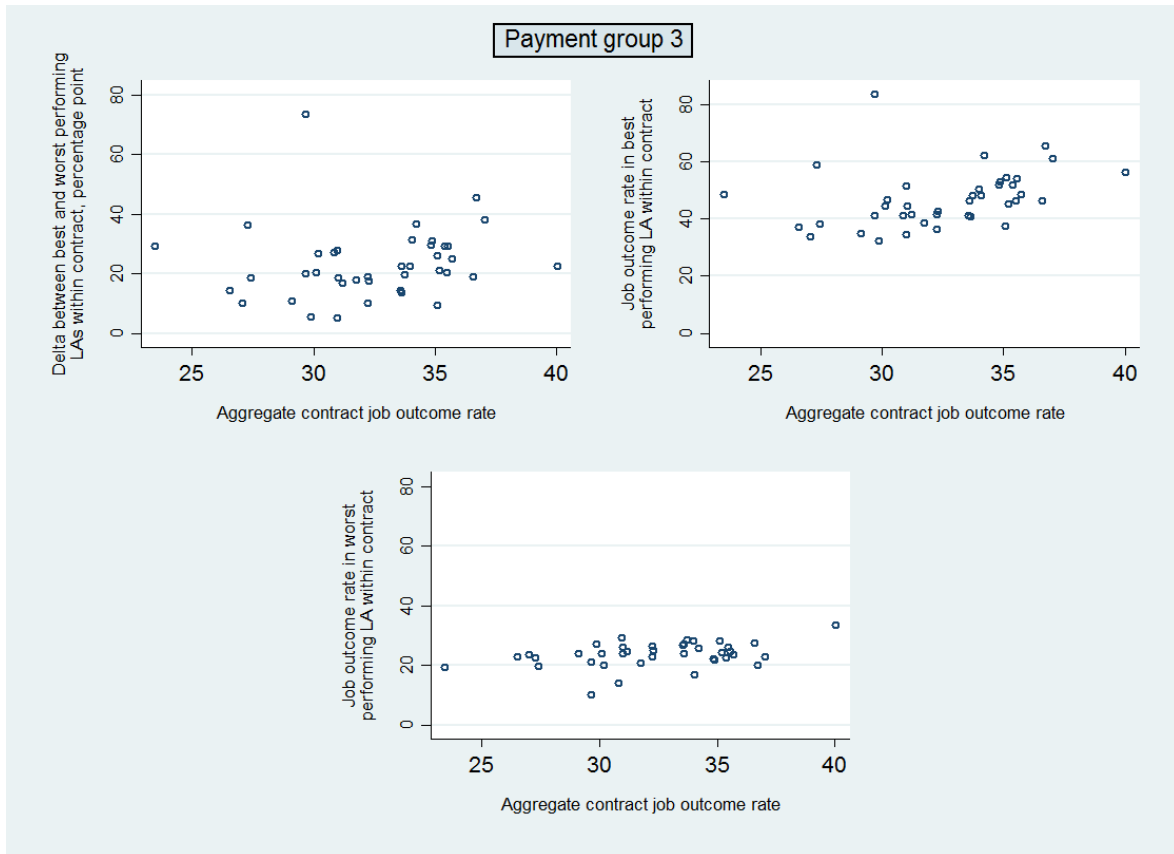


Figure 6.9 Within Prime contract local area performance variation for payment group 3

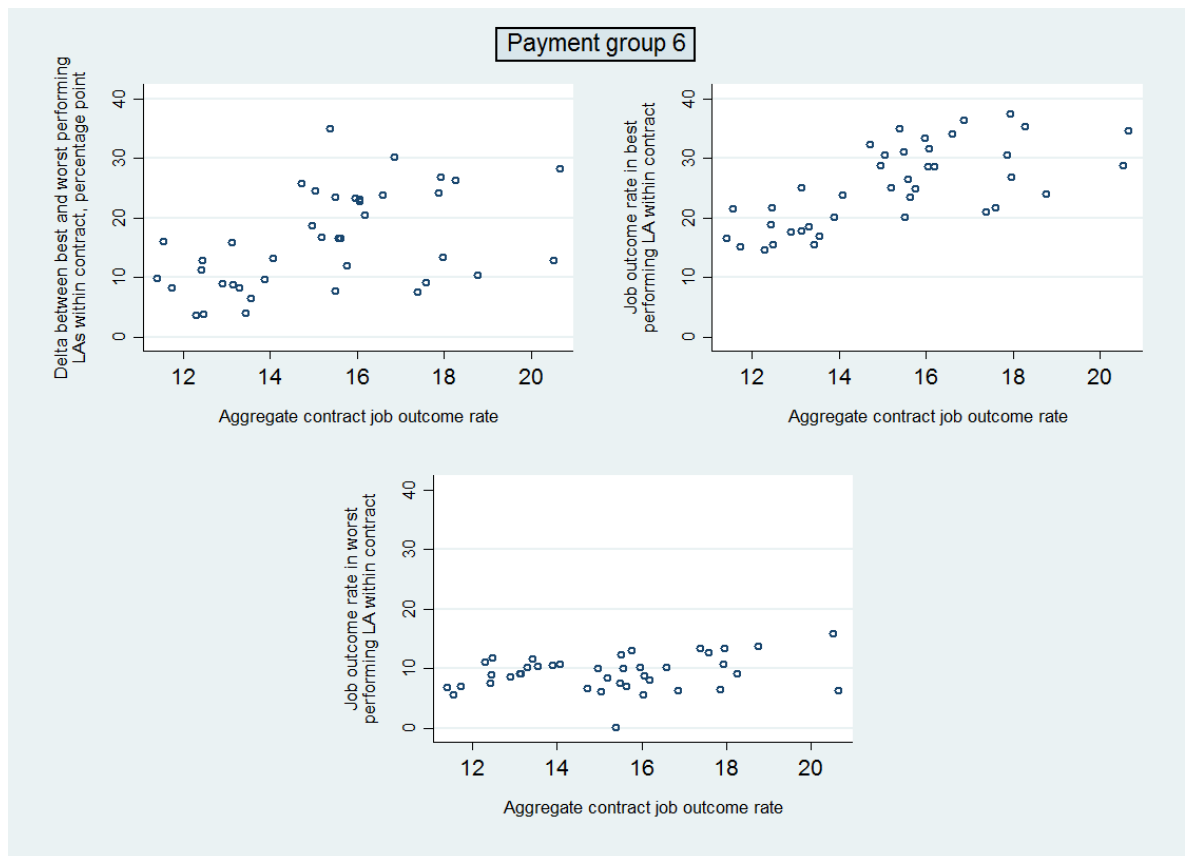


Figure 6.10 Within Prime contract local area performance variation for payment group 6

Across none of the payment groups is there a particularly clear relationship between aggregate contract performance and the gap between better- and worse-performing local authorities. There is a slight upward trend across each of the contract-level indicators. When tested as explanatory variables the relationships are not sufficient to suggest that improved aggregate performance at contract level has been achieved in conjunction with widening spatial inequality. However, the absence of clear associations does suggest that it is possible to ‘win’ and be a better performing Prime contract at CPA level without closing the gap in performance across local authority areas and without driving better performance in less well performing areas.

These figures, in conjunction with further statistical tests for the strength of association (not shown), indicate that contract allocation is not a strong predictor of local area performance. The following section begins to unpick these issues further by investigating other factors which might be associated with spatial variation in Work Programme performance.

6.6 Predictable patterns of success? Mapping Work Programme performance at local authority level

Beyond the wide variation in local authority job outcome performance, for those familiar with the economic geography of Britain the performance follows a clear spatial pattern. For reference a map indicating local authority deprivation levels across

Britain is shown in Figure 6.11 where lower levels of deprivation are indicated with paler shading.

Figure 6.12 to Figure 6.15 map the job outcome rate for each of the core payment groups at local authority level across Great Britain. In each map, local authority job outcome performance is shaded according to performance quintiles. The darker shading indicates those with lower performance (with the darkest shading representing the 20 percent of local authorities experiencing the worst job outcome performance rates) and lighter shading signifying those areas with above average performance (and with white shading representing the 20 percent of local authorities with the highest job outcome rates).

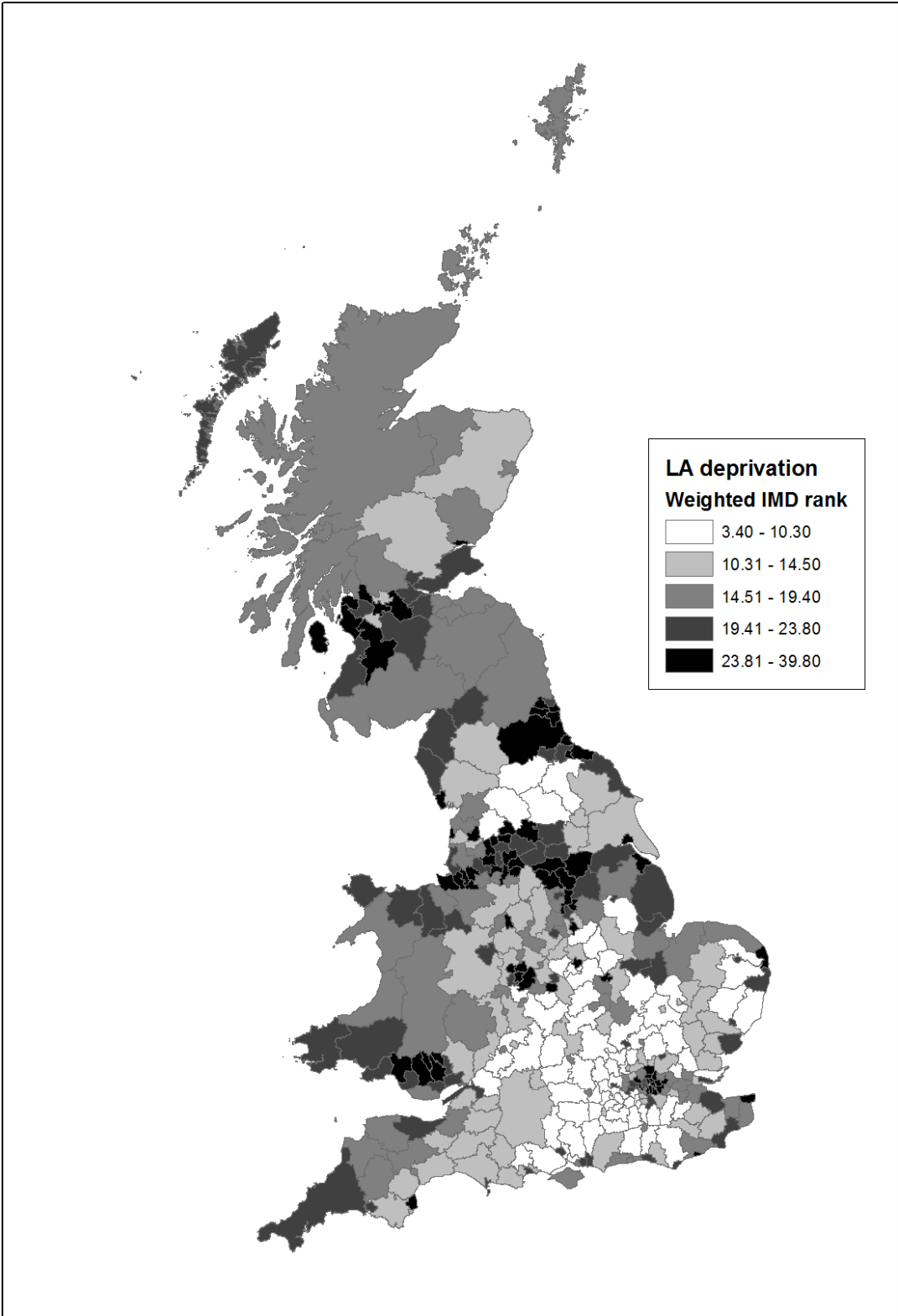


Figure 6.11 Local authority deprivation across Britain, standardised indicator sourced from CRESR and produced as described in Abel et al., (2016)

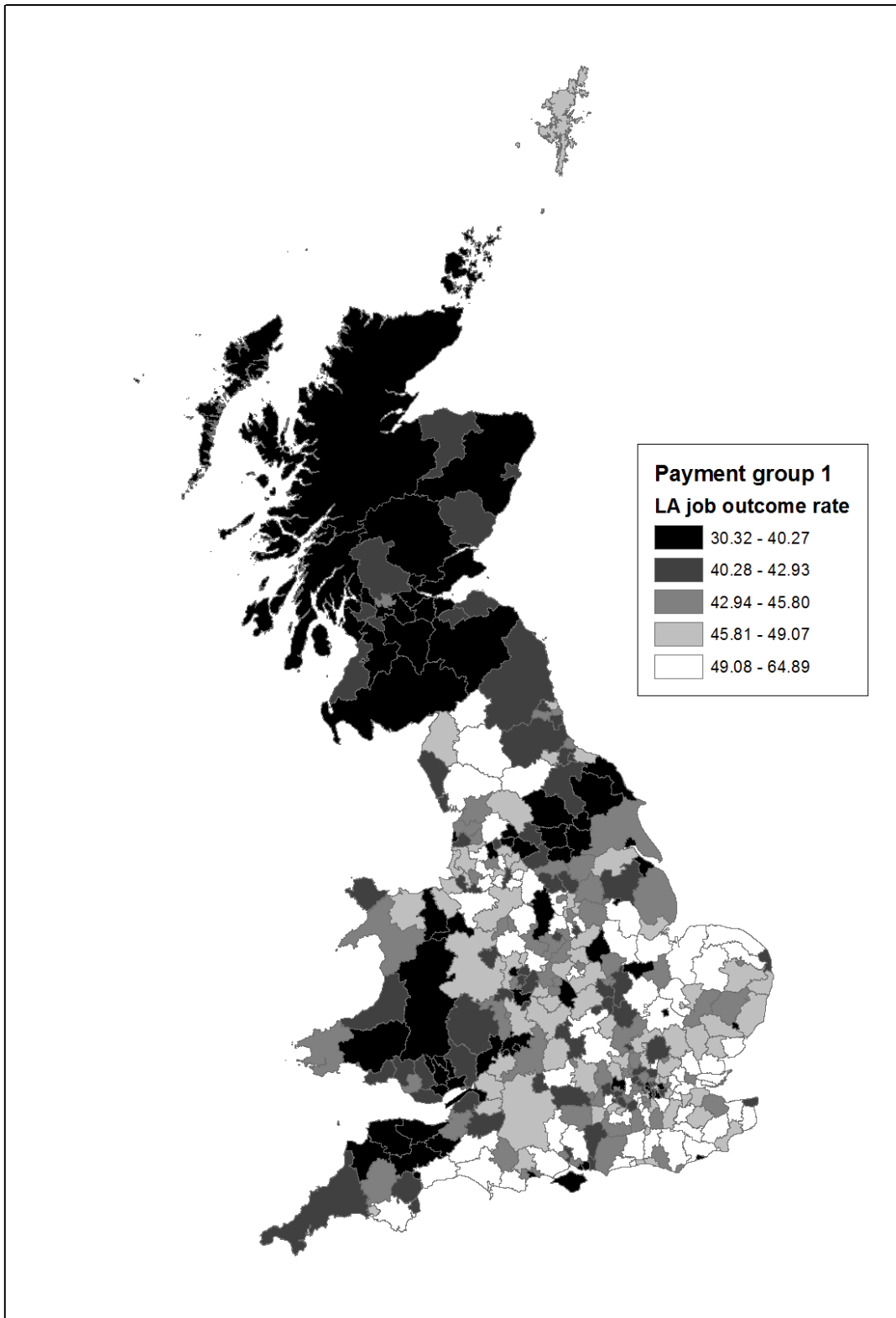


Figure 6.12 Proportion of payment group 1 participants achieving job outcomes at local authority level

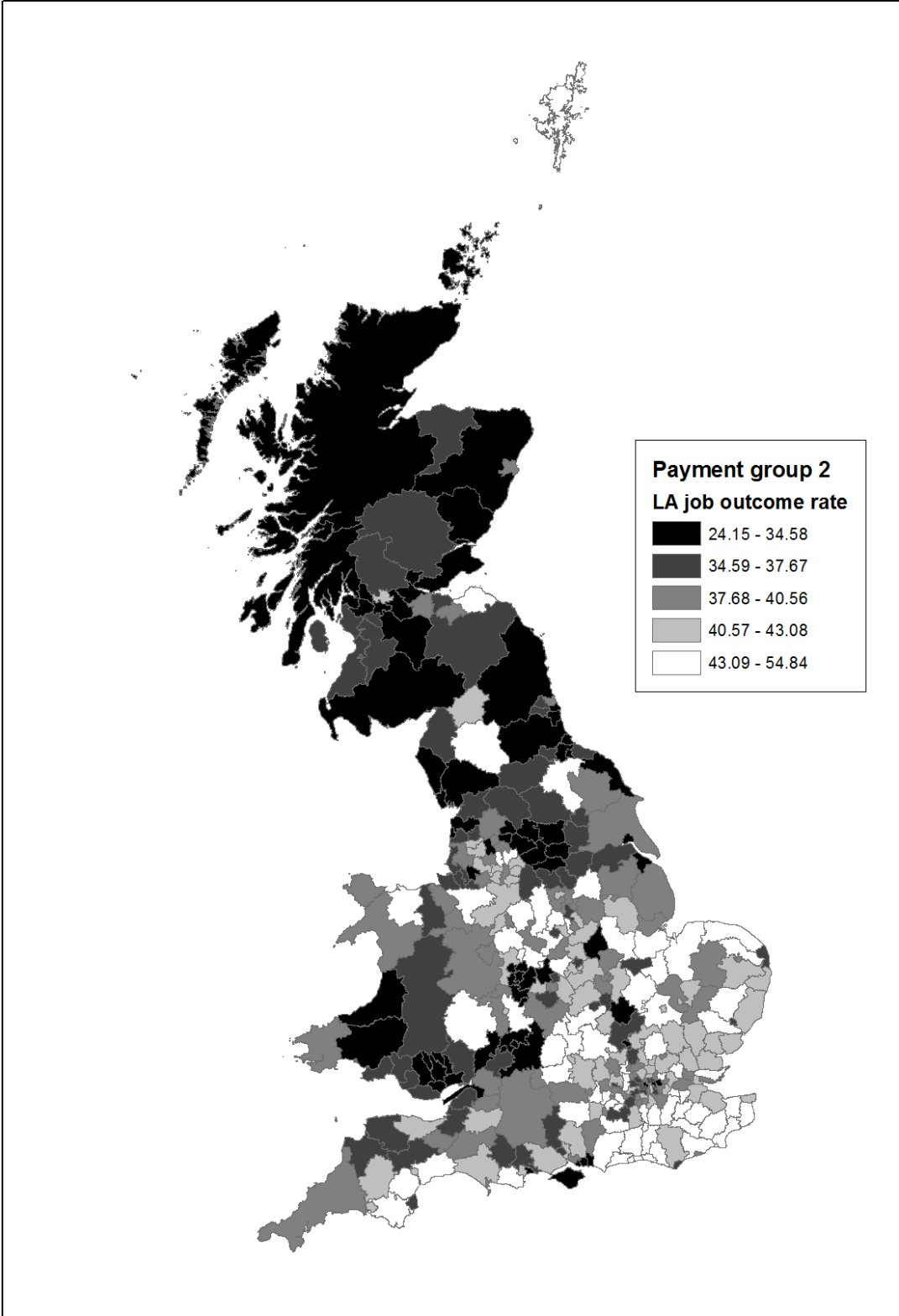


Figure 6.13 Proportion of payment group 2 participants achieving job outcomes at local authority level

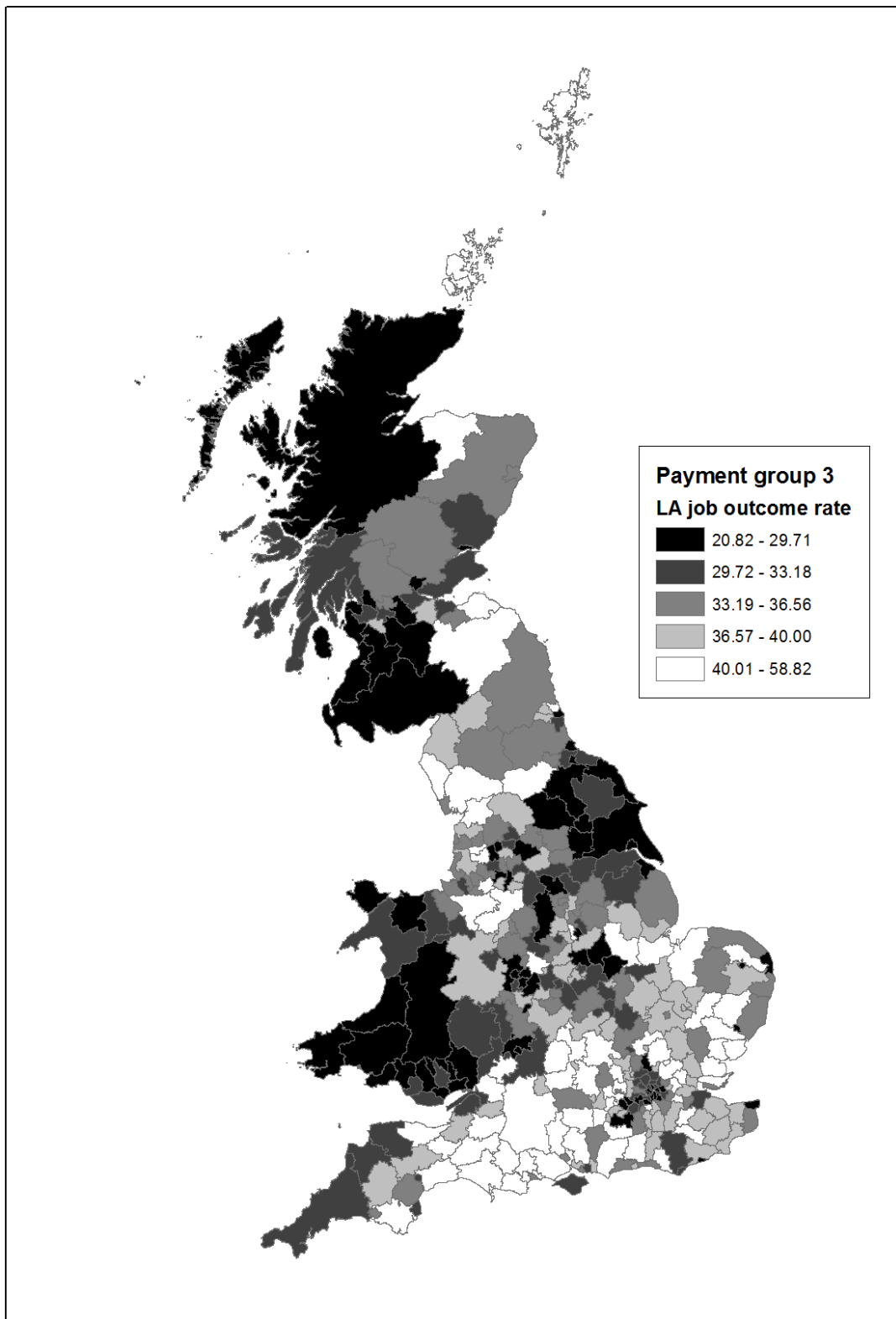


Figure 6.14 Proportion of payment group 3 participants achieving job outcomes at local authority level

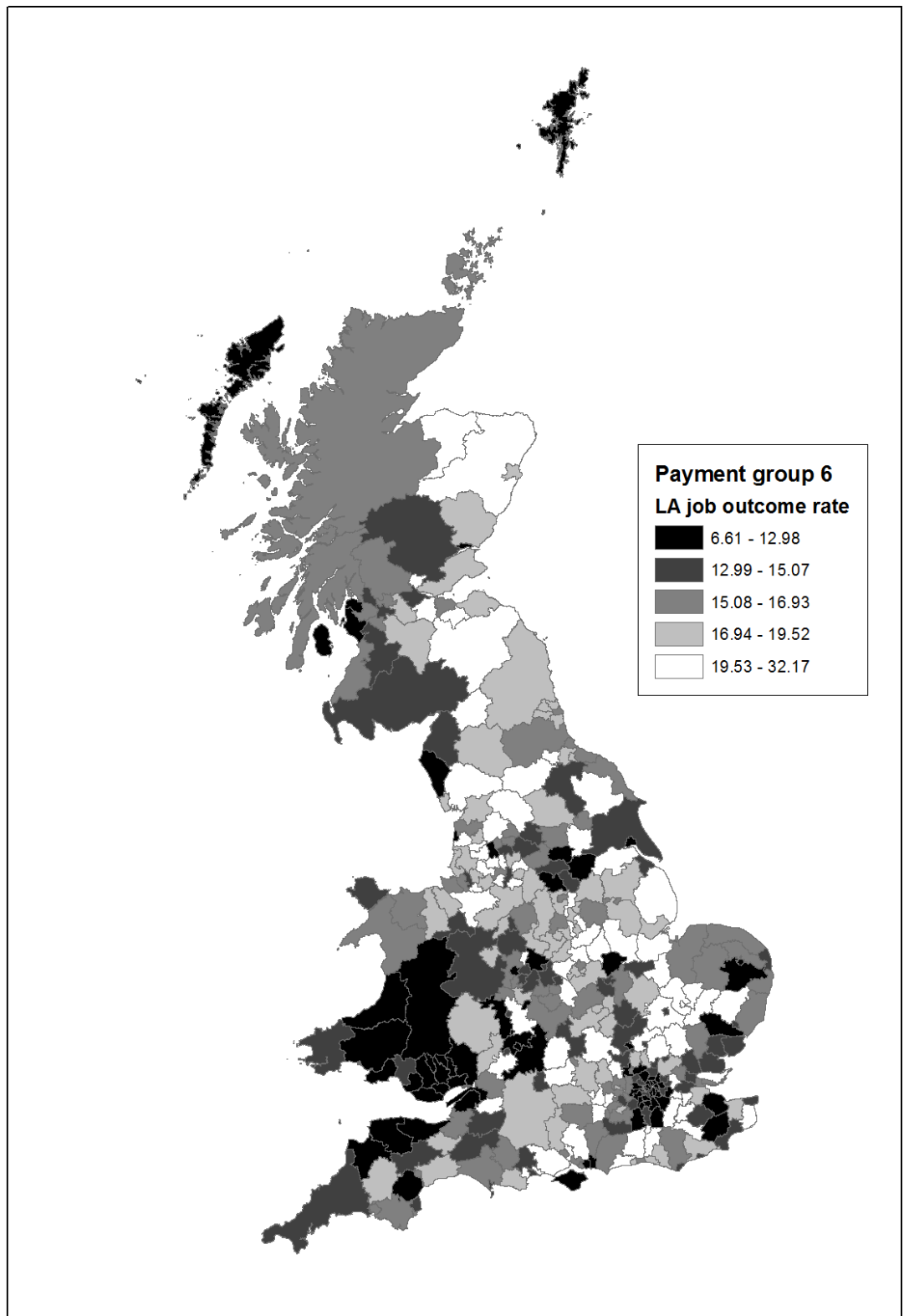


Figure 6.15 Proportion of payment group 6 participants achieving job outcomes at local authority level

Across each of the maps the places with the weakest job outcome performance are predominantly found in the de-industrialising areas of Northern England, Scotland and

the Welsh Valleys. Seaside towns and to a lesser extent some inner city areas also appear to be performing poorly. Contrastingly, a large band in the South East surrounding London contains almost all of the best performing authority areas. Whilst by no means perfectly aligned, the overlaps between locality deprivation and Work programme performance are evident visually.

6.7 Making sense of the variation

In order to promote a further understanding of this spatial patterning a series of potential independent variables were considered for inclusion in regression analysis, the intention being to identify those contextual variables that would contribute to the successful prediction of job outcome performance levels for local authorities.

Figure 6.16 and Figure 6.17 show the relationship between local authority job outcome performance and two key local authority explanatory variables. The first of these, shown in the upper panels, is the relative deprivation of the local authority. This is captured through the population weighted index of multiple deprivation scores for component small areas, standardised across Britain by Abel et al. (2016). Higher values indicate authorities with a higher incidence of local area deprivation. The second local authority variable explored is the percentage of the local authority working age population who have been referred to participate in the Work Programme (shown in the lower panes). Since programme eligibility is framed around long-term unemployment this proportion can be understood as capturing the incidence of long-term unemployment and hence the intensity of ‘need’ for the programme within the authority area. The working age population figure is calculated using the mean ONS mid-year population estimate for 2011 – 2014 sourced through NOMIS.

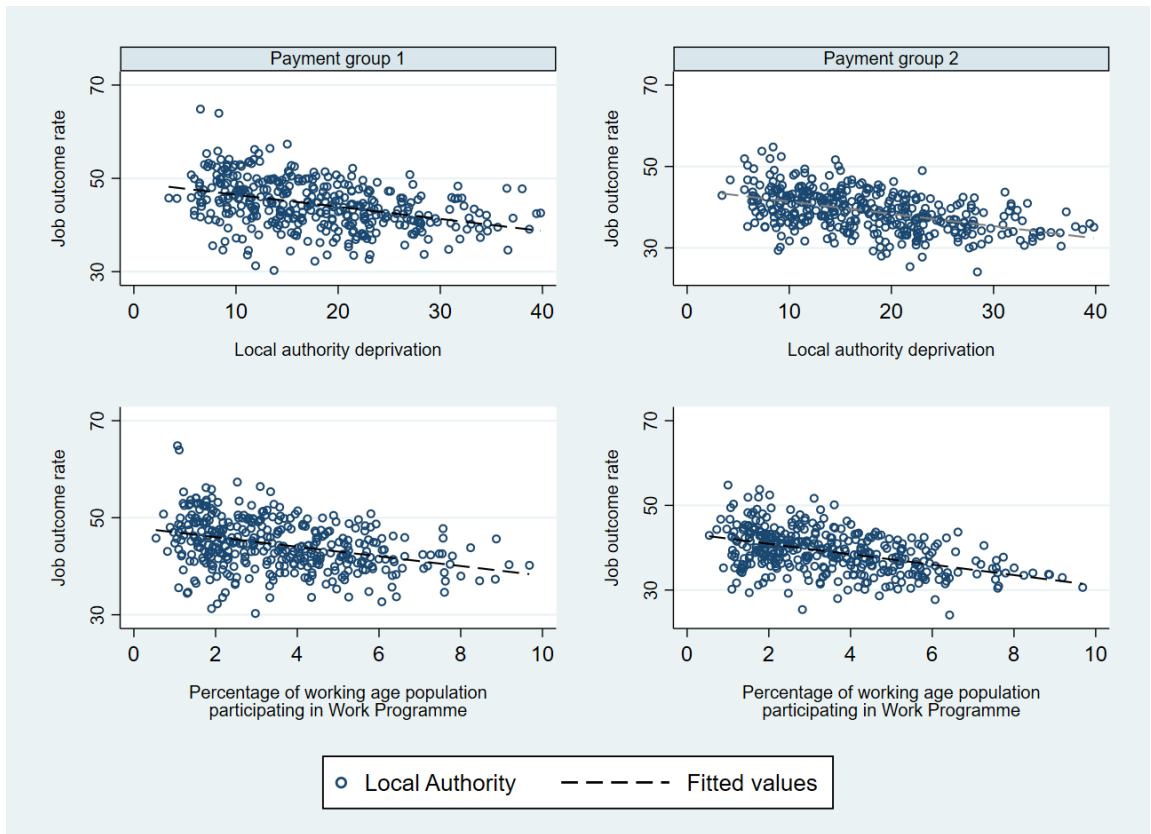


Figure 6.16 Association between job outcome rates and i) local authority deprivation; and ii) proportion of working age population participating in the Work Programme, for payment groups 1 and 2

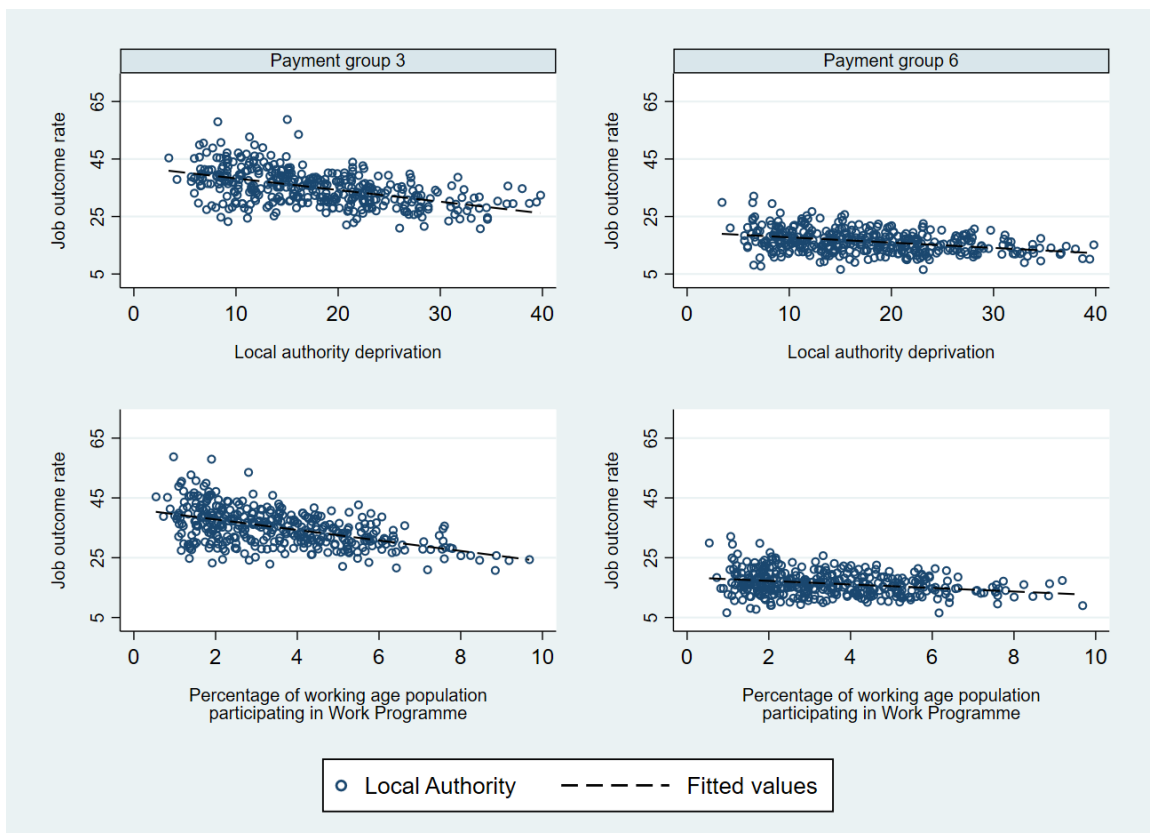


Figure 6.17 Association between job outcome rates and i) local authority deprivation; and ii) proportion of working age population participating in the Work Programme, for payment groups 3 and 6

The negative relationships between the indicators of local authority deprivation and need and programme performance visualised in Figure 6.16 and Figure 6.17 are confirmed using simple bivariate regression analysis (Figure 6.18). The predictive power of local context is stronger for the mainstream JSA groups (for payment group 2 the indicators separately account for over one-fifth of the variation in performance ($R^2 = 0.219$ and 0.202)) and local area context appears to be a less powerful predictor for the performance of payment group 6 (here the $R^2 = 0.128$ and 0.074), though the predictive power is understandably weaker given that there is less variation in the performance of payment group 6 which is almost universally poor.

The significance and gradient of the slopes in Figure 6.16 and Figure 6.17 is captured through the bivariate regression table in Figure 6.18 where the coefficient for each of the models gives a sense of the likely change in job outcome rates as we move across local authorities with different levels of deprivation and long term unemployment rates. For mainstream JSA participants in payment group 2, a single percentage point increase in the proportion of the local working age population participating in Work Programme is on average associated with a 1.2 percentage point reduction in the job outcome rate. Similar relationships hold for the other JSA payment groups 1 (JSA 18 – 24) and 3 (JSA early access), where a one percentage point increase in participation rates is associated with an average of -0.994 (payment group 1) or -1.753 (payment group 3) percentage point drop in job outcome performance. The gradient for the new Employment and Support Allowance participants in payment group 6, i.e. those who are newly in receipt of health and disability related unemployment benefits, is markedly shallower at -0.591 . For this group the gap in job outcome rates between local authorities with a low-incidence of long-term unemployment and those with a higher long-term unemployment rate is less sizeable.

The indicator for local authority deprivation, the weighted average adjusted IMD score, ranges from 3.4 (least deprived) to 39.8 (for the most deprived local authority). A one unit increase in this deprivation measure is consistently associated with a statistically significant reduction in local authority job outcome rates across all the payment groups with relatively similar effect sizes between roughly -0.2 and -0.4 percentage point reduction in job outcome rates.

**Local authority deprivation Proportion of working age
population referred to
programme**

	<i>Payment group 1</i>			
<i>B-coefficient</i>	-0.259	***	-0.994	***
<i>Constant</i>	49.084	***	47.994	***
<i>Observations</i>	378		378	
<i>R-squared</i>	0.147		0.121	
	<i>Payment group 2</i>			
<i>B-coefficient</i>	-0.302	***	-1.230	***
<i>Constant</i>	44.394	***	43.378	***
<i>Observations</i>	378		378	
<i>R-squared</i>	0.219		0.202	
	<i>Payment group 3</i>			
<i>B-coefficient</i>	-0.404	***	-1.753	***
<i>Constant</i>	42.297	***	41.296	***
<i>Observations</i>	378		378	
<i>R-squared</i>	0.250		0.262	
	<i>Payment group 6</i>			
<i>B-coefficient</i>	-0.183	***	-0.591	***
<i>Constant</i>	19.595	***	18.429	***
<i>Observations</i>	378		378	
<i>R-squared</i>	0.128		0.074	

Figure 6.18 Exploring local authority Work Programme performance. Outcome variable is job outcome rate ***
 $p < 0.001$

Taken together, on average the more challenging a local area is in terms of long-term unemployment and deprivation the lower the Work Programme job outcome rate. This is a counter-intuitive investment pattern for a programme seeking to best support those furthest from the labour market.

The implications of this patterning are spelled out in Figure 6.19 which uses the programmes own payment-by-results logic to investigate performance patterns. There is no official single metric which comprehensively captures programme performance. An all-inclusive metric would need to convey the ‘value’ of outcomes across a series of programme priorities capturing not only job outcomes but also the duration of sustainment and moreover valuing job outcomes differently according to individuals’ relative distance to the labour market. Ultimately the programme’s payment-by-results structure itself provides a key framework through which to trace this complex mix of performance priorities. Guided by the incentive structure a single measure traces the *maximum potential extent of payment* (i.e. should every programme participant in an area achieve full payable job and sustainment outcomes according to their payment group and sustainment payment cap). Following the logic of payment-by-results,

'better' performance ought to be equated with a higher level of disbursement against the potential payment pot. 'Proportion of potential PbR achieved' then conveys the ratio of achieved (and paid for) outcomes to the maximum potential within local authorities and Figure 6.19 shows the relationship between this aggregate metric of success and the two indicators of local need and deprivation.

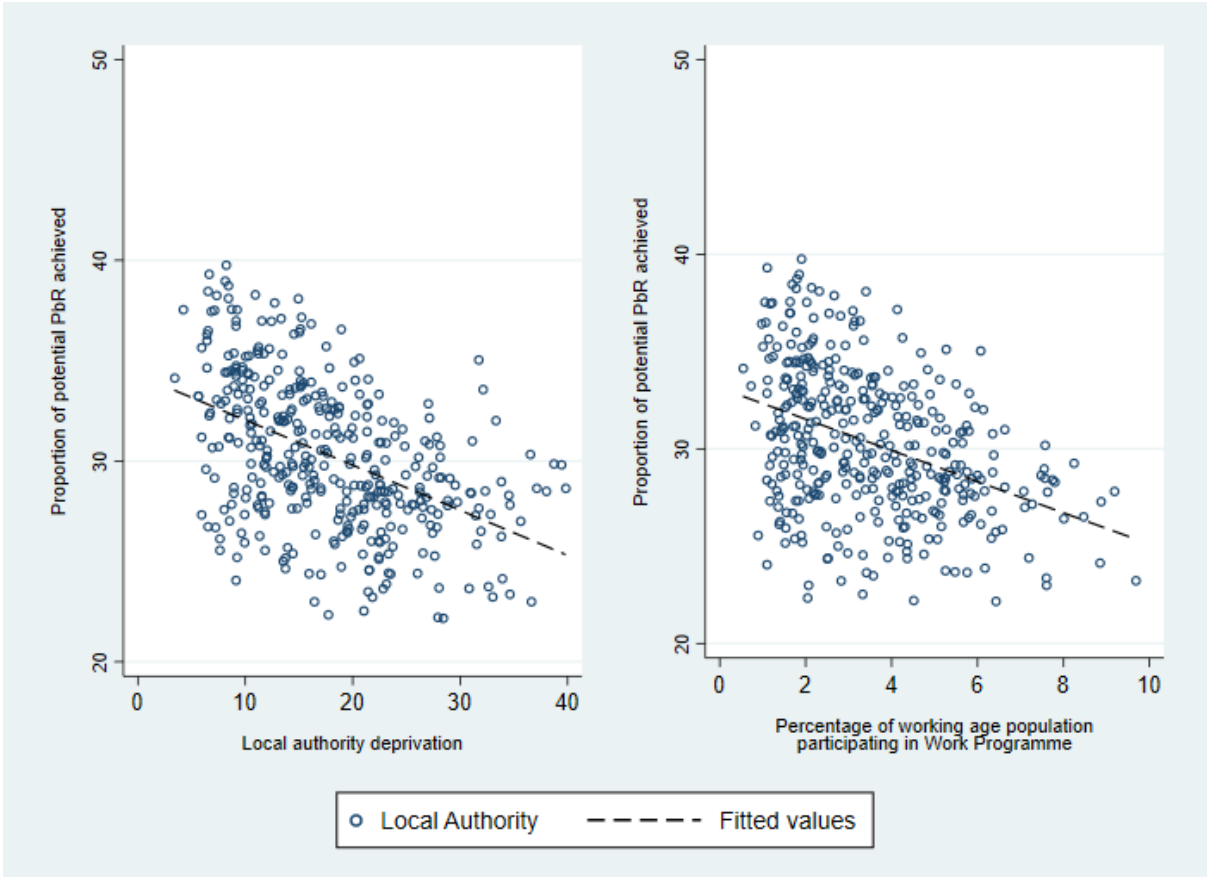


Figure 6.19 Variation in local authority payment-by-results 'success' by deprivation and proportion of programme participants amongst the working age population

The negative direction, relatively large effect size and high statistical significance of these effects shown in Figure 6.20 is particularly concerning. Aggregate programme performance is shown to be systematically weaker in those areas in which programme successes are most needed, be that expressed through high levels of long-term unemployment (right hand pane Figure 6.19) or greater extent of deprivation within the local area (left hand pane). Instead, as the need for Work Programme success to narrow spatial gaps increases the performance of the programme worsens: a spatially perverse result contrary to programmatic commitments and re-inscribing pre-existing patterns of geographical local labour market disadvantage.

	Local deprivation	authority ***	Proportion of working age population referred to programme	***
<i>B-coefficient</i>	-0.225	***	-0.802	***
<i>Constant</i>	34.287	***	33.132	***
<i>Observations</i>	378		378	
<i>R-squared</i>	0.220		0.156	

Figure 6.20 Exploring variation in local authority payment-by-results 'success' by deprivation and proportion of programme participants amongst the working age population. Outcome variable is proportion of potential PbR achieved

6.8 Putting people in their place

Knowing that the Work Programme's payment groups are poor proxies for person-level characteristics (Chapter 5), a potential criticism of the analysis presented above is that the work may be conflating individual and area-level effects and overstating the latter.

Area-level or 'neighbourhood' effects are notoriously difficult to identify as they are located and constituted amongst a number of circuitous and interrelated multi-level social processes (Buck, 2001). Indeed, questions have been raised as to whether it is appropriate to seek to partition the "false dualism" of context and composition since there is likely a mutually reinforcing relationship between people and place (Cummins et al., 2007, p. 1835). Nevertheless there seems little doubt that neighbourhood effects exist (Ellen and Turner, 1997; van Ham et al., 2012b; Dorling, 2001).

In order to 'put people in their place', Figure 6.21 gives an indication of the degree to which local context may matter independently from and in addition to person-level characteristics. This analysis reproduces the model used to predict job outcome likelihood developed in the preceding chapter but uses only characteristics sited at the level of the person and household (blocks of variables related to benefit history; employment history; health and disability; household and caring obligations). These are effectively 'individual level' predicted probabilities. Bands of people with the same individual level probabilities (i.e. the ten percent of participants estimated to be the most-likely to enter sustained employment given their personal characteristics, the middle ten percent, and bottom ten percent) are then identified and their job outcome rates are produced across deciles of local area unemployment benefit claimant rate. For example, the first blue bar on the left indicates the job outcome rate for individuals within the top ten percent of person-level predicted probabilities and who live in the ten percent of local authority areas with the lowest unemployment rates (captured by claimant rates).

In Figure 6.21 the downward gradient is clear for all levels of individual-level predicted probabilities. This is as would be expected from the previous regression analysis where explanatory variables capturing local area context retain significance when included alongside individual-level characteristics. Individual-level characteristics and experiences emerge as likely stronger predictors of job outcome success but place clearly ‘matters’. What can also be seen is that ‘place matters’ to a different extent depending on where individuals sit within the distribution of person-level predicted probabilities. There seems to be a non-linear interaction effect at play: local area context matters less for individuals with low personal predicted probabilities than for those whose individual-level characteristics make them likely candidates for (re)entering and sustaining paid work. Even amongst participants whose individual characteristics suggest they are close to work there is not a clear linear decay in job outcome likelihoods across local area context.

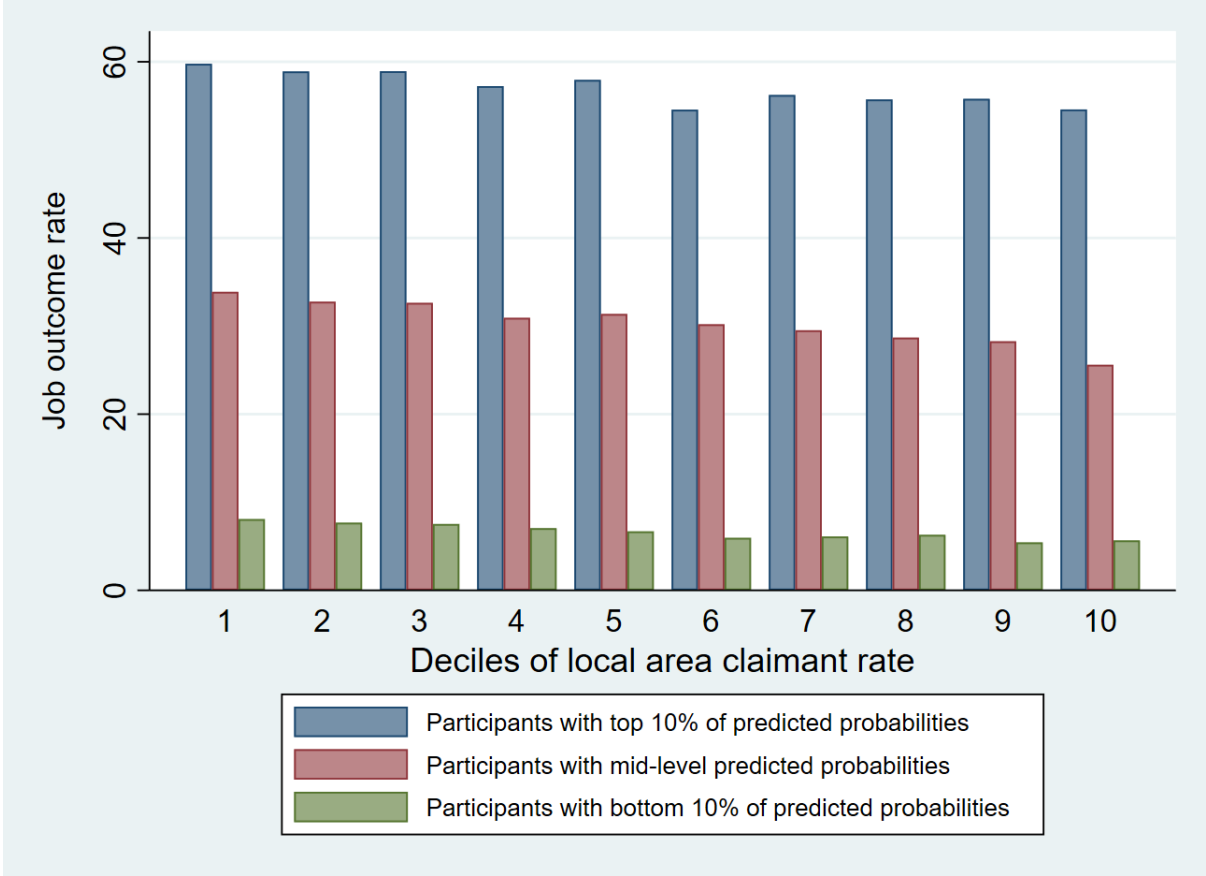


Figure 6.21 Plotting job outcome rates for 'people in place'

6.9 Discussion: controlling for composition and context

The flexibilities of the private power market, in conjunction with the general absence of contextual considerations in the design of key programme steering mechanisms, has left the Work Programme susceptible to spatial variation in performance. In practice there is highly variegated programme job outcome performance across Britain. The Work Programme’s geography of incentives is structured around two levels – at the level of the individual (since the bulk of payments are made once an individual achieves and sustains a job outcome) and at the level of the large scale regional CPAs

(where contractual performance expectations are set and assessed and where ‘high performance’ *within each CPA* is rewarded through market share shift).

Job outcome performance does vary across these extensive CPA regions and performance is consistently higher in more buoyant regions surrounding London and the South East. Competition is orchestrated between the two or three Prime providers operating within each of these CPAs but performance variation across contracts is small when compared to the variation in job outcome rates between local authorities *within* each CPA. Because the contractual agreements are sited at the level of CPA those local authorities with unacceptably low job outcome performance are contractually invisible, even if of significant spatial policy concern given their low performance, since they are diluted and concealed within the CPA aggregate performance figures.

As a result, there are three serious concerns around the current reliance on such large, undifferentiated regional CPAs as a unit for constructing key programme steering tools. Firstly, since the contractual minimum performance levels are standardised across Britain and this unadjusted national standard is applied indiscriminately, there is a risk that Primes are judged as ‘underperforming’ when their lower aggregate performance at the level of CPA may be informed by lower prevailing levels of a demand across a particular region. Providers in regions of higher unemployment are likely to find it more challenging to achieve job outcome rates at a level above the nationally set targets. Indeed, the OECD (2014, p. 201) note that “a provider might underperform against MPLs when in fact their performance is good” whilst providers in areas with better-performing labour markets might coast to performance above the MPLs. As a result the sanctions for underperformance may not be applied appropriately.

Secondly, and relatedly, none of the official programme performance indicators allow for meaningful comparison of performance *between* CPAs nationally. While it is possible to compare performance across the two or three Prime providers within the *same* CPA (as participants are randomly allocated between the providers) “it is not clear whether they are both high- or low-performers” (OECD, 2014, p. 28). Average job outcome rates do vary across the CPAs and according to current performance monitoring approaches adopted by the DWP it is not clear to what extent this reflects genuine variation in provider performance or, rather, differences in regional context and labour market characteristics.

Thirdly, and more substantively, the Work Programme design makes no attempt to neutralise the far greater variation in difficulty in entering the labour market *within* CPAs. The empirical findings presented above indicate stark sub-regional patterning in performance where more deprived local authority areas with proportionally larger long-term unemployed populations are experiencing significantly worse performance than more economically buoyant areas. The programme’s structure of regional CPA

competition and contractual minima appear to have done nothing to mitigate these spatial inequalities, nor would they be expected to function in this way.

While these patterns are consistent with a process of spatialised creaming and parking, from the available data it is not possible to determine whether this procyclical variation in performance – where better outcomes are achieved and paid for where they are needed least – has occurred due to the cynical and systematic behaviour of providers or because, despite provider efforts, outcomes are consistently harder and hence less-likely to be achieved in some parts of Britain. Nonetheless, it is clear that the high degree of spatial variation matters for participant outcomes. Given what is known about the well-being effects of non-meaningful activation interventions (Carter and Whitworth, 2017) this leaves participants living in deprived areas particularly vulnerable. The failure to calibrate for locally varied labour market contexts works against the programmatic ambition to reduce *gaps in performance outcomes between the easier- and harder-to-help* neighbourhoods and areas.

In sum, the undifferentiated geographical container of regional Contract Package Areas does not appear to be functioning as an appropriate geographical unit to uphold minimum performance standards or incentivise providers towards the pursuit of programmatic objectives. Potential responses to this situation can be contemplated on two fronts: there is a need to consider a more appropriate indicator of provider performance, which is sensitive to context; and there is the need for a re-scaling of tools to incentivise provider behaviours at a level *below* the CPA.

Turning to the first of these, there is a need to move beyond an ‘unadjusted’ indicator of performance to a metric which is sensitive to context in making performance comparisons. In the UK, a similar challenge has been tackled in the presentation of examination results for schools. The danger here is that through an unadjusted comparison of average examination results schools in less deprived areas who have higher attaining pupils at intake will tend to score more highly in absolute terms irrespective of the effectiveness of the schooling provided (Leckie and Goldstein, 2016). Differences in schools’ pass rates are “too often attributed solely to a supposed difference in the educational effectiveness of the two schools” (Leckie and Goldstein, 2016, p. 4). From 2006 – 2010 UK Government sought to better separate schools’ ‘true’ effects from the composition and context of their intakes using a regression-based ‘contextual value-added’ (CVA) approach.

Relatedly, the Job Services Australia model advances a form of performance measurement system which is sensitive to local labour market conditions as well as jobseeker characteristics. The “Star Ratings” system uses statistical regression analysis to allow the comparison of provider performances across Australia. By controlling for a set of established job seeker and labour market characteristics which impact most on the achievement of job outcomes, the Star Ratings model calculates the performance “providers could reasonably be expected to have achieved given the unique set of job

seekers they have assisted in their specific labour market” (Australian Government, 2012, p. 2).

Figure 6.22 shows the output from a prototype of Contextual Value Added performance for the Work Programme by introducing ‘contracts’ as dummy variables alongside the fully specified individual-level model to predict job outcomes developed in Chapter 5. This approach then controls for a detailed set of programme participant characteristics, an indicator of local area deprivation, and labour market conditions. Any ‘effect’ carried by the contract dummy variables can be understood as identifying ‘good’ and ‘bad’ relative performance, compared to the *expected* level of performance for the specific cohort and operating environment of the contract. The model is constructed as a binary logistic regression model and the ‘reference’ for the contract dummy variables is contract 1. In the exploratory analysis shown in Figure 6.22, the interpretation of odds ratios on the contracts is made in reference to contract 1, which in itself is a higher performing contract.

In Figure 6.22, moving down the pane the contracts are ordered from ‘worst’ to ‘best’ performance (the blue bars) according to DWP’s current assessment of performance, which is built only using payment groups and contracts. By contrast, the red bars represent an indicator of CVA – that is holding the full set of explanatory variables constant. Notably, the ‘best performing’ contract under the current assessment approach (contract 22) is *not* the best performing when controlling for composition and context. There are also a series of contracts where raw performance is well below the performance of the reference contract (i.e. blue bars are a considerable distance below an odds ratio of 1, where the dashed line indicates performance comparable to the reference contract) but where the red bars are nearly touching the reference line. These contracts are those which might be seen as ‘underperforming’ on the basis of current performance metrics but which are in ‘contextual value added’ terms performing comparably well.

The ‘Contextual Value Added’ approach piloted here begins to offer a more appropriate route to make comparisons of contractual performance across Britain and hence offers a more sensible basis on which to consider grounds for performance management interventions, contract extension or termination. However, provided that this approach is applied only at the level of CPA then such a metric in itself will do nothing to better calibrate incentives to counteract the important variation in context *within* the contract areas. To incentivise providers to work more intensively in less promising local labour markets within CPAs a series of reforms to the operation of the market structure and contracting is required.

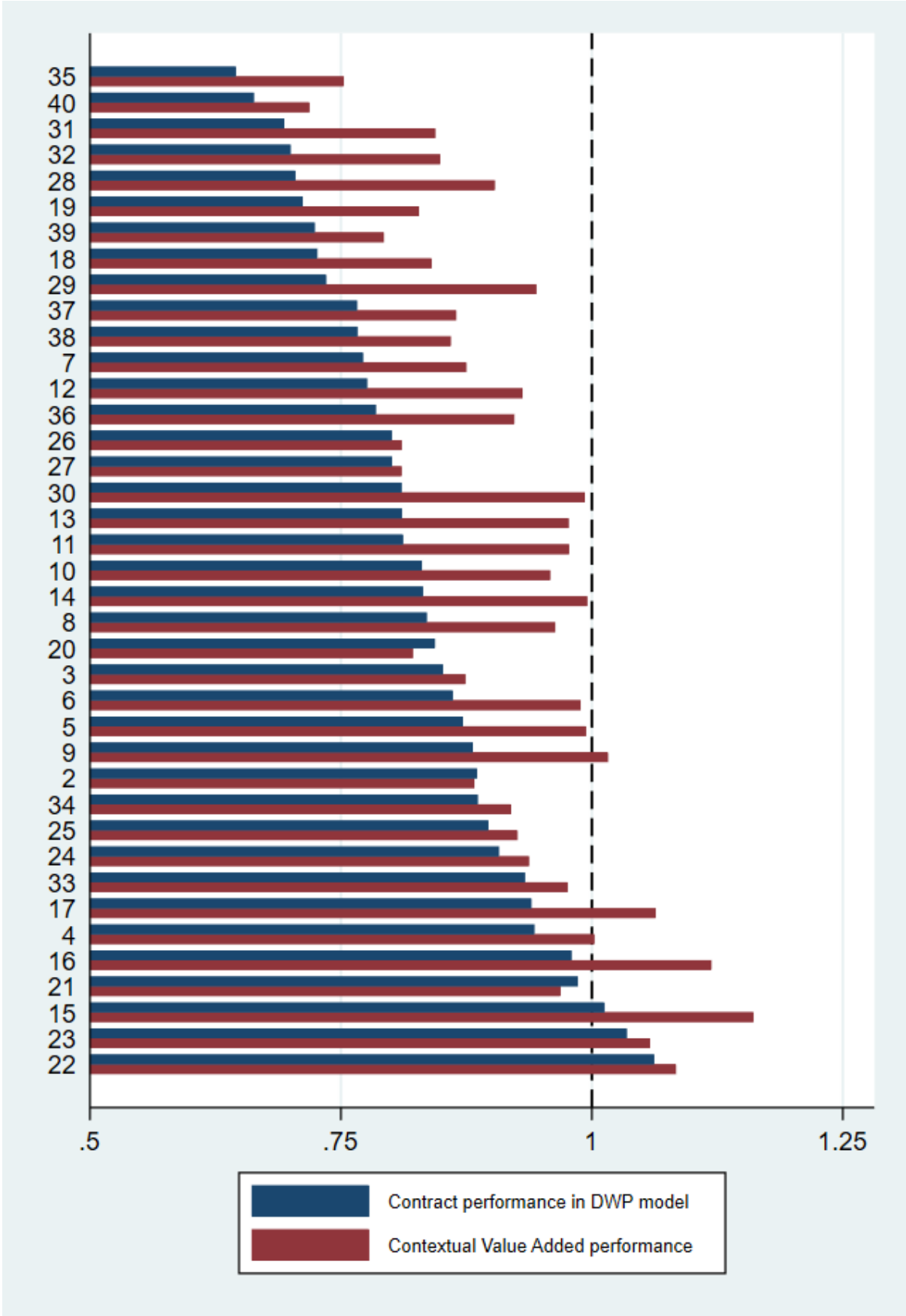


Figure 6.22 Alternative measures of contract performance: a 'Contextual Value Added' for Work Programme Prime providers

To achieve this contextual sensitivity and incentivise providers to work against the grain of pre-existing spatial inequalities, the unit of performance assessment and incentives will need to be substantially reduced in geographic extent. For example, in the 2015 Australian Star Ratings 751 distinct office areas are incorporated within the Star Rating analysis, noting that the working age populations of the UK and Australia are broadly similar, at 64.92 and 66.46 million respectively (OECD, 2018).

There are then choices as to whether to ‘steer’ on an ongoing marketised basis, for example using counter-cyclic price adjustments for these new sub-CPA labour market units or perhaps to actively construct and monitor contractual performance minima at this revised scale. Indeed, it may well be sensible to look beyond marketised levers to consider greater use of procedural accountabilities such as clearer and more intensive minimum service expectations.

The implementation of a more deep-seated set of reforms in this vein would necessitate a critical reflection on the appropriate geographical unit of performance assessment and incentive setting. In the analyses above, local authorities clearly offer a predictive improvement compared to CPA but they do not necessarily capture the ‘local labour market’ context for programme participants. Previous spatial analysis of British employment support programmes have pragmatically adopted the management structure or office units used for programme implementation (for example, Sunley et al., 2006, 2001). In the Australian case, although the full model specification is not available in the public domain the description of the Star Rating system suggests that the unit at which contextual variables is incorporated is based on “Bureau of Statistics Statistical Regions” (Australian Government, 2012). Conceptually, units to capture relevant contextual factors ought to include dimensions of the local labour market which inform programme participant job-seeking and employment experiences, for example: the speed of job potential job transitions; employment stability; pay; hours; contract-type; and care compatibility of employment options *within an appropriately commutable area*.

Background preparations for this chapter produced ‘Tailored Travel to Work Areas’ which flexibly incorporated vacancies and job densities across malleable ‘catchment areas’ of frequently commuted journeys for those working in lower skill jobs whilst residing in the ‘home’ Middle Super Output Area for programme participants. This represents one viable tailored small scale geographical unit around which to think more realistically about which localised labour market contexts matter for programme participants. It is beyond the scope of this chapter to resolve this issue, rather the aim is more modestly to highlight the importance of spatial considerations. There still remains a question as to the most appropriate scale at which to accommodate contextual factors. Ultimately the scale for incorporating context needs to promote a balance between being an appropriate and flexible ‘local’ labour market for participants (i.e. a geography with meaning) with the need for transparent and comprehensible computation and application. Simply overlooking or assuming away the structural contextual features of local labour markets however is not justified.

7 Crafting incentives for sustained employment: Investigating employment and earning trajectories for Work Programme participants

7.1 Chapter summary

The previous empirical chapters have focused, in different ways, on Work Programme's headline job outcome performance. This is a key metric for understanding programme performance, yet also central to the Work Programme's design considerations is the challenge of pursuing *sustained* employment outcomes.

Within Work Programme the specification of job outcomes – which require 6 months of cumulative or continuous employment for the majority of participants – combined with a long and intricately constructed 'tail' of sustainment payment are intended to encourage Prime providers to support programme participants into stable employment. The incentive structure aspires to break the well-known 'low-pay no-pay' cycle experienced by many on the lower rungs of the UK labour market.

To examine this issue of sustainment in unprecedented richness, the author's unique academic access to a set of Real Time Information on the month-to-month earnings of a sample of Work Programme participants facilitates the first of its kind analysis of earning trajectories using sequence analysis.

7.2 Introduction: The evolution of payable 'outcomes' within welfare-to-work payment-by-results systems

In the field of quasi-marketised welfare-to-work services there are multiple and diverse outcome standards that have been applied both internationally and over time in the UK (Finn, 2010a, 2010b, 2011b, 2011c, 2012; Martin, 2015). US states, for example, vary considerably in the complexion and composition of outcomes specification drawing on measures related to job retention, training and qualifications, wages and benefits, and earnings gains (Finn, 2010b).

In part, this variation in outcomes specification can be understood to stem from the variability in the overarching objectives of activation policies which – as noted previously in Chapter 3 – encompass a diverse range of intentions and which themselves vary over time and space (ranging from social inclusion and poverty reduction to tighter employment targets) (Bonoli, 2013; Eichorst and Konle-Seidl, 2008; Weishaupt, 2011). Importantly, these alternative outcome specifications do not appear to be solely associated with different institutional settings and objectives. Rather, outcomes definition seem to be suspended within a process of constant and incremental evolution as policymakers strive iteratively to strike the optimal balance between competing concerns around outcome clarity, complexity, measurability, attribution, and the avoidance of perverse incentives (Carter, Forthcoming; Considine, 2005; Finn, 2008, 2010b).

In the Work Programme a particularly strong emphasis has been placed on *sustained* employment outcomes. Within the programme's private power market there is a heavy reliance on financial incentives through payment-by-results to steer provider behaviours and shape service provision. Within this, the length of sustained employment outcomes required to obtain maximum payments for providers is novel in comparison to previous employment programmes in the UK and is also seen to be innovative compared to other bold marketisers internationally such as the Australian model (OECD, 2014).

The 'job outcome' metric in the Work Programme is something of a misnomer: no payment is made for entry to employment. The job outcome is more strenuous than the labour market attachment 'job entry' rewards used in previous schemes on two key dimensions. Firstly, for the majority of participants (those in payment groups 1, 2 and 9) providers need to secure 26 weeks of continuous or cumulative spells in employment before they can claim a job outcome payment. Even for those participants in 'harder to help' payment groups 13 weeks of employment is required before any outcome payments will be made, which is much longer than for previous programmes (NAO, 2012; OECD, 2014). Secondly, job outcome (and sustainment) payments are only paid for periods of employment where participants have moved off out-of-work benefits which means that they must be working for at least 16 hours per week (DWP, 2010a). Even where participants enter employment it may not be of sufficient duration or of sufficient quantity (in terms of hours) to qualify as a 'job outcome'.

However, it is through 'sustainment payments' for sustaining employment after initial job outcome payments that the vast majority of potential payments are located in Work Programme. Sustainment fees can be claimed every four weeks after the job-outcome payment for another one to two years, depending on payment groups. Work Programme's bold experiment in its financial elevation of sustained employment within the overall payment profile has not been formally evaluated however, and certainly not using sophisticated sequence analysis techniques, with previous analyses focusing instead (somewhat understandably) on the simplest and headline binary job outcome performance measure. This chapter seeks to respond to the uncertain impacts of the programme's incentive structure, which is designed in such a way as to reify sustained employment outcomes, on the employment and earning experiences of participants on the ground.

The discussion proceeds across three core sections. Firstly the chapter briefly traces the imperatives which have informed the growing focus on *sustained* employment outcomes at the heart of employment support programmes, including contextual considerations of employment stability and the low-pay no-pay cycle on the lower rungs of the UK labour market. The second section outlines the complexities of the incentive structure and the competing expectations from policy designers and expert practitioners in the degree to which the incentives were expected to hold traction for provider behaviours. The third core section offers original empirical analysis of the

RTI earnings data for a sample of programme participants in order to answer the research question: *What are the employment and earning trajectories of Work Programme participants? Has the use of sustainment payments broken the low-pay no-pay cycle?*

Importantly, the extreme use of sustainment payments cannot be evaluated through conventional impact evaluation techniques due to the simultaneous national roll out of the Work Programme's private power market and consequent lack of an appropriate comparator group. A key contribution of the chapter therefore is the creative introduction of an innovative statistical technique to investigate the granular month-to-month earning and employment experiences of programme participants. The empirical section implements sequence analysis and Optimal Matching for the first time on a set of earnings data for participant in a welfare-to-work scheme. This method has not previously been applied to the investigation of welfare-to-work outcomes in the UK context (or indeed, elsewhere). Yet this methodological approach enables a more holistic and considerably more advanced understanding of employment experiences than would be available through more conventional statistical techniques by producing a typology of earnings and employment trajectories.

This empirical contribution then facilitates critical reflection on the degree to which payment incentives are related to earnings (in)stability amongst Work Programme participants. While there has been considerable critical air-time given to the way that the Work Programme payment-by-results system fails in some regards – most notably, the way that price levels are irresponsive to participant 'need' (Chapter 5); local context (Chapter 6); and the way that distance travelled towards work is wholly unacknowledged (Morphy et al., 2012; Suleiman, 2014; WPSC, 2015) – there has been almost no consideration of the degree to which the complex payment structure has stimulated or stymied the achievement of genuinely sustained employment outcomes for participants, despite this being one of its key programme objectives and design considerations. This is considered for the first time here, and in unprecedented richness and methodological subtlety.

7.3 The low-pay no-pay cycle: a background to unstable employment outcomes

The importance of *sustainable* employment as an objective for policy designers has been enshrined through three interrelated features of Work Programme context. Firstly, the challenge of sustainment sits alongside concerns around low-pay no-pay cycling at the lower-skill end of the UK labour market. Secondly, the temporary nature of employment outcomes and programme 'churning' has been raised as an issue within previous UK employment support programmes. Finally, sustained employment outcomes have been inscribed deep within the justificatory narrative of the Freud report (2007) where the foundational logic for the Work Programme is presented as its ability to deliver 'savings' to HM Treasury which can only be achieved through longer employment spells. Each of these dimensions is introduced in turn, before outlining

the incentive structure intended to promote sustained employment within the programme.

Since the 2000s there has been a growing awareness that contemporary accounts of poverty and social exclusion overlooked the significance of low wages and insecure work, informed by Byrne's suggestion that "poor work is the big story" representing "the most significant kind of excluded life" (Byrne, 1999, p. 74; McKnight, 2002; Oakley, 2015). This sat uncomfortably with the defining political mantra of UK governments that "employment is the best route out of poverty", a proposition central to policy on welfare since that time (Shildrick et al., 2012, p. 8). Shildrick's extensive longitudinal qualitative research identified graphically the phenomenon of the 'low pay no pay cycle' where the predominant experience of low-income interviewees was one of moving in and out of low-paid, short-term jobs, repeatedly cycling on and off benefits (Shildrick, 2012; Shildrick et al., 2012).

Quantitative analysis bolsters the credence of these findings and several studies have used the DWP's administrative data to identify the issue of 'repeat claims' (Carpenter, 2006; Leitch, 2006; NAO, 2007). Amongst those who transitioned from benefits into work, more than one in five were found to reclaim the benefit within 13 weeks and 40 percent were reclaiming within 6 months (PAC, 2008). Importantly, the transience of these employment spells does not appear to be experienced by choice or by the inability of people seeking jobs to sustain work: the majority of those entering temporary jobs take them because they are unable to find a permanent position (Ashworth and Liu, 2001; Tomlinson and Walker, 2010).

Given the extent of low-pay no-pay cycling experienced by people exiting out-of-work benefits it is perhaps unsurprising that welfare-to-work programmes have historically struggled to achieve sustained employment for participants. Both international and UK evidence indicates issues of cyclical participation in employment support schemes. An evaluation of the much heralded Wisconsin Works scheme for unemployed women with families found that over a six-year period only 35 percent of programme participants averaged more than three-quarters of each year in employment (Wu et al., 2008). In the UK, across the 2000s the New Deals equally experienced problems with 'revolving-door' participation, since while initial job entry figures were promising it increasingly became apparent that participants moved from training into short-term employment, "and then back to unemployment, eventually repeating their participation in training" (Lindsay et al., 2007, p. 543). Around 20 per cent of those who participated in the New Deal for Young People in 2006 were attending the programme for at least the second time (ONS, 2006 in Lindsay et al., 2007).

There are important personal costs for individuals involved in low-pay no-pay cycling including reduced feelings of self-worth and confidence, reductions in future earning potential and declining chances for stable future employment (Booth et al., 2002; Goulden, 2010; Gregg, 2001; Mulheirn et al., 2009; Walker and Kellard, 2001). Yet it is parallel concerns of financial cost that appear to have been a preoccupation for

policymakers since employment sustainability is of paramount importance to the cost-effectiveness of employment programmes (Mulheirn et al., 2009; NAO, 2007). Given the high costs of achieving ‘additional’ job starts participants would need to remain in employment and ‘off benefit’ for considerable periods before the cost of interventions is effectively recouped (Mulheirn et al., 2009; NAO, 2007).

In the Work Programme the important connections between ‘value for money’ and duration of employment outcomes are further heightened since the large scale and universality of the scheme has in large part been justified by the savings expected to accrue to HM Treasury through reduced benefit spending and elevated revenue from taxation (Freud, 2007). Freud successfully argued for the DEL/AME switch – effectively uncapping the total amount that could be spent on outcomes in the Work Programme (where employment support is conventionally constrained by Departmental Expenditure Limits) – because “spending on welfare to work programmes could come from the very savings made from reducing the benefit bill as a result of them” (Haddon, 2012, p. 7). Because of this financing arrangement, in addition to the usual pressures experienced by civil servants to design a coherent outcome-based payment model, the ‘results’ payments in the Work Programme needed to be tied explicitly to welfare savings.

Importantly, whilst it is sometimes argued that issues of employment insecurity require interventions on the demand side, either by shifting sectors towards higher skill levels and hence higher wages and more secure contract terms or by tightening employment regulation to minimise the use of temporary or ‘zero-hours contracts’ for example, commentators are quick to point out that such considerations seem firmly off the table for policymakers (Hepple, 2013). The Coalition administration has shown no desire to hamper the hyper-flexibility of the UK labour market (Hepple, 2013; Wiggan, 2012). The residual challenge then is for employment support interventions to “get the best of both worlds: a flexible labour market in which people who would otherwise cycle in and out of employment are supported to remain in work and move between jobs without long and repeated periods of unemployment” (Mulheirn et al., 2009, p. 38).

The clear imperative then is to motivate Work Programme Prime providers to deliver sustainable job outcomes. But Mulheirn et al., (2009) also indicate there is a need for subtlety in terms of tying cash incentives to sustained job outcomes, in particular that the results system should be calibrated to avoid bias against helping those who are least likely to enter and retain employment (as discussed in Chapter 5) and moreover that incentives should be designed in such a way as to avoid making providers averse to promoting temporary employment at all. Several reports acknowledge the potential importance of temporary jobs to re-engage and support entry to stable employment over time (Booth et al., 2002; Zijl et al., 2004). The design challenge “is to structure incentives in a way to ensure that service providers turn temporary jobs into stepping stones to permanent employment” (Mulheirn et al., 2009, p. 19).

7.4 Coherent incentives? A complex response from designers in steering for sustainment

The design of payable outcomes within the Work Programme’s payment-by-results schedule clearly seeks to respond to these concerns as seen in the payment model schematic reproduced in Figure 7.1. The long tail of sustainment fees account for between 57 and 76 percent of the potential income from participants starting in the first year and with increasing relevance in subsequent years due to the incremental reduction and eventual removal of attachment fees (OECD, 2014).

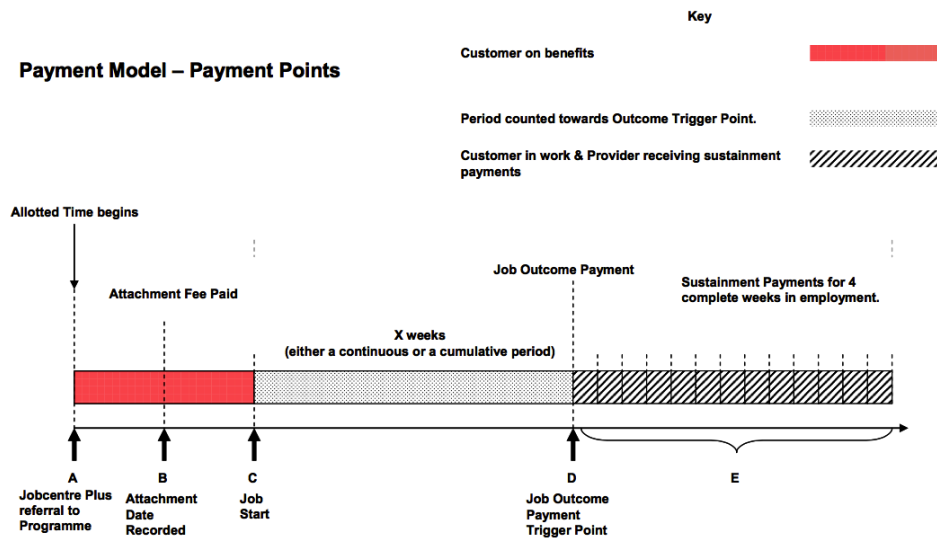


Figure 7.1 Illustration of Work Programme payment model, reproduced from DWP, 2010b, p. 11

Notionally, were the set of financial incentives manifest in the payment schedule responded to wholeheartedly (and ‘rationally’) by Primes it is anticipated that this would stimulate either (or all) of three responses, which would mark the programme as distinct from previous schemes. In addition to the support conventionally provided to target entry to employment (quickly, but regardless of job type or fit) it would be expected that providers would in these circumstances:

- 1) Deliver support to enable participants to remain in employment, overcoming any challenges or issues that might have been expected to lead to participants dropping out of the labour market (e.g. with issues around childcare) and potentially also to give support for participants to convert temporary jobs to permanent roles. In terms of intervention focus, the *type* of job entered does not change but *in work* support to retain employment is increased;
- 2) Provide a brokerage role once a participant has entered a (first) job to enable them to quickly enter alternative (second or third) roles once an initial temporary contract reaches its end. Again the initial intervention does not change the type of job first entered but *in-work* support looks to enable participants to quickly ‘step’ from one job to another;
- 3) Support participants to enter jobs that are markedly different in employment terms and quality from the employment typically entered by those exiting out-

of-work benefits. Services would shift from focusing on available employment to centre on jobs that offer permanent employment, which is compatible with participants' wider ambitions and responsibilities such that job entry becomes synonymous with long-term (>1 year) employment and there is minimal risk of a participant dropping out of the labour market. *Here the objective of initial intervention shifts from the speedy insertion of participants into any job, and instead offers a more 'choosy' approach to identify jobs with longer-term prospects.*

Importantly, however, overlaying the already complex set of job outcomes and sustainment outcomes, the incentives for sustainment vary over time and interact in somewhat complex ways with the timelines of programme engagement for each participant. While participants are within the 104 week initial window of 'allotted time' on the programme then employment may be built up cumulatively to contribute to the ultimate achievement of job outcomes. Likewise it is possible for periods contributing to 'sustainment' in employment to be accrued cumulatively, so there is – counter to the labelling – scope for there to be breaks in employment between sustainment payments. When a period in employment spans the time point 104 weeks after initial programme attachment providers are still able to continue to claim job outcomes and sustainment payments up to the maximum number. However, if there is a break in employment of 2-days or more after the 104 week marker then no further sustainment (or job outcome) payments can be claimed (DWP, 2010a).¹⁰ This means that whether and when it is deemed worthwhile for Primes to expend effort on the pursuit of sustained employment through the tactics outlined above is rather intricate and will further vary in accordance with the payment group into which a participant has been allocated.

7.4.1 Anticipated response to sustainment incentives – the 'known unknown' in Work Programme design

A key assumption on the part of programme designers then is that providers are both capable of interpreting and operationalising these incentives and that an appropriate toolkit of in-work support is available in response. Both assumptions have been contested.

In evidence presented to the Work and Pensions Select Committee the ability of providers to respond to the incentives for sustained employment outcomes is acknowledged as "one of the key known unknowns" of the Work Programme by experts in the field (Simmonds in WPSC, 2011, Q 30). Firstly, the complexity of the payment structure makes the incentives challenging for providers to operationalise. Secondly, the value accorded to sustained employment may or may not align with the relative difficulty of enabling different participants to sustain work (and given the

¹⁰ Similarly, if a period of employment which *would* have been eligible for a job outcome payment begins prior to the 104 week marker, but is broken after 104 weeks and before the full period of employment required for the achievement of a job outcome is achieved, then no payment will be made.

analysis in Chapter 5 showing that the differential job outcome payments fail to align with distance to labour market, it would seem sensible to be wary in relation to the alignment). Moreover, the financial value of connecting shorter spells of intermittent work varies at different time points across the participation window. Again, commentary from experts presented at Select Committee hearings emphasises this:

“It’s a very complex structure...In itself as a design it seems to answer some of the problems about sustainability, for example, but I think there are going to be some real issues in the delivery...and the question is whether it will tweak the right kinds of incentives.”

Finn in WPSC, 2011, Q50

Furthermore, that “[W]e can see the logic of including all those different payments...but there is a danger that the very complexity leads to an unravelling and a whole load of unintended consequences” (Simmonds in WPSC, 2011, Q50).

There is evidence that the specification of employment eligible for outcome payments within a payment-by-results framework has tangible impacts on the characteristics of work entered by programme participants. Hales et al., (2003) compared a cohort of participants in the Employment Zones programme with a similar group of participants in the Jobcentre Plus New Deal: although the net effect for treatment (the Employment Zone participants compared to New Deal) was not statistically significant for all jobs, Employment Zone participation showed a positive and highly statistically significant effect for jobs at 16-hours per week and over. Bruttel (2004) suggests that this was likely to have been achieved because providers concentrate on 16-hour plus jobs as they are only able to claim a full outcome fee for programme participants who are working at least 16 hours per week. The implication of this finding is that the specification of job outcomes and sustainment in the Work Programme *may* shape the *type* of jobs that providers support participants to enter.

The ultimate response to the incentive structure is highly ambiguous and there is a concern that the complexity may mean that at the front-line services proceed ‘as usual’ with an approach that pursues ‘work-first’ and ‘any available job’. In a situation where providers do seek to respond to the incentives and aspire to provide in-work support to those participants requiring this, it is both assumed that providers are able to effectively identify those participants who would benefit from such interventions *and* that effective in-work interventions to support sustained employment are available. Crucially, the evidence base on ‘what works’ for in-work support is acknowledged as being very limited at this stage (SASC, 2017; WPSC, 2016).

Hence the research and policy community finds itself presently in a position where it is not clear whether the incentives designed into the payment-by-results mechanism are capable of stimulating the desired results in terms of sustained employment for participants. It is here where UK Work Programme evaluations so far are highly limited. Whilst the nature of payment-by-results inevitably means that it is only

eligible portions of sustained work which are ‘paid for’ by commissioners this does not necessarily mean that the incentive structure is working effectively or as expected in terms of encouraging job entries to lead to sustained employment. There is considerable uncertainty as to how effectively Work Programme supports are facilitating translation of job entry to sustained, ongoing employment and in-work provision has been particularly questionable (Meager et al., 2014). Although it is only sustained work which is being paid for, the programme itself is not measuring or recording the degree to which this *is not* achieved. The perturbations, interruptions and cycles between unemployment and work are not documented.

It is here that this chapter drives forward the discussion by providing original empirical material which brings a more subtle analysis of employment and earning experiences. This more detailed investigation must necessarily draw on data beyond the official Work Programme performance figures and official payment records – since these are inevitably constrained by the way they record only those spells of employment which have successfully delivered job outcome payments and sustainment payments for providers. The essential facet for underlying data in this analysis is that it is able to convey *any* spell of employment – both those that ultimately contribute to the achievement of provider payments – but crucially – also those which do not. Additionally, one would ideally wish not only to understand in a binary sense whether there is or is not employment at any given point in time but also to go further in understanding something about the earnings during those employment spells. Here, the author’s unique academic access to a set of Real Time Information on the month-to-month earnings of a sample of Work Programme participants has facilitated the first of its kind analysis of earning trajectories using sequence analysis.

7.5 Putting employment trajectories in context: common earning patterns for Work Programme participants

Before presenting the results of the OM and cluster analysis it is important to position the empirical findings alongside a broader consideration of the employment and earnings experiences of Work Programme participants.

Figure 7.2 sets out the high-level characteristics associated with the sample. Only 45% of those attached to Work Programme between April 2013 and September 2014 feature in the RTI earnings dataset. This is comparable to the findings reported in the official programme evaluation, produced by Meager et al., (2014) where a statistically representative telephone survey found that 44% of participants had been in work *at some point* since their referral to the programme. The corollary is that around half of programme participants leave the scheme after two years of support without having spent any time at all in paid work (Meager et al., 2014).

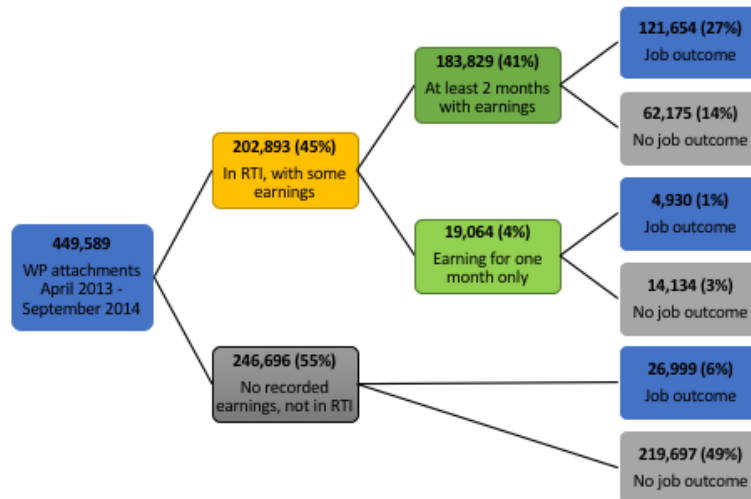


Figure 7.2 Earnings profile for participants attached to Work Programme between April 2013 and September 2014, final column presents overall percentage breakdown

7.5.1 The gulf between ‘job entry’ and ‘job outcomes’

Importantly, entry to employment – in other words, experiencing at least one month of non-zero earnings in the RTI dataset – is not tantamount to achieving a job outcome. Prior to the analysis presented here, the best indication of the mismatch between job entry and job outcome rates could only be considered by placing unofficial Work Programme performance figures published by the Employment Related Services Association (ERSA, the representative body for the employment support sector) alongside the official job outcome data provided by DWP. Positioning these sources alongside one another immediately suggested that there was a discrepancy between entry to employment and official job outcome rates. ERSA do not position their figures as a proxy for government statistics, rather the publications are intended to provide a more timely report on Work Programme performance. ERSA’s figures do however helpfully illustrate the discrepancy between the numbers of participants who have spent any time in paid work and those who ultimately achieve job outcomes. ERSA suggest that between 65% and 85% of job starts will convert into job outcomes (ERSA, 2013). The initial breakdown of job outcomes for the RTI sample indicates that the conversion from job entry to job outcome is likely to be towards the lower end of this range: the job outcome rate for those participants who feature in the RTI is 62%. This suggests that around one-third of those who enter work are unable to sustain this sufficiently to achieve a job outcome.¹¹

There is also the potential for further decay in terms of job stability since not all those who achieve a job outcome sustain employment and hence fail to deliver the maximum possible number of sustainment payments to their Prime provider. DWP’s official statistics suggest that for early waves of referrals with sufficient time following a job

¹¹ It should be noted that those participants referred towards the end of the RTI window (from May 2014 – September 2014) will not have been attached to the programme for the full 24-month period when the Work Programme administrative data was merged to the RTI, and hence, the job outcome rate for those featuring in the RTI dataset is ultimately likely to be marginally higher.

outcome payment, of those who achieved a job outcome only 65.2% go on to attain the maximum possible number of sustainment payments (DWP, 2013c, p. 2013). Unfortunately, the limitation of a 15-month window for the RTI data serves as an impediment to the further investigation of longer-term stability of employment amongst those who have achieved job outcomes. The analytical focus in this chapter provides a more general exploration of employment and earning experiences, rather than focusing only on longer term sustainment amongst those with job outcomes.

Overall, the initial hints of a ‘missing third’ of job entries which have not been successfully converted to job outcomes for Work Programme participants – as seen through the ‘gap’ between ERSA and DWP figures – is supported more concretely by the highly imperfect job outcome rate amongst those who feature in the RTI dataset. The job outcome rate amongst those participants who have earnings recorded in the RTI dataset is 62 percent, suggesting that the experience of entering employment is not automatically or easily converted to sustained work.

The complexity of these employment and earning histories and their fractured and potentially circuitous nature has not previously been investigated but it is to these patterns and trajectories that the analysis now turns.

7.6 Understanding employment and earnings over time

Informed by the growing sequence analysis literature (Brzinsky-Fay, 2007; Dorsett and Lucchino, 2014; Quintini and Manfredi, 2009), the analysis traces individual unemployment-earning trajectories and classifies them into several transition pathways. This approach seeks to unpack the entire employment and earning trajectory experienced by programme participants following attachment to the programme, including the nature of each month’s labour market status and the ordering of spells within the wider transition pattern. As discussed in detail in Chapter 4, a pragmatic, policy-informed approach is taken to constructing the distinct monthly labour market status types.

Element	Definition
Z	No earnings recorded for this month in RTI
L	‘Low earning month’ earnings are equivalent to earning less than 16hr per week at NMW
M	‘Middle earning month’ earnings are equivalent to earning more than 16hr per week at NMW but less than 35hr per week at NMW
H	‘Higher earning month’ earnings are equivalent to earning more than 35hr per week at NMW
U	No RTI data available

Figure 7.3 Definition of distinct labour market elements within participant employment trajectories

Conveying diverse patterns in temporal data is analytically challenging and hence there is great value in pursuing analytical approaches that are more sensitive to variations in patterns of progression over time. The charts in Figure 7.4 offer two preliminary graphical representations of the earning and employment status for the full Work Programme cohort attached between April 2013 and September 2014.

The upper panel shows, for each month following the month of attachment (horizontal axis), the proportion of Work Programme participants in each of the five element types (vertical axis). This has the advantage of being straightforward to ‘read’ but has important limitations, as highlighted by Quintini and Manfredi (2009). Given that it does not follow individuals over time the upper chart cannot be used by itself to determine the characteristics of a group of sequences. For example, while a significant portion of Work Programme participants are experiencing ‘0 earnings’ at the end of the observation period it is unclear whether these are the same individuals who were without earnings at the start of the tracking window, a wholly different group who became unemployed at a later point in the window, or indeed, some combination of the two.

For the purposes of sequence analysis then the lower portion of Figure 7.4 provides a more helpful representation. Here entire individual earning trajectories are plotted horizontally, meaning that each horizontal stripe when read from left to right represents the 16 monthly activity statuses experienced by one individual. Five shades are used to denote the different element types, and each change in colour corresponds to a change in status.

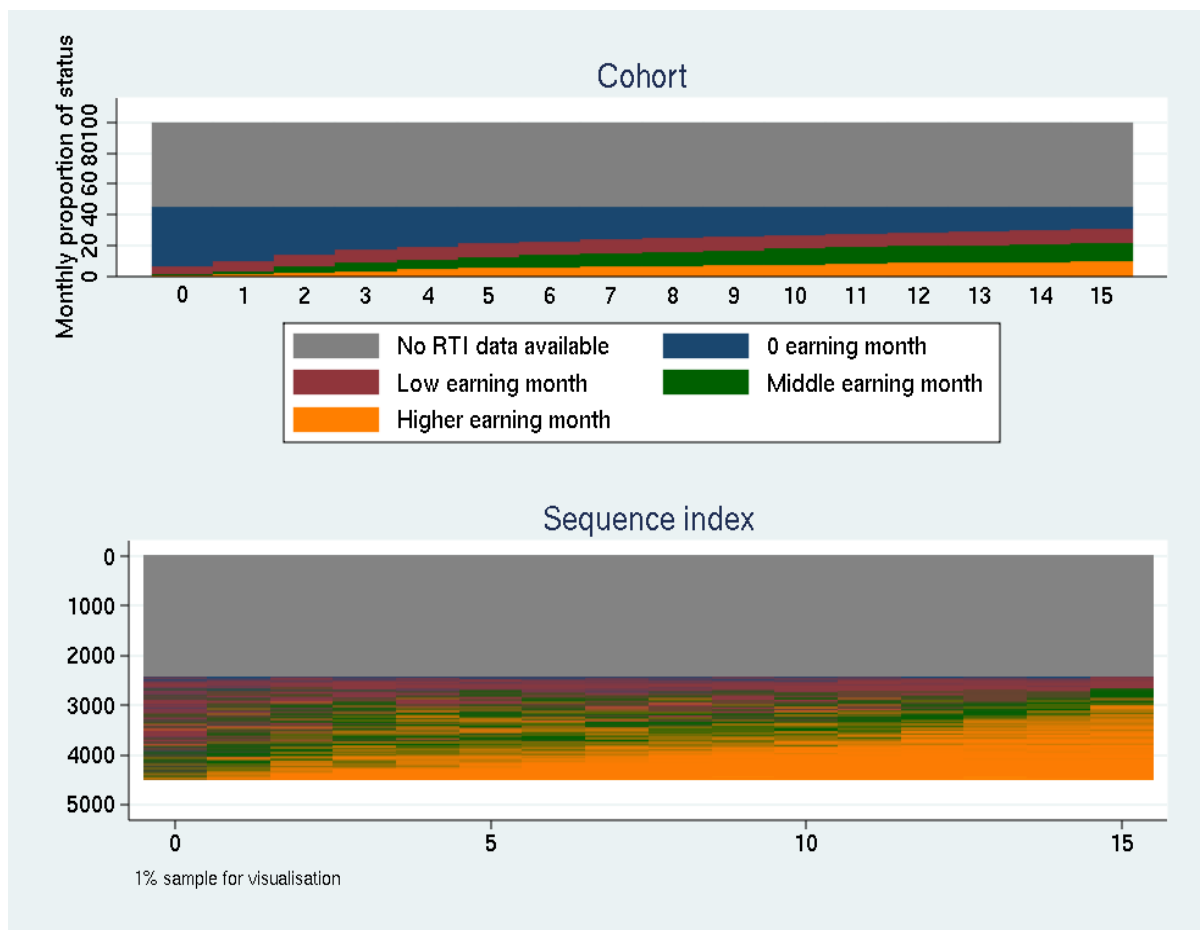


Figure 7.4 Investigating employment and earnings pathways: Comparing cohort-level representation to sequence index plot for Work Programme participants

Recalling the method described in Chapter 4, within Optimal Matching the sequence for every participant – i.e. the month-to-month series of particular employment and earning ‘element’ types (Figure 7.5) – is compared to that of every other participant.

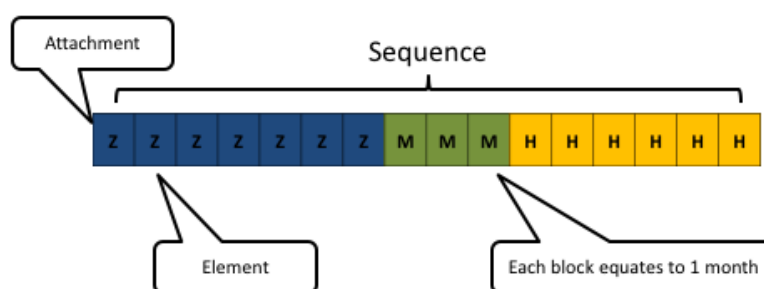


Figure 7.5 Earnings trajectories as a sequence of distinct elements

The number and type of changes (substitutions, or insert/delete functions as in Figure 7.6) that would be needed to convert one sequence to the other is referred to as ‘alignment’. This process is used to produce a measure of dissimilarity and the comparison of more ‘dissimilar’ sequences effectively results in larger difference values (the Levenshtein distance). The key output from the OM process is a distance

matrix: a symmetric matrix where both rows (i) and columns (j) represent the individual trajectories in the sample and each cell a_{ij} contains the distance between the sequence of individual i and of individual j. Cluster analysis is then used to bring clusters of sequences together in such a way that the resultant clusters *maximise the similarity* of sequences *within* each cluster and *maximise the differences* or dissimilarities *between* different clusters.

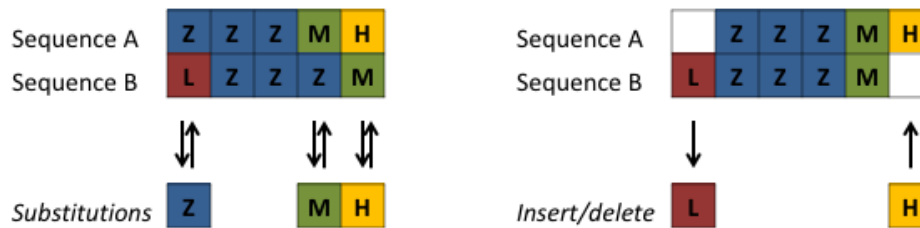


Figure 7.6 Optimal Matching operations

Running OM is a highly computationally intensive process and therefore the analysis that follows has been developed using a random sample of 20,000 Work Programme participants who feature in the RTI dataset. Cluster analysis is applied to the derived measures of dissimilarity in order to group more similar trajectories together.

A degree of discretion and judgement are applied in order to identify the most appropriate number of clusters since there is no formal statistical method for identifying the ‘best’ solution. Cluster analysis stopping rules have been developed to support this task and the commonly applied (and well supported, Milligan and Cooper, 1985) Caliński-Harabasz pseudo-F index (Caliński and Harabasz, 1974) is displayed in Figure 7.7. This was produced using functions within the ‘clustermat stop’ command in Stata 10 which calculates the Caliński-Harabasz index from the pairwise dissimilarities in the distance matrix (Halpin, 2016 and personal correspondence with Halpin).

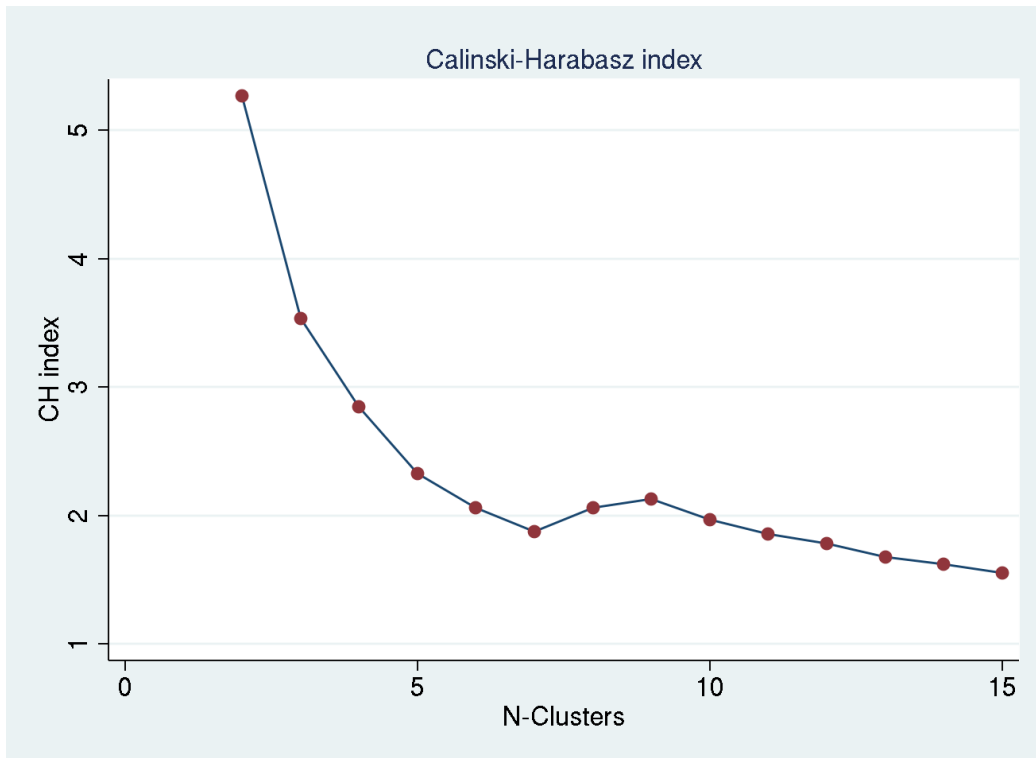


Figure 7.7 Identifying the optimal cluster solution: Calinski-Harabasz pseudo-F index for Ward's method

For interpretation, larger values of the Calinski -Harabasz index are indicative of distinct clustering (StataCorp, 2013) and so the slight uplift in the curve at the 9-cluster mark suggests that this might offer an appropriate solution. Brzinsky-Fay (2007) notes that contextual arguments should be used to define the appropriate number of clusters and suggests that the observation of analytically meaningful groups each containing a sufficient number of cases should be used to define the appropriate number of clusters. Further investigation of each of the solutions from 2 – 15 clusters confirms that the 9 cluster solution offers a meaningful resolution and discussions of the results with policymakers and analytical staff at DWP confirm that this cluster solution is analytically meaningful. As Ward's is a hierarchical agglomerative clustering method (i.e. it works by fusing clusters from the bottom up) it is also possible to consider the nine clusters as nested within 4 overarching 'super-clusters' as shown in the abbreviated dendrogram in Figure 7.8.

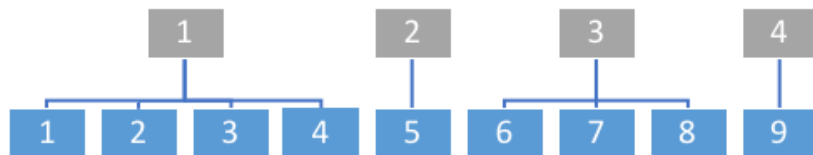


Figure 7.8 Ward's linkage identifies 9 clusters (blue) within 4 'super-clusters' (grey)

7.7 A typology of employment and earnings trajectories amongst Work Programme participants

Overall, therefore, nine groups of earning trajectories were identified amongst those Work Programme participants who appear in the RTI dataset and these are grouped into four high-level categories or ‘super-clusters’. The trajectories typifying each of the groups are visualised through sequence index plots which show the full employment and earning histories for each group member. The high-level categories are named according to the dominant features of the trajectories in each group and are described in turn below.

7.7.1 Super-cluster 1: ‘slow and unstable’

The earning trajectories experienced by Work Programme participants in the four clusters within the first ‘super-cluster’ are characterised either by the instability of earning trajectories or through the relatively long time periods taken to first enter employment (or both), hence the selection of the title for this super-cluster as ‘slow and unstable’.

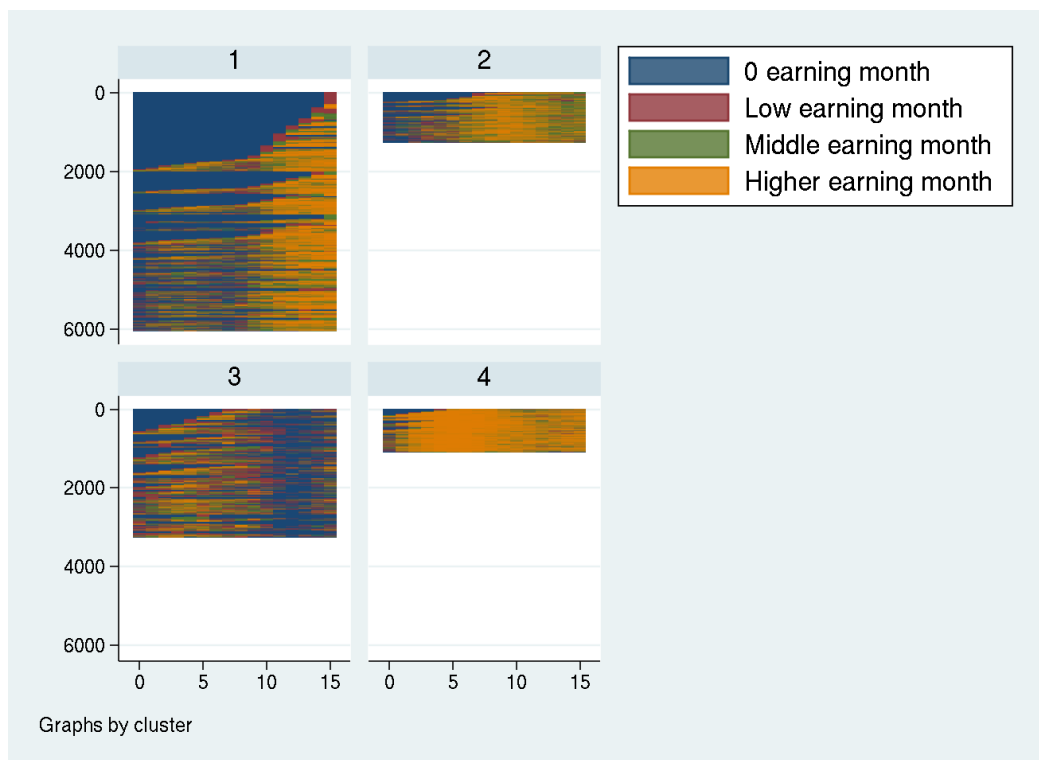


Figure 7.9 Sequence index plot for super-cluster 1: ‘slow and unstable’ earning trajectories

Cluster 1 shows trajectories where it has taken particularly long periods for participants to enter paid work, conveyed by the large expanse of non-earning months in the left-hand phase of the sequence index plot. On average, participants in this cluster take nearly ten months to enter employment for the first time (Figure 7.14 presents a high-level summary of characteristics for each of the clusters). Those trajectories in the lower portion of the sequence plot show that where employment has been experienced more quickly, nil-earning months are again experienced after this

initial spell in work. This is by far the most populous of the clusters identified and 6,042 people sit within this group. Just over half (56.1%) of those experiencing this trajectory type go on to achieve job outcomes.

Those whose earnings trajectories are grouped in cluster 2 move into work more quickly (the mean time to enter employment for this cluster is 5.7 months) but experience several different earning levels across the RTI tracking period. While it seems common for individuals in this group to experience higher earnings levels at around 10 months following attachment, the following months are characterised by fragmented or lower earnings levels. Only 69 percent of this group ultimately achieve job outcomes.

Cluster 3 groups together some of the most chequered earning trajectories where employment instability is the norm. Members of this group experience a ‘Joseph’s coat’ of earning statuses and most earning trajectories in this group are punctuated by several months of unemployment (zero earnings) at ten-fourteen months subsequent to attachment. Unsurprisingly, the employment instability of this group is associated with particularly low levels of conversion from employment entry to the fulfilment of job outcome criteria and only 24 percent of this group achieve official job outcomes. Average earnings for participants in this group across the RTI tracking period are the lowest of all the clusters, at £1,171.

Cluster 4 has the smallest population of the clusters and contains only 1,080 participants. The trajectories initially appear quite similar to those of cluster 2 but with quicker initial employment transitions to generally higher earnings levels. The earning trajectories here dip or become fragmented and tend to fall away from the ‘higher earning’ phase from around 9 months onward. The cohort visualisation of the clusters shown in Figure 7.13 assists in clarifying the distinction between clusters 2 and 4 and the ‘dip’ in terms of earnings and employment occurs at an earlier point for those in cluster 4. Despite the preponderance for trajectories to dip to lower earnings levels in the later portion of the trajectory, around 70 percent (71.3%) of cluster members do trigger job outcomes. This is a much higher translation rate of employment entry to job outcome achievement than experienced in other clusters within this overarching super-cluster.

By making a reasonable assumption that the random sample of 20,000 trajectories drawn from the full RTI data for inclusion in the sequence analysis are representative of the full RTI dataset, and assuming that those attached to the Work Programme from April 2013 – September 2014, are broadly representative of the programme’s cohort as a whole, then it would be expected that just over a quarter (26.2%) of all those attached to the Work Programme would experience an employment and earning trajectory of the sort characterised by the ‘slow and unstable’ super-cluster.

7.7.2 Super-cluster 2: 'stuck in low-pay'

Super-cluster 2 contains only one cluster and trajectories for participants in this group are dominated by the low earning elements, with the title 'stuck in low-pay' reflecting this.

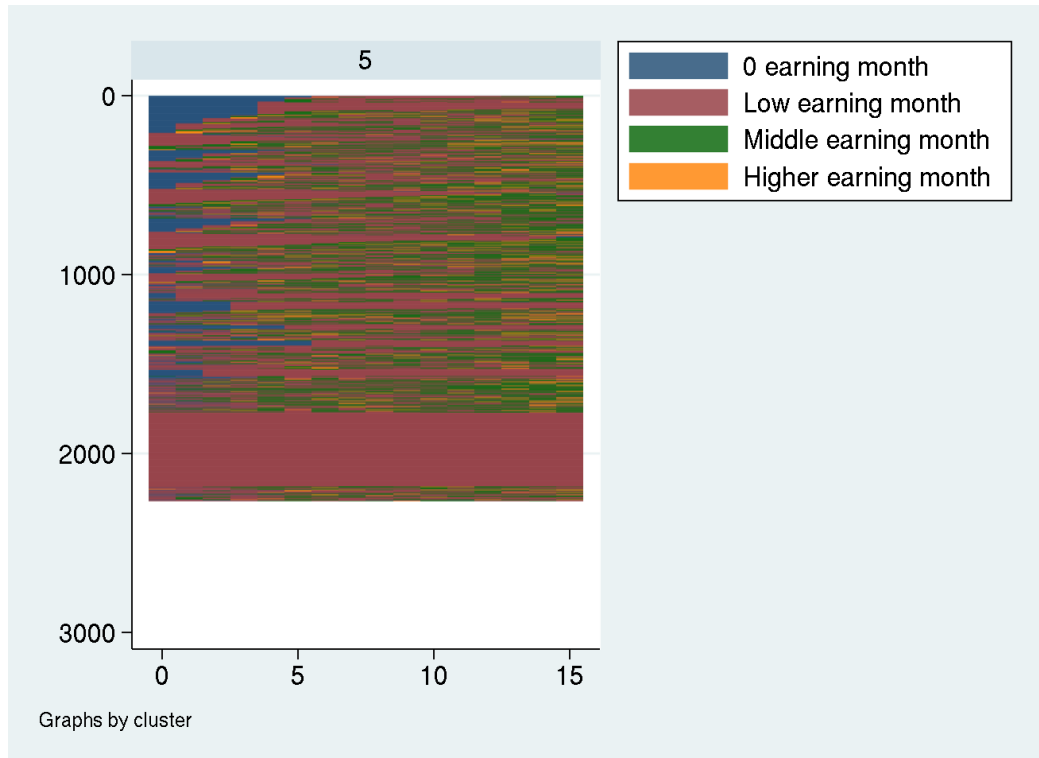


Figure 7.10 Sequence index plot for super-cluster 2: 'stuck in low-pay' earning trajectories

Again acknowledging the assumptions of representativeness outlined above, it would be expected that 5% of people attached to the Work Programme would experience this type of trajectory. Although the average time taken to enter employment is the quickest of all the clusters (1.3 months), average earnings for participants in this group across the 16-month window are considerably lower than for many of the other trajectory types at £4,421. The trajectories of participants in this group contain very few elements that are not low earning months. Indeed, in this cluster, 415 people were in paid work (and earning at a low-level) in the month they were attached to the programme and record low level earnings for every month within the 15-month RTI tracking period, never progressing above the equivalent of 16 hours a week at NMW. Around half of people (53%) experiencing this trajectory type ultimately achieve job outcomes for their providers.

7.7.3 Super-cluster 3: 'general earnings success'

The third super-cluster contains three clusters whose trajectories can be considered to convey 'general earnings success'. The trajectories of super-cluster 3 would be expected to represent the experiences of 9.8% of Work Programme participants.

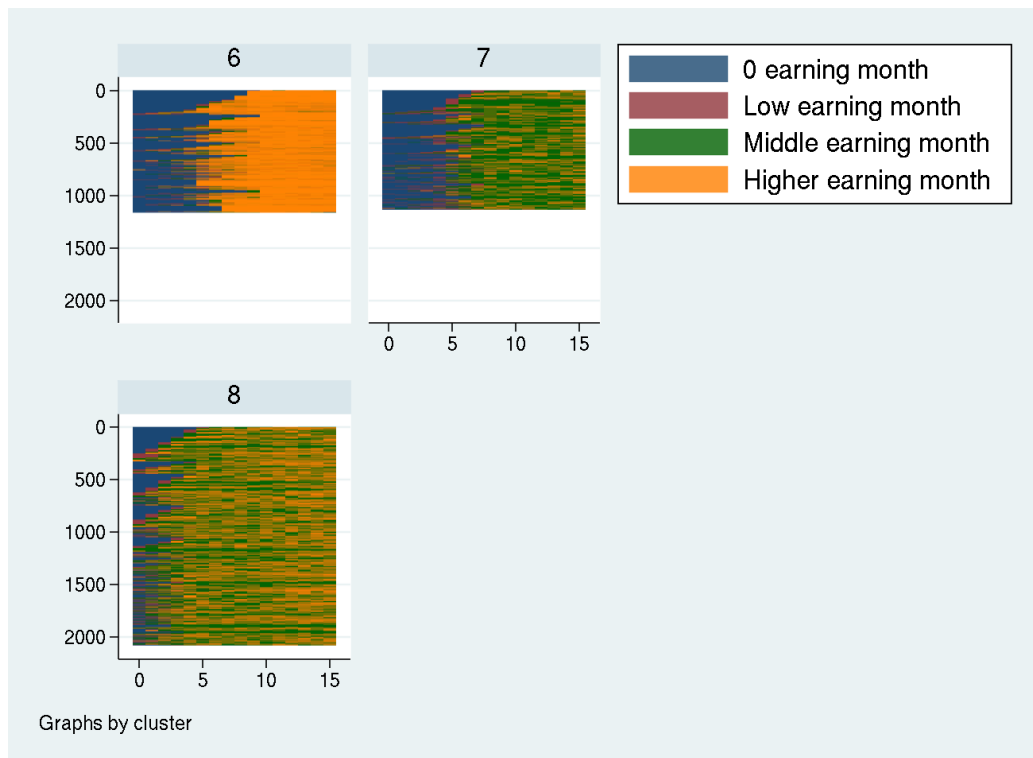


Figure 7.11 Sequence index plot for super-cluster 3: 'general earnings success' earning trajectories

Cluster 6 contains 1,156 people – equivalent to 2.6 percent of all Work Programme participants – and trajectories in this group mostly convey a single step from unemployment to higher-level earnings. On average, people in this group take just below 6 months to enter employment (5.8 months) and once employment has been achieved this is generally sustained for the remainder of the tracking period. This is reflected in the relatively high job outcome rates amongst group members at 90 percent.

Cluster 7 shows trajectories where a relatively speedy employment entry (the average time taken to enter employment for this group is 5.2 months) leads to more of a mixed-earning-level experience. The green 'middle-earning' elements dominate the right hand portion of the graph, suggesting that these participants are entering work that is part-time (less than 35-hours a week at national minimum wage) and for many that earnings levels change across the later portion of the window, with middle earning months often punctuated by one (or more) periods of higher earnings above the 35 hour NMW threshold. Although earnings for this group may be modest, with average earnings at £6,653 across the full RTI window, this group is particularly successful in securing official programme job outcomes, with 92 percent of group members doing so.

The trajectories in cluster 8 again express a pattern of 'general earnings success'. The initial transitions into paid work by members of this group are even quicker following attachment than for the visually similar cluster 7, suggesting that participants here are very much 'ready for work' at programme start. Indeed, some of the trajectories in this cluster show that participants entered paid work within their month of programme

attachment. Unlike cluster 5, however, the earnings profile for these participants appears to be one of progression to ‘higher earning’ elements, suggesting an increase in either hours and/or pay for group members across the tracking window. Compared to cluster 7, trajectories for members of this group show a greater preponderance of higher-earning months, although again month-to-month earning levels fluctuate. As with other trajectory types in this cluster there is a high conversion rate of employment entry to official employment outcomes and 89 percent of the 2,073 people in this cluster achieve payable job outcomes.

7.7.4 Super-cluster 4: ‘gold standard’

This super-cluster contains only one sub-group and describes the trajectories of 1,748 participants. Of all the super-clusters identified it is the smallest in terms of population is equivalent to just below 4 percent of the Work Programme population as a whole. Rapid transitions to employment are achieved, on average taking just 1.9 months for participants to enter paid work. Earnings are then sustained at stable, higher levels, shown through the on-going occurrence of ‘higher earning’ months throughout. There is a high (but not perfect) rate of job outcome achievement with 90% of group members attaining successful job outcomes.

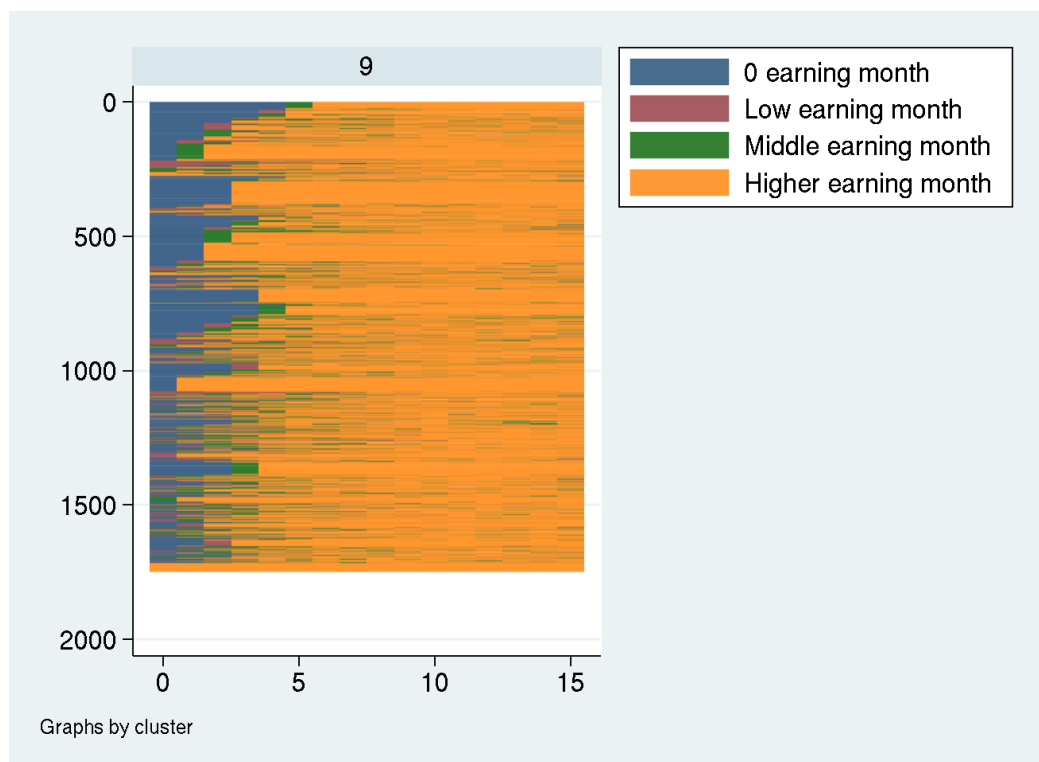


Figure 7.12 Sequence index plot for super-cluster 4: ‘gold standard’ earning trajectories

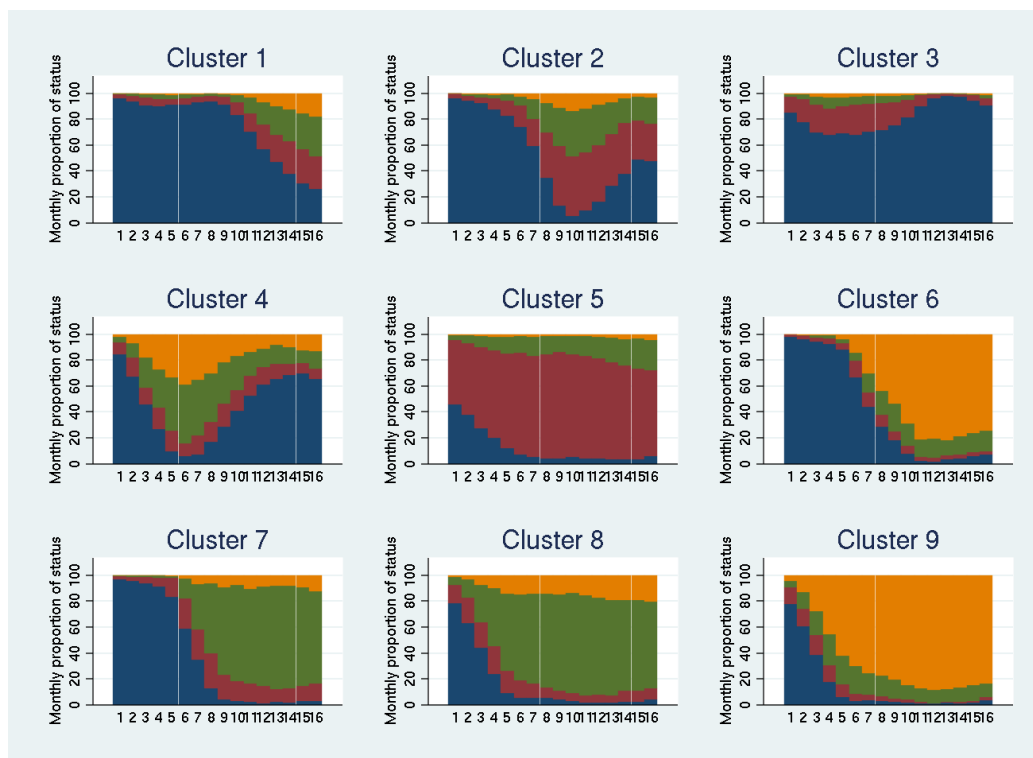


Figure 7.13 Cohort representation of earnings trajectory clusters for Work Programme participants

Cluster	Cluster size	Job outcome rate	Mean time to first earning (months)	Mean earnings (£)	Expected proportion of all WP attachments
1	6,042	56.1	9.6	2,611	13.6
2	1,255	68.8	5.7	3,987	2.8
3	3,252	24.3	3.6	1,171	7.3
4	1,080	71.3	2.2	7,216	2.4
Slow and unstable (summary)	11,629	50.0			26.2
5 'stuck in low pay'	2,264	53.3	1.3	4,421	5.1
6	1,156	89.8	5.8	11,285	2.6
7	1,129	91.9	5.2	6,653	2.5
8	2,073	89.3	2.0	9,912	4.7
General earnings success (summary)	4,358	90.1			9.8
9 'gold standard'	1,748	89.8	1.9	16,996	3.9
Grey non-RTI group		6.0	NA	NA	55

Figure 7.14 Employment and earnings trajectory clusters and relative size

7.8 Reflections and considerations on employment and earning trajectories

The sequence analysis appears to have been successful in terms of producing an analytically meaningful typology of earnings trajectories for its large random sample

of Work Programme participants. The identification of four distinct earning pathways – ‘slow and unstable’, ‘stuck in low-pay’, ‘general earnings success’, and ‘gold standard’ – and the relative proportions of Work Programme participants was considered an insightful and helpful contribution by policymakers and analysts in May 2016 who were involved in reflecting on the Work Programme and designing its successor scheme.

In considering these distinct earning trajectory experiences it is important also to recall the large monotone block of grey in the lower portion of Figure 7.4 representing over half (55%) of all Work Programme participants who do not feature in the RTI. This is a crucial policy weakness. The majority of Work Programme participants will not be in paid work at the end of their allotted time on the scheme and return to Jobcentre Plus to receive ongoing employment support once their 2-year period of attachment has ended.

Of those participants who do feature in the RTI data, as a result of experiencing at least 1-month of earnings from paid work during the tracking window, the analysis identifies two of the four overarching trajectory classes as fulfilling the intentions of programme designers where participants undertake transitions to paid work which then extend over lengthy time periods and which are associated with earning levels compatible with working more than 16 hours a week at NMW. The sequences illustrated by the ‘gold standard’ group of participants can be understood as the most ‘successful’ in terms of the ambitions both of the Work Programme and wider policy objectives around shifting the UK to a “higher wage, lower welfare economy” (Oakley, 2015, p. 4). Unfortunately, the trajectories in these two groups convey the minority experience for participants: only 13.7% of all Work Programme participants would be expected to follow these successful earning pathways.

Particularly concerning in relation to this chapter’s focus on *sustained employment beyond job entry* is the programme’s seemingly limited ability to convert entry to employment into sustained, ongoing earning pathways at levels compatible with non-claiming of unemployment benefits. Super-clusters 1 and 2 are both worrisome in this regard. The slow transitions of those participants in the ‘slow and unstable’ type are in themselves not troubling, indeed the intention of the programme’s prolonged attachment window is to stimulate provision even for those participants who may take some time to initially enter paid work. Rather, the issue is with the instability and disruption to employment once job entries do occur. Providers do not appear to be working effectively to enable temporary work placements to be ‘joined up’ or sustained. For many, temporary employment does not appear to have functioned as a stepping stone to more secure work, but rather offers a teetering journey of unstable earnings.

7.9 Policy implications and challenges in predicting cluster membership

Understandably, individuals in the two less-successful trajectory groupings will be of greatest interest to policymakers and service providers. In particular, the ability to

identify those programme participants who are most at risk of cycling between low-pay and no-pay, or who remain on a low-earning trajectory month-to-month, will be important in targeting and designing appropriate and effective additional supports (assuming that long-term earnings success remains the ultimate policy objective). Figure 7.15 offers a preliminary description of the demographic characteristics of programme participants who fall within each of the employment and earning clusters.

	1 'Slow and unstable'	2 'stuck in low-pay'	3 'general earnings success'	4 'golden ticket'	5 'no earnings record'	Full cohort
Male (%)	64	44	59	76	56	59
Age (mean)	33	41	33	36	39	37
Has a disability or health condition (%)	28	39	22	21	56	43
Has children (%)	28	33	28	22	31	30
Lone-parent (%)	12	21	14	5	10	11
Substance issues (%)	1	1	0	1	3	2
Offender (%)	6	1	4	7	6	6

Figure 7.15 Demographic characteristics of programme participants within trajectory groups

Relative to the two more successful earning trajectory types, those who are stuck in low-pay are more likely to be women, have children and be a little older. Lone parents are also more commonly found in this group. Offenders and other more vulnerable jobseekers who are eligible for membership of Payment Group 3 are more prevalent in the 'slow and unstable' group. People with a disability are most commonly found not to feature in the RTI dataset.

Following Dorsett and Lucchino (2012), statistical techniques are also used to explore whether there are any distinctive characteristics at the point of programme referral which could help predict an individual's future earnings trajectory type. Multinomial logit estimators are used to predict membership of high-level trajectory types and separate binary logistic models are also developed to predict membership of each of the overarching groups individually. The ambition here is not to establish causality but rather to assess the presence of strong correlation and hence predictive power sufficient to identify in advance who is at risk of an unsuccessful earnings trajectory.

This model development work is not shown since the models to predict membership for each of the earning trajectory groups have relatively weak explanatory power. A particular challenge in predicting membership of each of the trajectory clusters is that many of the available explanatory variables function in a similar way – i.e. with the same direction and similar effect size – for each of the earning trajectory groups. The implication here is that while there are characteristics which function as effective predictors of entry versus non-entry to employment, the same set of variables do not seem to offer discriminatory power across the different earning trajectory types. It may

be that with a richer set of explanatory variables, particularly with greater detail on employment history, the terms of previous employment, and qualifications, that stronger models could be built, capable of predicting which individuals are likely to enter each of the trajectory groups. If so, providers are better placed than commissioners to establish and use these data. It may be, however, that it is simply highly challenging to establish a set of explanatory variables that robustly predicts members of these longitudinal sustainment types.

7.10 Conclusions

By paying only for results within a Prime-provider quasi-market there is a consensus that the Work Programme has delivered broadly similar employment outcomes for a lower cost per participant than under previous programmes (WPSC, 2015). However, there is an important distinction between ‘cheap’ results and an appropriate incentive structure that has desirable effects in terms of steering provider behaviour and shaping outcomes achieved. To date the employment and earning experiences of Work Programme participants have only been explored through the high level information relating to the numbers of participants achieving official job outcomes and sustainment outcomes (DWP statistical releases) or crude counts of those entering employment (for example, ERSA, 2013). This chapter has for the first time used sophisticated longitudinal classificatory methods and uniquely rich earnings data to investigate the employment trajectories of participants. The empirical findings suggest that there may be a disconnect between the intended incentive structure and the experience of participants.

A key contribution is the use of Optimal Matching on a newly available dataset of month-to-month RTI earning records to examine employment and earning experiences in a holistic way and with greater sensitivity to *whole* transition pathways than afforded through conventional statistical approaches. Drawing on this method, and combining it with cluster analysis, enables the identification of nine distinct trajectory clusters which sit within four higher level categories. These groups of trajectories can be summarised as ‘slow and unstable’, ‘stuck in low-pay’, ‘general earnings success’, and ‘gold standard’ earnings. The latter two trajectory categories can be understood as embodying the characteristics desired by policymakers: they generally feature relatively quick transitions into sustained employment, with earning levels which each month tend to be above an income level equivalent to 35-hours per week at NMW. Although programme designers are likely to take some satisfaction from these ‘successful’ earning trajectories, these pathways in practice would only be expected to apply to less than 14 percent (13.7%) of people attached to the programme. By far the most common experience is never having entered paid work at all.

The first two trajectory groups offer particularly striking findings for a programme that has placed such considerable effort in designing a payment structure that rewards ‘sustained’ employment outcomes. These two groups – which likely account for nearly a third of participant experiences (31.3%) – show trajectories where participants enter

employment but experience either an unstable employment pathway or will be stuck in a low-earning employment pattern.

Whilst it is not possible to draw systematic comparisons with previous employment support schemes, judged by the ongoing presence of unstable and low-paid employment patterns the Work Programme's complex sustainment payment structure does not seem to be incentivising providers to develop provision and supports that enable entry to employment to be successfully converted to sustainable, well-paid employment trajectories (as is the programmatic objective). For most Work Programme participants no work is the outcome whilst for those participants who do successfully move into paid work the low-pay no-pay cycle all too often persists.

8 Making markets in employment support: does the variety of quasi-market matter for people with disabilities and health conditions?

8.1 Chapter summary

This chapter offers a more synoptic reflection on the Work Programme's private power market. It returns to the analytic framework on varieties of quasi-market developed in Chapter 2 and through an intelligently constructed quasi-experimental analysis assesses the employment and earnings outcomes of two alternatively configured quasi-marketised employment support programmes which have run in tandem in the British context since 2011: Work Programme and Work Choice.

8.2 Introduction: the uncomfortable gulf in job outcome rates for people with health conditions and disabilities

With the application of the provider-directed quasi-market configuration in the Work Programme there has been a large and persistent gulf in the employment outcomes for those participants who have a long-term health condition or disability and those others who do not (Figure 8.1). The experiences of people with health conditions and disabilities jars against the programmatic ambition to “ensure that providers have strong incentives to help *all* of their customers”, and close the performance gap between the easiest- and hardest-to-help (DWP, 2012, p. 6, emphasis added). The gulf in outcomes now also sits uncomfortably against the Government's manifesto commitment to tackle the disability employment gap and “get 1 million more people with disabilities into employment over the next ten years” (Conservative Party, 2017, p. 57). Voluntary sector organisations and think tanks have been united in arguing that the ambition for such high levels of employment cannot be reached on the basis of the current performance of welfare-to-work programmes (Disability Rights UK, 2016; Oakley, 2015; Purvis et al., 2014).

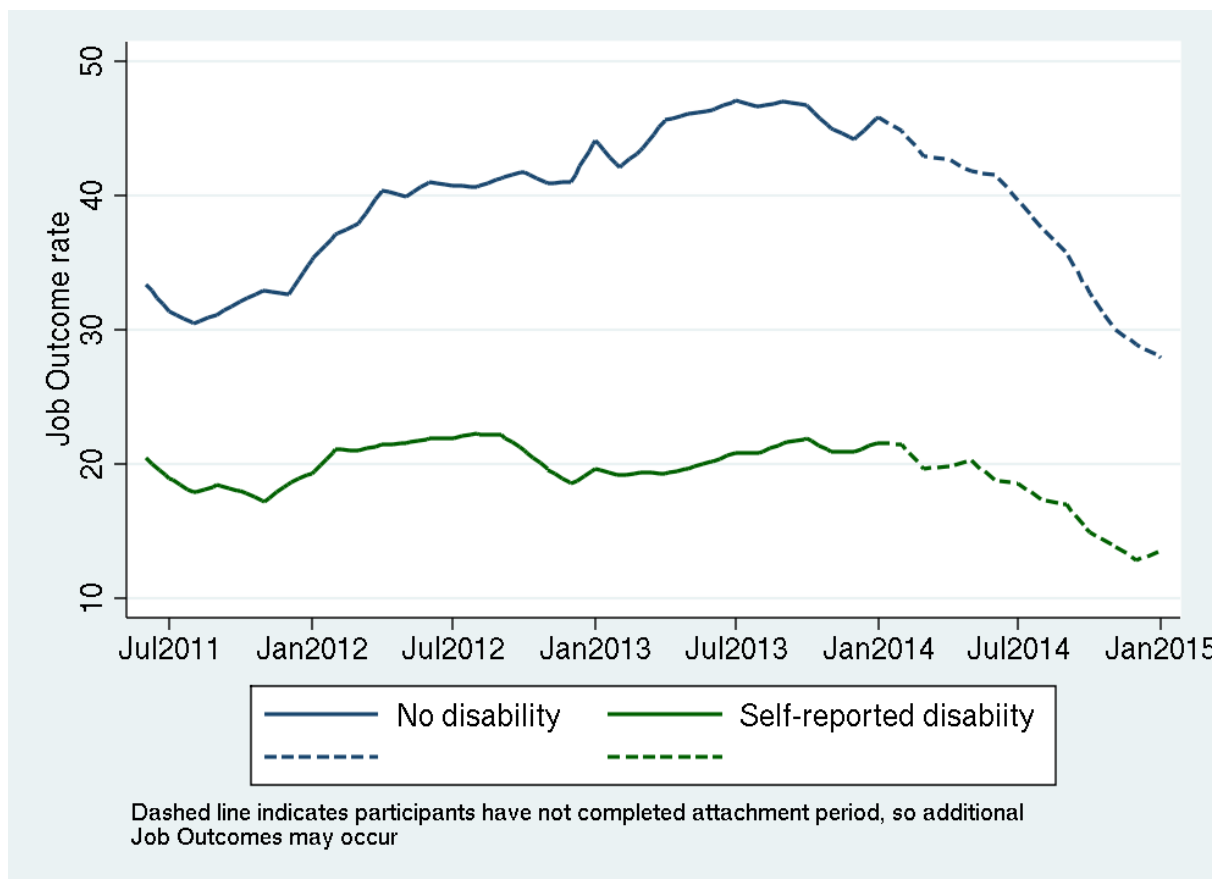


Figure 8.1 Job outcome rates for Work Programme participants over time, by self-reported health condition and disability status

Somewhat peculiarly, given policy-maker's confidence that the 'universal' Work Programme can effectively and simultaneously serve a large cohort of participants with a highly diverse set of characteristics, work experiences and employment support needs (DWP, 2012; WPSC, 2013), there is a parallel programme: 'Work Choice'.

Work Choice is a specialist disability welfare-to-work programme commissioned by the DWP and which runs in tandem to the Work Programme across Great Britain. Work Choice is a voluntary programme focused on those individuals whose health and disability-related support needs mean that standard Jobcentre Plus support is unlikely to be appropriate and is delivered by a range of employment support providers from the private, voluntary and public sectors. Work Programme and Work Choice are grounded in the same commissioning strategy, but demonstrate important differences in the degree to which they embrace highly-marketised provider-directed governance arrangements (Jantz et al., 2015) and particularly in the extent to which they implement payment-by-results principles. Given the mixed findings from studies of marketised outcome-based commissioning across policy fields, but particularly in the area of welfare-to-work, the twin track approach of the Work Programme and Work Choice provides a powerful opportunity to consider potential differences in the impacts of the two alternative quasi-marketised service configurations. This chapter implements a novel quasi-experimental analysis to investigate the implications of these alternate quasi-market formulations for delivery and outcomes, drawing on the original analytical framework advanced in Chapter 2.

8.3 The particularities of the quasi-market configuration embodied by the Work Programme and Work Choice

Both the Work Programme and Work Choice were commissioned subsequent to the DWP 2008 Commissioning Strategy, which was heavily informed by the work of Lord Freud. Despite being commissioned according to the same overarching strategy however, there are important variations in the design and consequent quasi-market configuration of the two schemes and it is this variation that this chapter exploits analytically and empirically.

8.3.1 Applying dimensions of variation as a comparative tool: *variation on the allocation dimension*

Returning to the analytical lens outlined in Chapter 2, the allocation dimension which relates to the financing and regulation of service provision is reprised in Figure 8.2. The first component conveys whether services are collectively financed by society or whether provision relies, in full, or in part on private resources of individuals. The second allocation component captures the degree and strength of regulation for provider activity.

There is no variation between the two programmes in terms of the formal financing component of the allocation dimension. As with all UK employment support programmes these two schemes are both collectively funded – hence the stacked position of the programme markers at the left-hand side of the top axis in Figure 8.2. The key driver of variation across the allocation dimension in UK activation programmes is therefore the regulation of service quality (Wiggan, 2015a).

Under the Work Programme service quality is intended to be held up by ‘minimum service guarantees’ (MSGs), which are designed and published by Prime providers themselves. As noted in earlier chapters, “weaknesses in the substance, consistency and, in some cases, even the possible enforceability of providers’ minimum service guarantees render these a far less useful and reliable protection than they could be” (Carter and Whitworth, 2015, p. 281). The National Audit Office has criticised the variability of MSGs and has suggested that in practice this system provides only limited safeguards for participants who may receive few services or experience ‘improper practice’ (NAO, 2012).

By contrast, in Work Choice, there is a high degree of clarity over contractual service minima. The Invitation to Tender (DWP, 2009) set out a contractual requirement for providers to devote a minimum number of support hours per week to each participant across the different service modules (see Figure 8.4 for detail on module service expectations). Programme evaluation material suggests that the minimum contractual requirements appear to have set high expectations in terms of provision, that frontline staff are highly aware of these support commitments, and that under the auditing approach providers appear to adhere to the high levels of structured support (Purvis et al., 2013).

The lower axis of Figure 8.2 conveys this divergence between the programmes by positioning a considerable distance between the Work Programme – which sits at the provider-directed, light touch end of the service regulation spectrum – and Work Choice – which is more congruent with state-directed standards and more robust quality assurance.

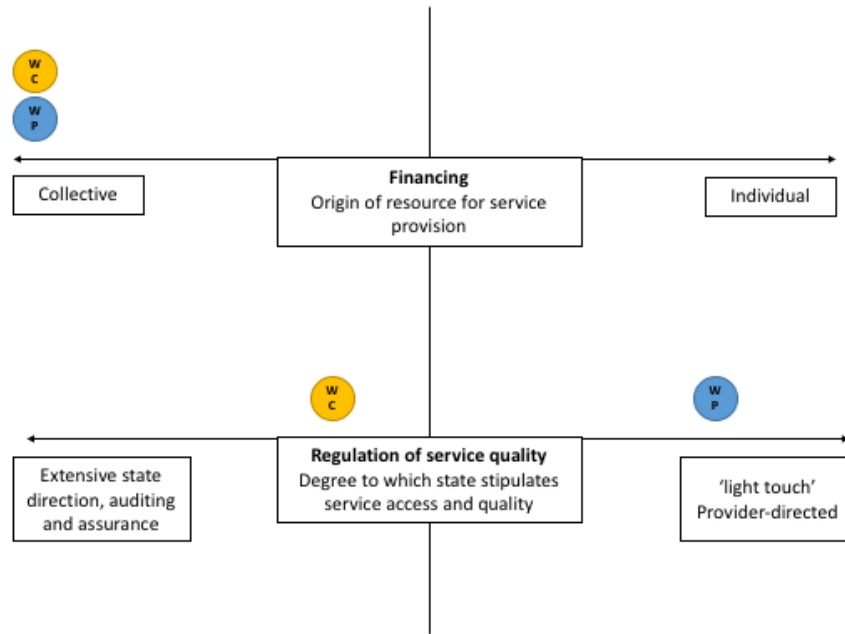


Figure 8.2 Visualising the allocation dimension for the Work Programme and Work Choice quasi-markets

8.3.2 Applying dimensions of variation as a comparative tool: *variation on the production dimension*

The production dimension covers choice and competition – the way that a service is ‘produced’ within the quasi-market (described more fully in section 2.7.2). The first sub-component (upper portion of Figure 8.3) relates to how ‘open’ the market is to new provider entrants and captures the degree to which the quasi-market arrangements service the hegemony of specific or incumbent providers. This spans from an open, highly contestable and accessible market (far left) to a situation where access is constrained or limited.

The second axis captures the control apportioned to the state versus service providers in the design and stipulation of provider activity. At the left-most extent, there is a high degree of public sector control over provider activity, since payment to providers is explicitly tied to the delivery of fixed service components. Shifting rightwards, the dimension portrays an increasingly outcomes-led approach whereby service providers are afforded the freedom to innovate and design flexible services geared to pursuing specified outcomes but where the public sector takes a position of agnosticism on the *means* of securing these outcomes.

The final axis in the lower portion of Figure 8.3 captures the function of user choice and voice within service production. Here, the left-most portion of the axis conveys a situation where service users are powerful agents whilst at the right-hand extent service user perspectives have no mechanism for shaping services. For each axis within Figure 8.3 a position closer to the extreme right-hand side reflects a stronger provider position be that through minimal competition (upper axis), minimal state control (middle axis) or minimal responsiveness to user preferences (lower axis).

Considering the relative positioning of Work Programme and Work Choice on the first component – that of market access – both programmes were procured through variants of a ‘Prime Contractor’ model which limits the direct contractual relationship between DWP to a relatively small number of top tier, ‘Prime’ providers who are then responsible for managing supply chains of provision (Armstrong et al., 2010; DWP, 2008). There are however, distinctions between the programmes in the degree to which particular forms and sizes of provider have a stronghold over provision as Prime contractors. Despite its smaller size (in terms of participant numbers, and hence, contract value), Work Choice has a greater number of Contract Package Areas (CPAs) and contracts were awarded to Primes in 28 CPAs compared to Work Programme’s 18. Where Scotland operates as a single CPA in the Work Programme, provision is divided across 4 CPAs in Work Choice for example. The tender process for Work Choice was preceded by a pre-qualification questionnaire but unlike Work Programme this had no hard thresholds in terms of capital requirements or turnover. This, in conjunction with the smaller contract sizes meant that, compared to Work Programme, a wider range of organisations were in a position to compete for contracts.

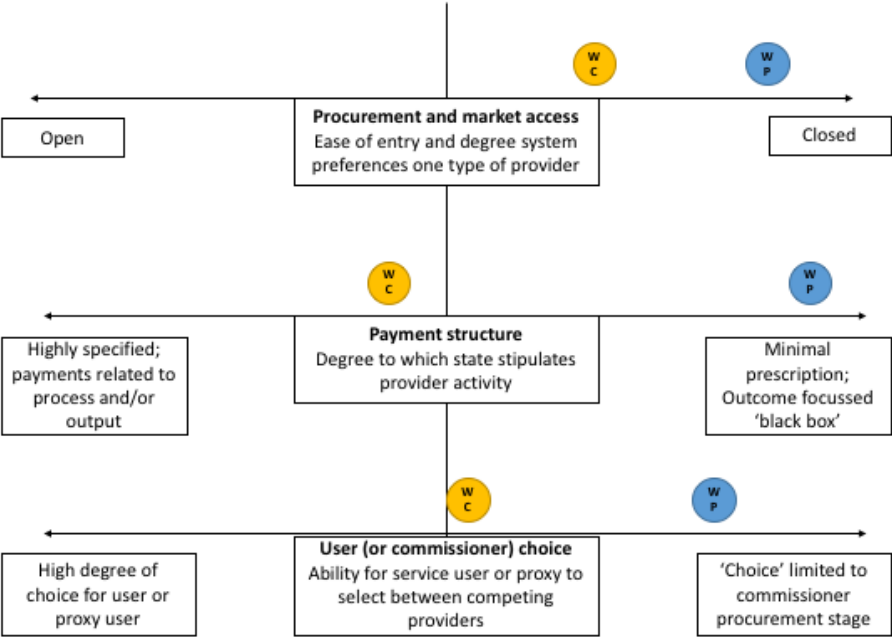


Figure 8.3 Visualising the production dimension for the Work Programme and Work Choice quasi-markets

The size and stipulations associated with Prime contracts in the two schemes has clear implications for the types of providers involved. Of the eight organisations who successfully became Work Choice Prime providers four are from the private sector and four are from the voluntary sector.¹² When considering the number of contracts held by these Primes, over 70 percent of provision is led by voluntary sector providers (Post Tender Discussion Documentation in Thompson et al., 2011). For the Work Programme there were considerably tighter requirements for bidding as the DWP sought to ensure that Prime contracts were held by organisations capable of financing upfront investment in services and shouldering the significant financial pressures of back-ended outcomes payments at a large scale.

The Work Programme is therefore positioned as comparatively more ‘closed’ on the market access dimension (upper portion Figure 8.3) since tendering requirements reduced competition for contracts and situated existing private sector providers (and those on the ERSS framework) in a dominant position to tender for future contracts. Work Choice is comparatively more open, although the long contracting periods do inculcate some degree of ‘lock in’.

Both programmes implement outcome-based funding arrangements for providers but there are key differences in the payment models for the two schemes, most notably in the extent to which payment-by-results dominates the payment profile. Work Choice providers receive a monthly service fee equivalent to 70 percent of their contract price with around 30 percent of programme funding being contingent on outcomes (Thompson et al., 2011). Purvis et al. (2013) note that in Work Choice there is a tension between the Commissioning Strategy principle of minimum service prescription intended to facilitate a flexible and personalised approach and the need to guarantee robust minimum levels of service delivery, which is seen as the quid pro quo for the substantial service fee element of the programme. Under Work Choice the commissioning Department has established a highly specified set of modular components which offer a high level of structured contact between participants and service providers whilst aiming to ensure progression towards the ultimate goal of unsupported employment, where this is appropriate (DWP, 2009; Purvis et al., 2013, see Figure 8.4). This balance between prescribed components and provider discretion can be considered a ‘grey box’.

¹² In practice, one of the organisations classed as ‘private sector’ is a special purpose vehicle created by two voluntary sector organisations.

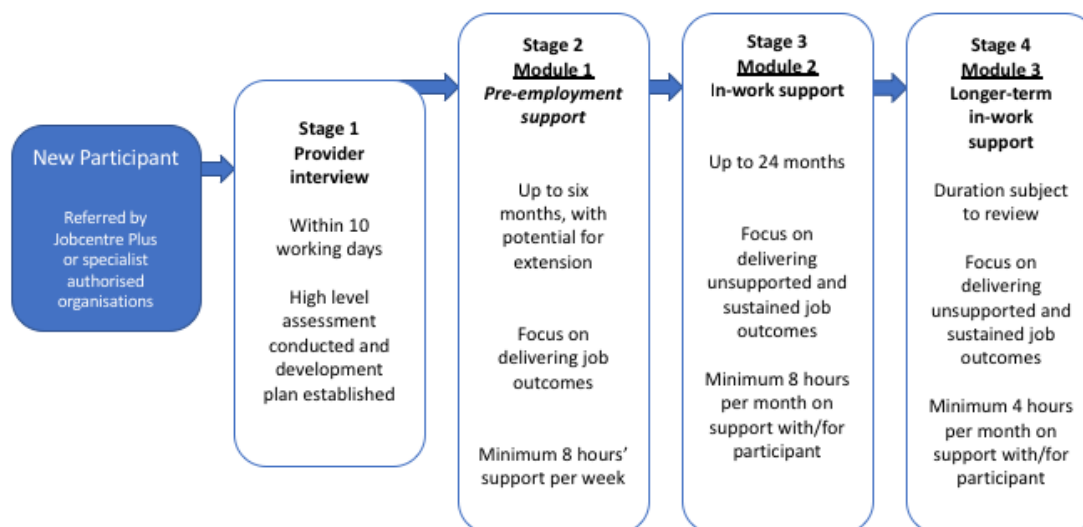


Figure 8.4 Contractually specified modular intervention structure of Work Choice (adapted from Purvis et al., 2013)

The Work Programme holds a much more extreme position on this dimension due to its full embrace of the ‘black box’ delivery model whereby Primes have almost complete discretion over the nature and extent of their intervention and where provider payments are based almost entirely on job outcome results (Rees et al., 2014). To counter tendencies for providers to neglect those participants whose barriers to work are greater (i.e. where payable job outcomes are less likely to be forthcoming) the DWP relies on its differential payment-by-results structure. In the Work Programme, over the course of the contract the ratio of sustained job outcome fees to attachment fees is intended to be 80:20 (Wiggan, 2015a). Importantly, however, the fixed attachment fee reduces to zero across the early years of the programme so since 2014 the Work Programme has effectively been operating under ‘pure’ payment-by-results, where no portion of provider payment is guaranteed. Correspondingly, Work Programme is positioned at the most extreme right-hand position on this axis, whilst Work Choice sits in a softened hybrid position.

On the final dimension relating to the function of choice there is again deviation between the two programmes. For service users, Work Choice is a voluntary programme whilst the majority of Work Programme participants are mandated to engage as a condition for their ongoing receipt of unemployment benefits (though importantly, people with long term health conditions who are understood to be distant from the labour market may be afforded the option of voluntary participation). In the main, choice in terms of programme exit is not an option if people wish to protect their income. User choice within the market *between* providers has been almost non-existent as a feature in UK welfare-to-work provision (Wiggan, 2015a). Under the Work Programme participants are randomly allocated to a Prime provider. Under Work

Choice, participants do have a ‘choice’ between Remploy¹³ provision and the single contracted Prime provider in their region, although the official programme evaluation suggests that Jobcentre Plus Disability Employment Advisors will be involved in this process and it is not clear that user choice exerts influence on provider practice.

8.4 Pure provider-directed or hybrid forms of quasi-market?

Along these dimensions then the Work Programme clearly sits as the embodiment of the ‘provider-directed’ market (sitting as it does at the right-hand edge of each of the axes in Figure 8.3). Indeed, as discussed in Chapter 3 the Work Programme specifically embodies a private power market.

Contrastingly, although Work Choice was also commissioned according to Freud’s guiding principles it adopts a softer, less-extreme position as a ‘provider-led’ market formulation. Across the allocation dimensions, Work Choice bolsters the state’s power to regulate service quality (a preference of service users). Within the production axes, Work Choice leans *towards* a provider-directed configuration. The Prime provider model, outcomes-related payment, and grey box specification all point to a strong position for providers, and yet it fails to sit squarely at the ‘provider-directed’ right-hand portion of these axes. The state has a much stronger role in stipulating activity and service users have *some* degree of choice (a double choice as to whether to participate and which of two service providers to work with). Where Work Programme is a full-throated version of a private power market, almost exclusively reliant on accountability levers of price and competition, Work Choice offers ‘Freud-lite’. Work Choice takes a hybridised form – where provider prerogatives are mediated by bolstering *both* the state- and user- preferences – and does not neatly sit within the brackets of Gingrich’s framework.

Intriguingly, although we cannot attribute *causation* to this softer form of provider-direction in the Work Choice configuration, concerns in relation to service quality, creaming and parking of clients and market stability have not dogged Work Choice in the same way that they have plagued the Work Programme. But it is to outcomes that the chapter’s empirical focus turns. Given the persistent performance critiques of the Work Programme in its support of participants with health conditions and disabilities (and, indeed, with other wider more challenging barriers, Carter and Whitworth, 2015; Newton et al., 2012; WPSC, 2013) the parallel operation of Work Choice offers something of a quasi-experiment through which to assess whether the softening of provider-directed marketisation levers can be seen to have led to differing experiences for participants as illustrated by the patterning of employment outcomes. The key empirical question informing the remainder of this chapter is whether these two programmes (embodying Freud-inspired provider-directed marketised regimes to

¹³ Remploy is a non-departmental government body which is directly funded to deliver Work Choice by DWP under Grant in Aid arrangements and offers Work Choice provision nationwide alongside the contracted providers. Historically Remploy has specialised in supported employment provision and supported employment routes are removed from the subsequent empirical analysis.

different degrees) are differently effective for participants who providers may view as 'harder to help' because of their disabilities or limiting health conditions.

8.5 An empirical exploration of different quasi-market configurations for employment and earning outcomes

The implication of the distinct quasi-marketised configurations which are implemented within Work Choice and the Work Programme is that these alternate formulations may exert different pressures and incentives on the allocation and production of services for similar cohorts. This chapter for the first time offers an innovative quasi-experiment using rich RTI earnings data to present a cross-programme comparative analysis of the employment and earnings outcomes achieved by a matched sample of participants in these two distinct quasi-marketised formulations.

Dimension	Work Programme	Work Choice
Size (participants)	c. 1 million in year 1	c. 20,000 per year
Eligible participants (on paper)	All long-term unemployed people or those at risk of becoming so including those with disabilities and LTCs	People with disabilities who cannot be effectively supported into work through mainstream employment programmes or Jobcentre Plus provision
Participants (in practice)	As above, sizeable Jobseekers Allowance population who self-identify as having a disability	Restrictions on referral numbers and less focussed on those with 'most significant needs'; majority of participants are claiming Jobseekers Allowance
Compulsion	Largely mandatory, sanction backed	Voluntary, though potential for some participants to opt for WC in light of upcoming WP mandation
Outcomes and objectives	'Sustained' job outcomes (13 or 26 weeks) with monthly sustainment payments	Sustained, 'unsupported' job outcomes

Figure 8.5 High-level programmatic comparison of Work Programme and Work Choice

The comparison of Work Choice and Work Programme can be understood as a natural experiment since although *on paper* the programmes targeted different cohorts of service users in *practice* ambiguities and tensions in the referral processes for the two programmes mean that participants in each scheme are much more similar in terms of their disability-related employment challenges than when the programmes were initially conceived. Figure 8.5 provides a summary of programme characteristics and eligibility criteria. The empirical analyses bolster this overlap further with the use of propensity score matching techniques.

Many Work Programme participants are facing multiple and complex barriers to employment, including significant disability-related support needs (Newton et al., 2012; Purvis et al., 2013). The official Work Choice evaluation has identified “a number of situations where disabled people with complex support needs, who might have been suitable for Work Choice support, were being referred to the Work Programme” (Purvis et al., 2013, p. 145). In parallel, there has been slippage in the degree to which Work Choice participation has been targeted at those with the most profound disability-related employment support needs: the majority of Work Choice participants are claiming Jobseekers Allowance rather than a specialist, disability-related benefit as DWP expected at the programme design stage. Each of these programmes has experienced stretch and tension in relation to referral routes and eligibility criteria and this leads to a situation where there is marked overlap in terms of the characteristics displayed by participants attending each of the schemes.

Thus, in summary, there is a group of people who are claiming out of work benefits, have health and disability-related barriers to entering the labour market, and who could

have been referred to either scheme. This partial overlap of participants with similar characteristics at the intersection of the two programmes unlocks the potential for quasi-experimental analysis and specifically raises the ability to use propensity score matching for the construction of an artificial comparator group. Considering the Work Choice participants as the ‘treated’ individuals, matching is used to find a non-participant from the control group of Work Programme participants with the most similar observed characteristics possible (Gertler et al., 2016). This method has been used widely in the evaluation of British welfare-to-work policies; has an intuitive appeal arising from the way it mimics random assignment through the construction of a control group post hoc; and is based on fewer assumptions than conventional regression based approaches (Bryson et al., 2002).

Access to participant-level administrative data held by the DWP has uniquely facilitated the analysis presented here. The evaluation uses administrative data for all individuals attached to the Work Programme and to Work Choice from April 2013 to September 2014. The participant level information is based on National Benefit Data records coupled with the specific administrative datasets for each of the programmes. P45/P60 records are used to construct pre-programme employment histories. This delivers a rich set of matching covariates including pre-programme benefit claim detail and employment durations, fixed individual demographic characteristics, and time coded programme attachment details used to ensure that participants in both treatment and control groups are experiencing interventions in a shared macroeconomic climate. The logistic regression model used to produce the propensity scores is shown in Figure 8.6.

The requirement for a standardised outcome metric for treatment and control participants is facilitated through the author’s unique academic access to HMRC RTI which provides immediate, regular data on employee earnings (Tarr and Finn, 2012). For the sample of Work Programme and Work Choice participants this data holds a record of gross earnings received by each participant, each month, for the month of attachment to their respective programmes and for 15 consecutive months.

Historically an ‘off-benefit’ measure has been used within DWP programme evaluations but the earnings data presented here offer a more subtle analysis that more closely proxies programme objectives around employment and earnings rather than benefits off-flows. For both Work Choice and Work Programme the overarching intent is to secure stable employment for participants that endures for the long-term and this objective occurs in a wider policy context which aspires to shift the UK to a “higher wage, lower welfare economy” (Oakley, 2015, p. 4). Hence paid employment which is both more stable and higher paying is the desirable policy outcome across both programmes. Therefore, both the number of months with earnings (where a greater number of months with recorded earnings is the preferable outcome) and earnings value (where greater income levels are seen as superior) serve as appropriate metrics to capture programme ‘effect’ in a comparable manner across the two interventions.

After data cleaning, the preparation of standardised variables and the removal of those Work Choice participants who were in supported employment, there were 405,417 Work Programme cases and 23,916 Work Choice cases available for analysis.¹⁴ A range of information relating to individual demographics, household characteristics, employment and benefit claiming histories and local context are used to produce the final model to predict Work Choice participation and the propensity scores (Figure 8.6). There is a general consensus in the propensity score matching literature that either a logit or a probit model may be used to estimate treatment (1 or 0) and that there is no strong advantage to using one or the other (Heinrich et al., 2010). A logit model is used here.

¹⁴ Within Work Choice there is a pathway to payable outcomes for ‘supported employment’ and those participants who are routed to this supported employment outcome (3,255 individuals) are excluded from the analysis. It is not deemed appropriate to compare job outcomes achieved in the open economy to employment income derived from a supported business; the analysis rules out any additional employment or earnings effect from Work Choice being associated with supported or ‘sheltered’ employment.

	Odds Ratio	Standard Error	P value
<i>Self-reported disability (binary; reference: 0)</i>	14.427	0.470	0.000
<i>Days with active ESA claim in 2-year prior to attachment</i>	0.991	0.000	0.000
<i>Days with active JSA claim in 2-year prior to attachment</i>	1.002	0.000	0.000
<i>Days in paid employment in 2-year prior to attachment</i>	1.003	0.000	0.000
<i>Days with active ESA claim in 5-year prior to attachment</i>	1.002	0.000	0.000
<i>Days with active JSA claim in 5-year prior to attachment</i>	1.002	0.000	0.000
<i>Number of JSA spells in 2-year prior to attachment</i>	1.120	0.007	0.000
<i>Number of ESA spells in 2-year prior to attachment</i>	0.554	0.009	0.000
<i>Number of paid employment spells in 2-year prior to attachment</i>	0.895	0.005	0.000
<i>Mixed ethnicity (reference: white)</i>	0.844	0.062	0.020
<i>Asian or British Asian ethnicity (reference: white)</i>	0.707	0.034	0.000
<i>Black or black British (reference: white)</i>	0.923	0.043	0.086
<i>Chinese or other ethnic group (reference: white)</i>	0.724	0.056	0.000
<i>Unknown ethnicity (reference: white)</i>	1.127	0.049	0.006
<i>Disadvantaged due to drugs, alcohol or ex-offender status (binary; reference: 0)</i>	0.167	0.013	0.000
<i>Cumulative measure of limitations associated with disability</i>	1.185	0.008	0.000
<i>Male</i>	1.158	0.021	0.000
<i>Age</i>	0.978	0.001	0.000
<i>Has dependent children (reference: no dependent children)</i>	0.442	0.013	0.000
<i>Lone parent (reference: not lone parent)</i>	2.688	0.137	0.000
<i>Has ICD code (reference: no ICD recorded)</i>	12.132	0.337	0.000
<i>Local Authority dummies not shown</i>			
<i>N</i>	429,333		
<i>Pseudo R²</i>	0.459		

Figure 8.6 Logistic regression model to predict Work Choice participation

OLS models to predict earnings value and time in employment are also constructed and demonstrate that the set of explanatory variables are also associated with employment outcomes. This test ensures that the explanatory variables are associated with both ‘treatment’ and ‘outcomes’ and is a key criterion for propensity score matching. The large number of significant variables in the propensity score model indicates that there were important differences in the composition of Work Programme and Work Choice programme populations, and Dorsett (2004) has argued that this highlights the potential for matching to be an effective approach. The final model to predict participation in Work Choice provision has a Pseudo R² value of 46% and can be seen as a relatively powerful model. The analysis therefore proceeds on the premise

that although the final propensity score model is unlikely to have controlled exhaustively for differences between Work Choice and Work Programme participants there is little remaining unobserved heterogeneity between the groups that is systematically correlated with participation *and* outcomes.

8.5.1 Propensity scores and matching algorithm

The predicted probability of participating in Work Choice is used as the propensity score for each participant in the Work Programme ‘control’ and Work Choice ‘treatment’ groups. The distribution of propensity scores is shown in Figure 8.7.

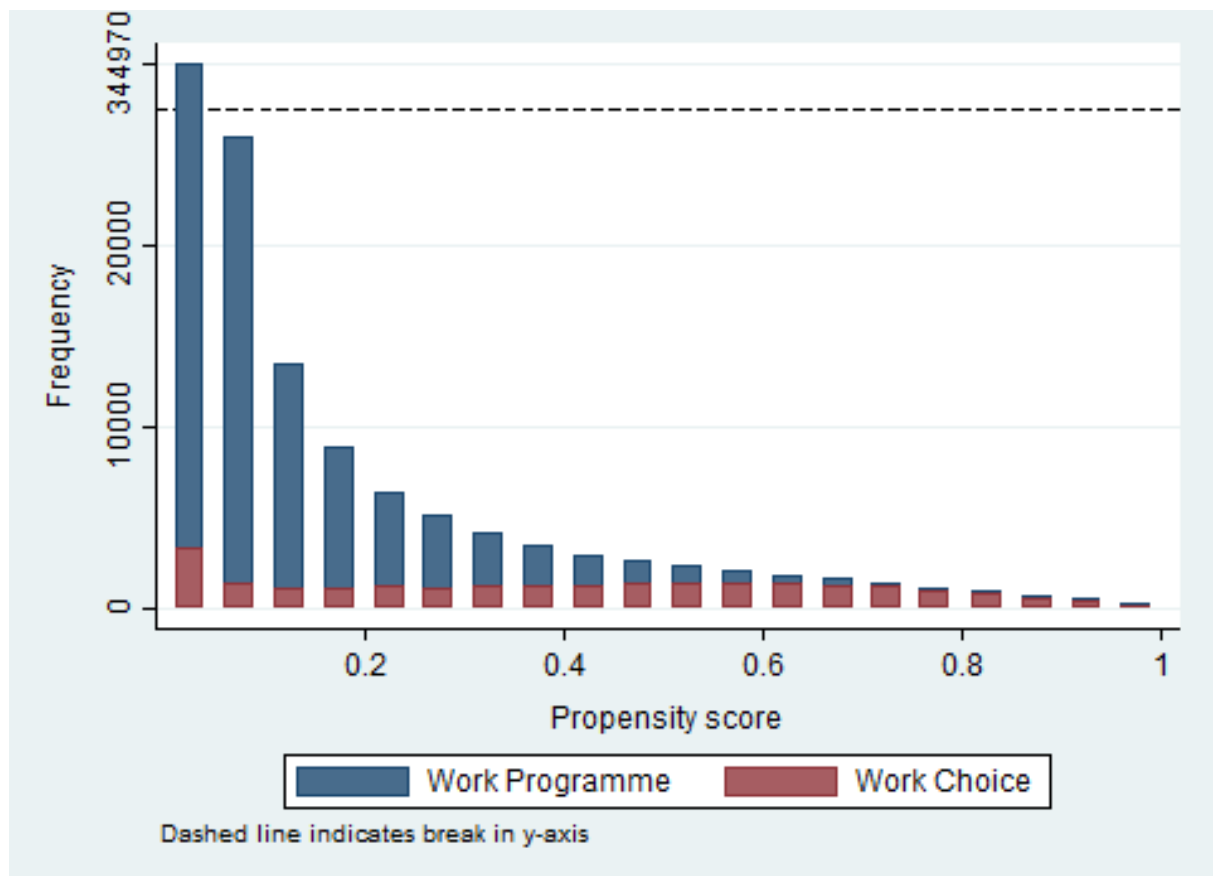


Figure 8.7 Distribution of propensity scores for Work Choice and Work Programme participants

Once the propensity scores have been produced for each case there are a number of alternative mechanisms by which the process of matching can take place. Studies reveal that in practice the choice of matching algorithm makes little difference to the resultant estimates of effect size (Smith and Todd, 2002 in Dorsett, 2004). Dehejia and Wahba (2002) found that enabling control participants to be used as matches to multiple treated individuals through replacement improved the performance of the match. Using ‘replacement’ in the match of Work Programme participants to the Work Choice treatment group is particularly relevant, since at the highest levels of propensity score (seen in Figure 8.7) there are a greater number of treatment individuals than there are control pool match candidates.

The 1:1 nearest neighbour match with replacement is implemented using the `psmatch2` programme (Leuven and Sianesi, 2003) in Stata. Following convention, those treatment cases that have propensity scores falling beyond the maximum extent of corresponding propensity scores for the comparator group are discarded. This results in the removal of just 53 cases from the analysis. This constitutes a low level of rejection (other studies have removed close to one-third of treated cases) and is not a concern regarding the representativeness of the results.

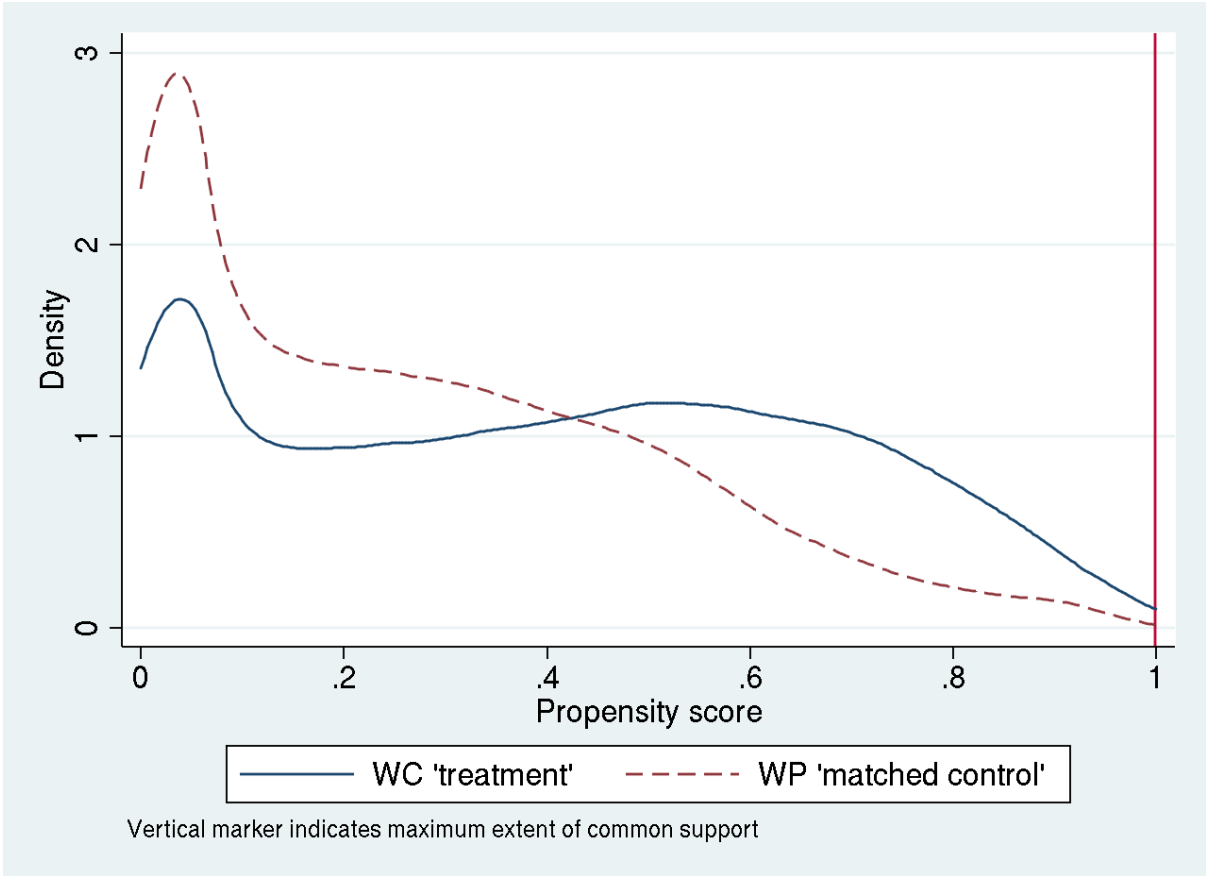


Figure 8.8 Propensity scores of Work Choice ‘treatment’ and matched ‘Work Programme ‘control’ group

8.5.2 Quality of the ‘match’

Figure 8.8 shows the distribution of propensity scores *post-match*, that is, only the treatment group cases (with support) and their matched control group. As expected, it can be seen that the density of control group cases is higher at lower propensity scores. To facilitate the closest match some Work Programme individuals have been used multiple times and this has occurred most often at higher levels of propensity score. Dorsett (2004) notes that the crucial aspect of the performance of the match is the degree to which it achieves balance across the variables included in the propensity score model. This is effectively a summary of the degree to which the synthetic control group expresses, in aggregate, the same characteristics as the retained treatment population.

Variable	Sample	Treated WC (mean)	Untreated WP (mean)	% bias	% reduction bias	t-test	p>t
Self-reported disability	Unmatched	0.94	0.43	128.7		156	0.000
	Matched	0.94	0.92	3.1	97.6	5.11	0.000
Days ESA (2-year)	Unmatched	16.36	135.10	-70.6		-81.03	0.000
	Matched	16.39	19.83	-2	97.1	-5.03	0.000
Days JSA (2-year)	Unmatched	259.78	249.78	4.4		7.18	0.000
	Matched	259.74	300.30	-18	-305.6	-16.94	0.000
Days employment (2-year)	Unmatched	197.22	105.59	41.9		78.2	0.000
	Matched	197.12	180.31	7.7	81.7	7.13	0.000
Days ESA (5-year)	Unmatched	339.97	389.45	-8.7		-12.36	0.000
	Matched	339.90	308.11	5.6	35.7	7.21	0.000
Days JSA (5-year)	Unmatched	524.11	384.08	34.7		58.56	0.000
	Matched	524.01	584.82	-15.1	56.6	-13.58	0.000
JSA spells (2-year)	Unmatched	1.74	1.87	-8.6		-11.13	0.000
	Matched	1.74	1.78	-2.8	67.3	-2.74	0.006
ESA spells (2-year)	Unmatched	0.49	0.71	-26.2		-35.13	0.000
	Matched	0.49	0.46	4.7	82	6.26	0.000
Employment spells (2-year)	Unmatched	1.65	1.36	13.6		21.07	0.000
	Matched	1.65	1.49	7.4	45.9	8.48	0.000
Ethnic groups	Unmatched	4.66	5.39	-4.1		-6.01	0.000
	Matched	4.66	4.94	-1.6	62.2	-1.73	0.084
Disadvantaged due to drugs, alcohol or ex-offender status	Unmatched	0.01	0.07	-30.9		-35.73	0.000
	Matched	0.01	0.01	-0.3	99	-0.71	0.480
Daily limitations of disability	Unmatched	2.27	0.98	91.3		133.47	0.000
	Matched	2.27	2.41	-9.3	89.8	-9.2	0.000
Male	Unmatched	0.65	0.59	12.5		18.47	0.000
	Matched	0.65	0.67	-4.8	61.3	-5.39	0.000
Age	Unmatched	38.43	36.76	12.8		19.57	0.000
	Matched	38.43	37.94	3.8	70.1	4.32	0.000
Has dependent children	Unmatched	0.11	0.22	-31.4		-42.07	0.000
	Matched	0.11	0.09	4.9	84.5	6.54	0.000
Lone parent	Unmatched	0.04	0.10	-24.5		-31.36	0.000
	Matched	0.04	0.03	2.5	89.8	3.75	0.000
Has ICD code	Unmatched	0.68	0.28	87.2		132.92	0.000
	Matched	0.68	0.63	11.7	86.6	12.31	0.000

Figure 8.9 Comparison of baseline and post-match characteristics across treated and untreated groups

T-test based comparisons after propensity score matching are controversial and are particularly challenging in situations of very large sample sizes. Whilst there do appear to be some residual differences between the characteristics of the two groups post-match (shown in Figure 8.9) the residual differences are considered to be small enough not to undermine the analysis. Post-matching, the Work Choice cohort on average seem to have spent slightly less time with active out of work benefit claims within the preceding two-years and in aggregate have a lower average value than the Work Programme control group on the cumulative disability-related barriers variable. A higher proportion of the Work Choice group do have formal record of an ICD code, formalising their disability or health condition.

8.6 Findings: Employment and earning effects across the two programmes

The estimated employment and earning effects of Work Choice compared to Work Programme participation is presented in Figure 8.10. There are two outcome variables: firstly, the average value of employment income earned within the 15-month tracking window; and secondly, the number of months, on average, individuals are in paid employment within the same time frame.

Prior to the propensity score matching process, the average earnings of Work Choice participants and the *full* Work Programme cohort are fairly similar. Work Choice participants on average earned around £336 more than Work Programme participants over the tracking period. For the *unmatched* cohorts, that is, comparing Work Choice participants to the full Work Programme cohort there is only a very small difference in the average number of months in paid work between the two groups. Work Choice participants are on average in paid work for 1.3 more days than the full Work Programme population. The standard error for the time in employment effect is nearly as large as the effect itself, implying that there is no substantive difference. When considering the employment performance for the programme populations *as a whole* then – all Work Choice compared to the full Work Programme – there is little to separate the two interventions in terms of the employment effects achieved by participants.

Comparing the performance of the two schemes in aggregate however, is not appropriate given what is known about the Work Programme's highly diverse participant population, and the scheme's known abilities to deliver job outcomes for those participants who are closer to the labour market. The real question of relevance here is the degree to which these alternate programmes and quasi-market configurations have been successful in supporting those with longer-term health conditions and disabilities.

Crucially, comparing the employment effects of the two programmes *after* matching a markedly different picture emerges. When the analysis is focused on those who are similar in terms of their employment histories and health-related characteristics (as per the full list of variables in Figure 8.6), the value of average earnings and the amount of time in employment for Work Choice participants within common support is

significantly higher than that demonstrated by their matched Work Programme counterparts. Work Choice participants, on average, earn £1,288 *more* than the matched Work Programme control group within the tracking window. Since average earnings for the matched Work Programme population are only £1,559 (within the 15-month window) this is a sizeable effect. The earnings level of Work Choice participants is 183% of the typical earnings for the Work Programme matched control group.

Variable	Sample	Treated (WC)	Controls (WP)	Difference	S.E.	T-stat
Earnings (£)	Unmatched	2848.57	2511.69	336.88	30.97	10.88
	ATT	2846.17	1558.66	1287.51	60.68	21.22
Time in employment (months)	Unmatched	3.71	3.56	0.15	0.03	4.28
	ATT	3.70	2.48	1.22	0.07	17.55

Figure 8.10 Employment effect of Work Choice compared to Work Programme

Given such considerable differences in the average employment and earnings effects experienced by Work Choice participants compared to matched Work Programme participants questions are immediately raised as to the manner by which these improved outcomes are achieved.

We can begin to consider this by exploring the four potential routes through which aggregate outcomes can be raised: is this through a greater proportion of Work Choice participants entering employment? Speedier transitions into employment? Different earning levels within work? More stable employment trajectories? Or any combination of these features? Here again the approach of sequence analysis using Optimal Matching and cluster analysis offers a novel route to investigate potential variation across the patterns of employment and earning trajectories experienced by participants in the two schemes. The same definition of policy-relevant ‘element types’ is used as in the preceding chapter.

The chart below (Figure 8.11) presents the matched Work Programme and Work Choice samples through sequence index plots. This representation plots entire individual earning trajectories horizontally, meaning that each horizontal stripe when read from left to right represents the 16 monthly activity statuses experienced by one individual.

The majority of participants in both Work Programme and Work Choice are without any earnings record in the RTI (see the upper portion of Figure 8.11). Notably, a larger proportion of Work Programme participants are without earnings than their Work Choice counterparts (67% of matched Work Programme participants have no earning record, compared to 59% of the Work Choice cohort). So, in part the greater employment effects for the Work Choice programme can be understood as being driven by the greater number of participants on that scheme who do ultimately end up achieving paid employment. Work Choice participants do also enter employment more

quickly following programme attachment, illustrated by the sharper gradient for colour changes. On average Work Choice participants enter employment after 4.4 months following programme attachment whilst matched Work Programme participants on average take 5.1 months.

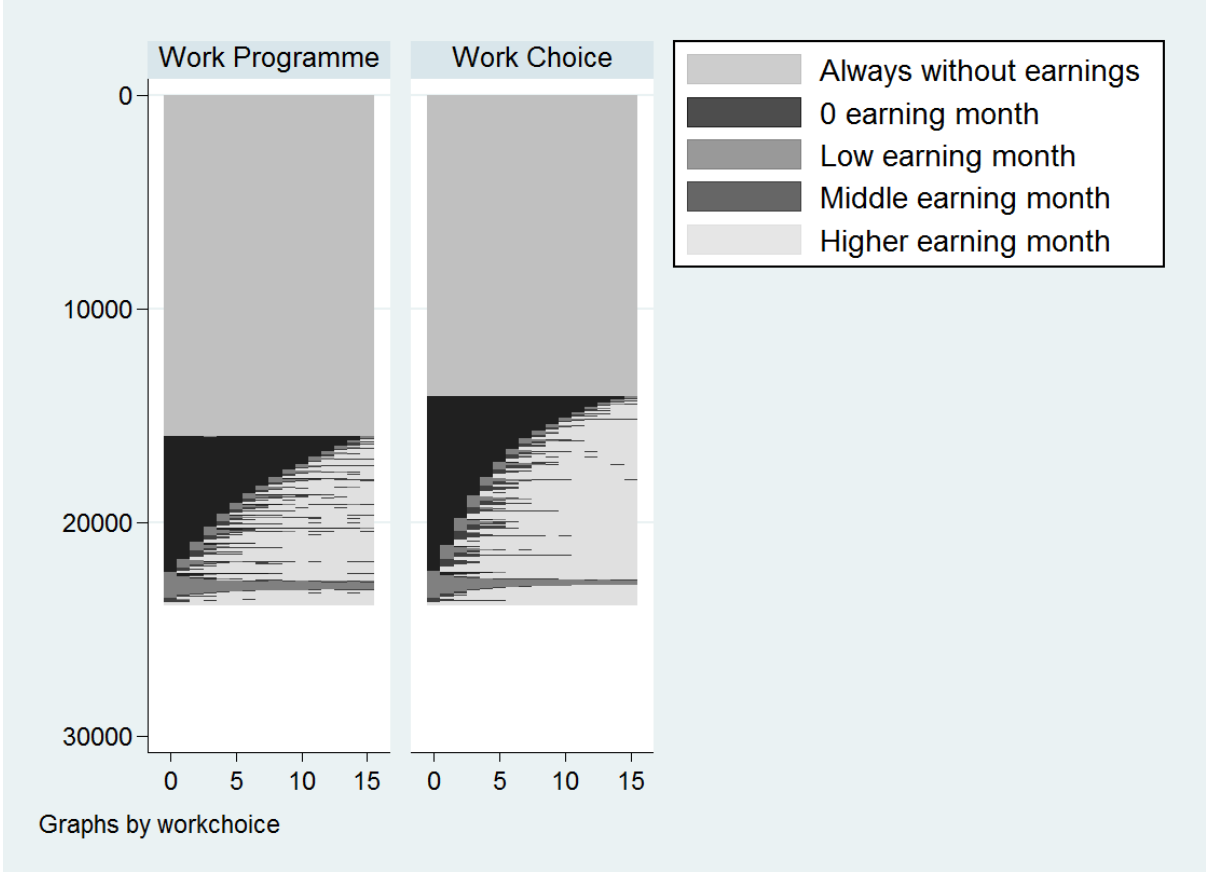


Figure 8.11 Graphical representation of earning trajectories for matched Work Programme and Work Choice participants

The value of average monthly earnings amongst Work Choice participants is also higher. Across the months with non-zero earnings, the mean monthly earnings for Work Choice participants is slightly higher (at £772.84) compared to the average monthly earnings for Work Programme participants (£631.04). Implementing Optimal Matching and cluster analysis on the combined Work Choice and matched Work Programme control group identifies six earning trajectory clusters which are nested within three ‘super-clusters’. These are described below:

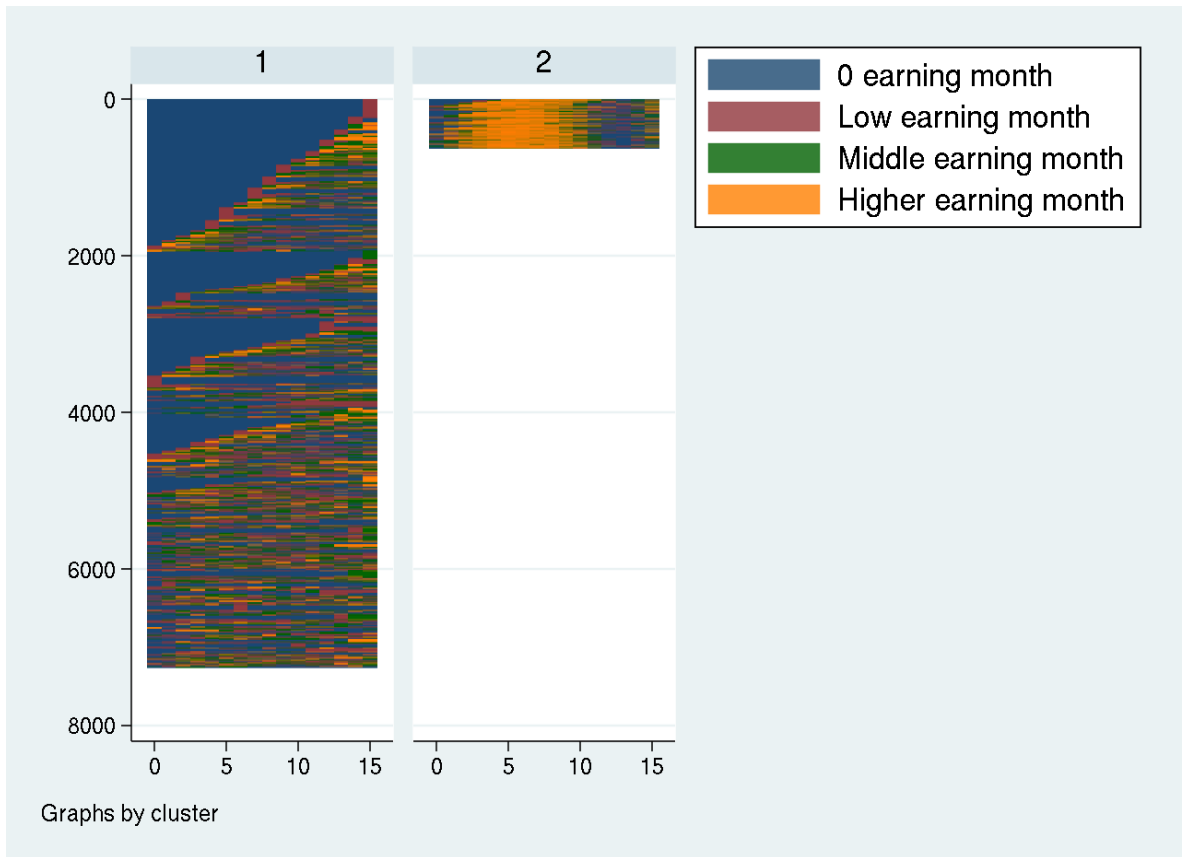


Figure 8.12 Super-cluster 1: slow, chequered and unstable

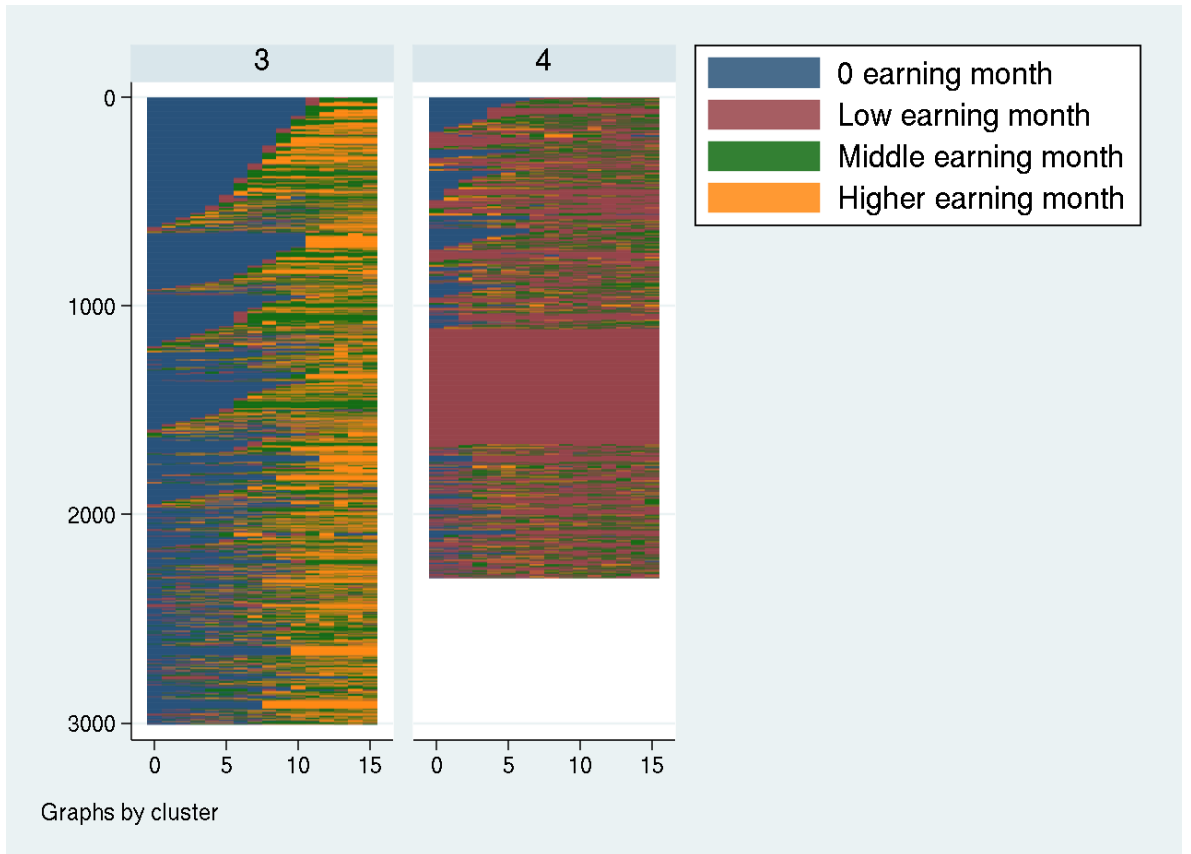


Figure 8.13 Super-cluster 2: stuttering transitions and low-pay

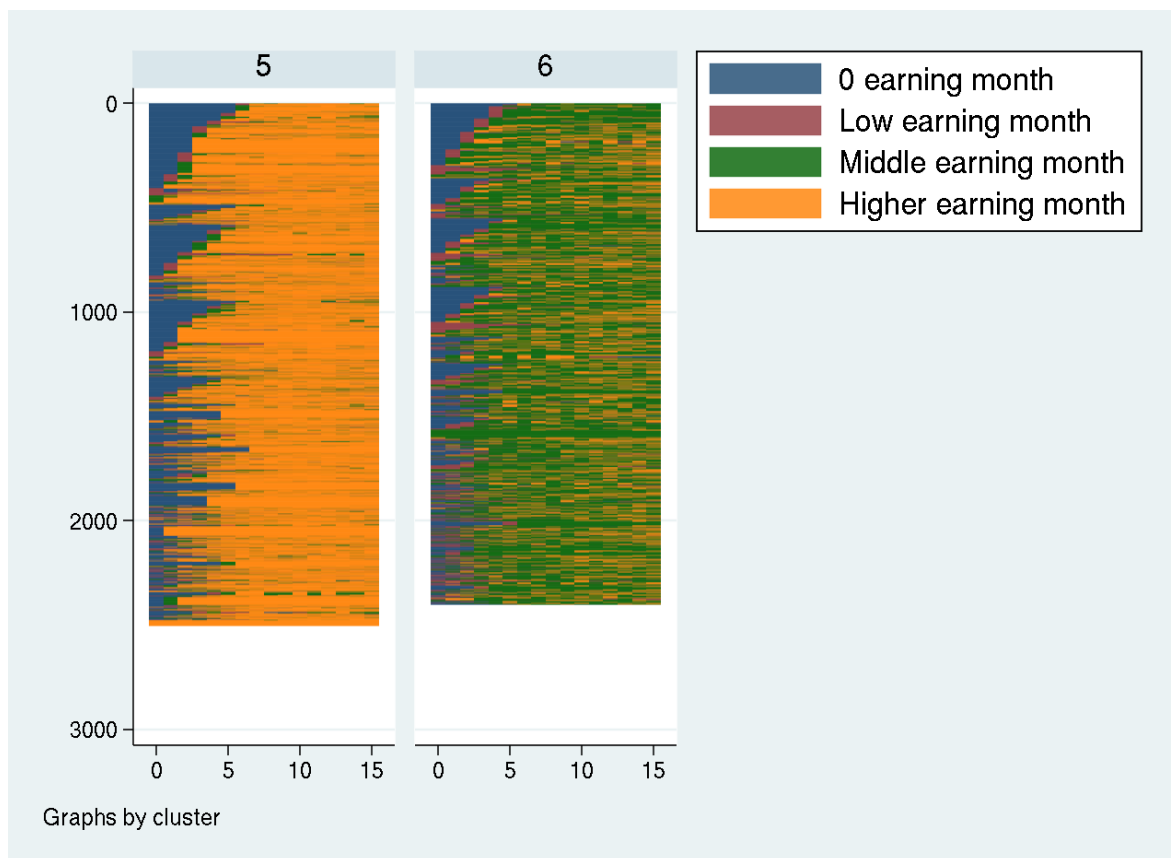


Figure 8.14 Super-cluster 3: quick entry to stable employment

Earning trajectory cluster	Work Programme	Work Choice	Proportion of WP RTI cases in each cluster	Proportion of WC RTI cases in each cluster
1	4004	3251	48.9%	32.8%
2	210	412	2.6%	4.2%
3	1351	1655	16.5%	16.7%
4	1084	1220	13.2%	12.3%
5	718	1784	8.8%	18.0%
6	816	1584	10.0%	16.0%
Total	8183	9906		

Figure 8.15 Work Programme and Work Choice participant presence within each of the earning trajectory clusters

Of the matched Work Programme cases which feature in the RTI dataset close to half (48.9%) are found in the ‘slow and chequered’ group. Work Choice participants in contrast are more commonly found in stable earnings trajectories: over a third of Work Choice participants in the RTI sit within the ‘quick entry to stable employment’ group. Work Choice participants therefore appear less likely to experience an employment trajectory punctuated by spells of very low or no employment earnings.

The analyses presented above suggest important differences are taking place for similar types of service users across these two differently configured programmes, with large impacts on key employment and earnings outcomes. Without a full cross-programme linked process and impact evaluation it is not possible to be explicit on the specific ways and reasons for such variation in user experiences and outcomes across

the schemes. Their different quasi-marketised configuration seems of central importance, yet it is not possible in these analyses to isolate and identify the precise programme characteristic or, more likely, constellation of programme characteristics that explain this large evidenced performance gap.

8.7 Discussion and conclusions

By utilising the thesis' original analytical framework presented in Chapter 2, this chapter operationalises the dimensions through which to understand variation in the way that quasi-markets may be constructed. Depending on the design of particular quasi-markets and their configurations with respect to these dimensions, providers and their service delivery activities will be differently responsive to the priorities and preferences of either the state, service users or providers themselves. This in turn is expected to have implications for the quantity, quality and distribution of services and consequently on any social outcomes achieved through such provision.

The analytical framework's ability to detect subtle variation in the formulation of quasi-markets is then operationalised within the empirical study of the Work Programme and Work Choice. Whilst on paper these programmes were commissioned according to the same commissioning strategy which sought to 'unleash' the innovation and entrepreneurialism of providers, the specific market formulation across the schemes varies in important ways. The positioning of the Work Programme against the dimensions described above suggest that it sits as a wholehearted exemplar of a provider-directed quasi-market configuration.

Work Choice, in comparison to the Work Programme's wholehearted positioning as a provider-directed market, offers a more balanced or *hybridised* market position. The Work Choice arrangement counterpoises provider dominance through an enhanced role for the state in stipulating activity and regulating service quality and by furnishing service users with an (albeit limited) degree of choice. Although it leans *towards* the provider-led configuration Work Choice retains important elements of both a state-directed and user-responsive arrangement.

The chapter draws on uniquely rich RTI earnings data and links this with an innovative methodological approach which connects the construction of a cross-programme quasi-experimental study of matched Work Programme and Work Choice participants with longitudinal sequence analysis to explore these earning pathways in unrivalled depth and detail. Critically, the enhanced outcome performance promised within the Work Programme's provider-led approach is *not* what we witness in practice. The empirical work shows that amongst matched programme participants with health conditions and disabilities Work Programme performance is significantly *worse* than that achieved through Work Choice in terms of both employment and earnings outcomes.

The balanced and hybridised approach to quasi-market configuration advanced under Work Choice appears to be important when reading the empirical findings offered here alongside the now sizeable body of evidence on the Work Programme from within DWP's official evaluation (Lane et al., 2013; Meager et al., 2014; Newton et al., 2012), from select committees (PAC, 2012, 2013, WPSC, 2013, 2011) and from academic research (Carter and Whitworth, 2015; Rees et al., 2014).

This suggests that providers have not responded to the Work Programme's provider-led arrangement with innovative service offers tailored to the personalised support needs of programme participants. Instead, providers appear to have utilised their freedoms to maximise profits and reduce risk by cynically deprioritising services for those who – in providers' eyes – may be more difficult to move into work. This tallies with the findings here which indicate that poor outcomes performance for those with health conditions and disabilities are *not inevitable* and that the employment and earnings outcomes for these groups are markedly better within the hybridised quasi-market configuration. Under the Work Programme the interests of providers are dominant and in practice this arrangement appears to function in such a way that disadvantages unemployed programme participants with disabilities and health conditions, whose employment and earnings outcomes are significantly and sizeably lower than what they would have been, had they participated in the alternative Work Choice scheme.

In sum, the variation in quasi-marketised form between the Work Programme and Work Choice emerges as an important mediator of programme participant experiences across the two schemes. Whilst the cause of this performance gap cannot be surmised from these analyses, the analytical framework, as well as official Work Choice evaluation material (Purvis et al., 2013)(Purvis et al., 2013), points a suggestive finger towards stronger state stipulation of relatively intensive pre-employment engagement and guaranteed in-work support for Work Choice participants allied with a notably less aggressive payment-by-results profile for this cohort.

9 Rest In Peace Work Programme: Epitaph for a bold quasi-market experiment

9.1 Beginnings of the eulogy...

In Summer 2018 the Work Programme is in its dying days: the final cohort of participants were referred in March 2017 and providers will continue to operate up until 2021 but with dwindling numbers of participants (Commercial Directorate Correspondence Team, 2016). Now is therefore an opportune moment to pause and take stock of the experiences and understandings that have emerged from this bold experiment in steering services through a private power market.

This concluding chapter is structured across four key sections, the first three of which outline the core contributions, study limitations, and key conclusions across conceptual, empirical and policy domains respectively. The final section considers the emergent research agenda as a result of these Work Programme experiences and findings.

9.2 Conceptual and analytic contributions

The original conceptual framework developed across Chapter 2 and applied in Chapters 3 and 8 offers a unique and subtle contribution to the academic understanding of ‘the difference thesis’ (Powell, 2015) by extending and unpacking the dimensions of variation within alternative configurations of public service quasi-markets. By extending from previous work investigating varieties of quasi-marketisation in public service delivery (Gingrich, 2011; Wiggan 2015a; 2015b), the conceptual framework developed and utilised in this thesis introduces to the scholarly literature a set of richer graduated dimensions through which to better understand variation in the way that quasi-markets may be constructed.

Within this frame, the allocation dimension captures the form of financing for services (ranging from wholesale public funding to fully private provision) and the regulation of provider activity by the state (stretching from very light-touch or negligible regulation through to comprehensive scrutiny and service quality control). The production dimension captures three underlying axes: how ‘open’ and competitive the market is, spanning from an open, highly competitive and accessible market to a situation where access is constrained or limited; the specification of provider activity, ranging from a high degree of public sector control over provider activity to an increasingly outcomes-led approach, where the public sector takes a position of agnosticism on the means of securing specified outcomes; and the role of service user choice and voice, spanning from a situation where the preferences and concerns of participating citizens act as a powerful steering mechanism for provider behaviour to a state where service user choice and/or voice is constrained to such a degree that it has no traction for service provider actions.

Depending on the configuration of particular quasi-markets with respect to these dimensions, providers and their service delivery activities are expected to be differently responsive to the priorities and preferences of either the state, service users or providers themselves. This in turn is expected to have implications for the quantity, quality and distribution of services and consequently on any outcomes achieved through such provision.

A key conceptual contribution of this thesis's original framework as a series of graduated axes is firstly to offer a richer articulation of important underlying dimensions against which quasi-marketised arrangements vary. A second core intellectual contribution stems from the re-conception (and visualisation) of these dimensions as graduated. Through these continua the framework is more alert to subtle variations in the extent as well as to the type of quasi-market under study. While quasi-markets *may* embody the 'pure' alternative types of either state-, user- or provider-directed forms, in practice *gradients* are likely to exist across these types. Quasi-market arrangements in practice are messy and the formulations on offer may straddle the clean-cut boundaries suggested in Gingrich's (2011) framework. This notion of softening, or *hybrid quasi-market forms*, has parallels with other work analysing employment support provision (Heins and Bennett, 2016, p. 40) which has identified classification complexity when it comes to understanding "the new welfare providers". Although Heins and Bennett's work centres on service providers, their conclusions – that a process of hybridisation and boundary crossing is increasingly important in order to understand and address concerns around organisational change and service orientation – have clear echoes here.

The conceptual framework's practical utility and ability to detect subtle variation in the formulation of quasi-markets is then utilised within the detailed appraisal of the Work Programme as well as in the comparative investigation of Work Programme and Work Choice. Compared to Gingrich's classification, the framework developed here offers a greater ability to trace the rich, multi-dimensional differentiation of welfare-to-work programmes and brings enhanced applied analytical power to comparative scholarly investigation. Though applied here specifically to the field of British welfare-to-work arrangements, the device has been conceived as a sufficiently flexible tool so as to have wider applicability to alternate quasi-market contexts both beyond the UK and beyond the field of welfare-to-work.

The framework reveals itself to be valuable as an identification tool. Reading the Work Programme against the dimensions of this analytical framework signals the scheme to be a wholehearted exemplar of a provider-directed market. Considering the production dimension (Figure 3.6), in terms of market access the Work Programme configuration is relatively closed, prioritises incumbents and is only very weakly competitive. In relation to payment specifications Work Programme fully embraces payment-by-results and 'black box' commissioning such that the state has minimal control in specifying services. Finally, there is limited scope for user choice and voice to inform the production of services. This orientation reflects a strong provider position: through

minimal competition (upper axis); minimal state control (middle axis); or minimal responsiveness to user preferences (lower axis). The consequence is that *provider* preferences are the dominant set of interests shaping the functioning of this market. Implicitly, service-user desires for quality and state ambitions for efficiency are subsumed beneath a provider logic that seeks to maximise profit from job outcomes within a system that gives a high degree of discretion in pursuing this aim.

Importantly, tighter than expected cost-pressures have been exerted on Work Programme Primes from the outset due to a combination of a more difficult than expected macroeconomic environment which affected job outcome (and hence payment) levels, the strong prevalence – and apparent success of – discounting practices at the bidding stage (NAO, 2012), and some have suggested unrealistic performance expectations were set by the DWP for some groups which would have the effect of depressing profit margins against projected levels (Mulheirn, 2011). Private providers are strongly motivated to retain profit margins and this will become particularly apparent under such extreme cost pressures. In the Work Programme it is expected that providers are more able to cascade the ‘cost’ of poor or partial provision to service users due to weak regulation of service quality on the allocation dimension. Therefore, the conceptual framework devised in this thesis more specifically enables the identification of the Work Programme as a form of *private power market*.

The analytic framework developed in the thesis therefore enables greater conceptual depth and precision to be placed on the Work Programme’s particular quasi-market construction. Work Programme is not just a quasi-market, it is a particular type of quasi-market – a private power market. Compared to existing conceptual frameworks, this fuller and more precise specification more powerfully identifies and articulates the particular and acute constellation of risks surrounding the programme. In any marketised arrangement there will be risks that private providers pursue cost-cutting innovations to improve crude efficiency at the expense of unspecified or poorly captured aspects of services – such as quality (Williamson, 1975). Such risks however will be amplified within a private power market such as the Work Programme where there are intentionally under-specified contracts (outcomes only) and producers have great control. This ‘look Mum, no hands’ form of market configuration leaves the Work Programme promising innovation but with acute hazards around rent-seeking and uncontrolled cost-cutting by providers at the expense of efficient or high-quality production.

The final benefit of this thesis’s original conceptual device is its explicit acknowledgement of hybridity and the important, but often neglected, ways in which quasi-market alternatives may flow from the overlaying and intersection of additional non-market (or further marketised) accountability tools. The dimensions of variation provide hints to academics and policy makers as to the accountability levers that might be used as correctives to recalibrate quasi-market systems in situations where perverse behaviours, as a result of a single stakeholder’s dominance, are leading to unintended

and undesirable outcomes. This conceptual contribution is discussed further below in relation to its ensuing policy contributions.

9.2.1 Critical reflections and limitations of the conceptual framework

The conceptual framework developed within the thesis provides an important contribution via its rich articulation of quasi-market varieties and their underlying graduated dimensions (and hence hybridised forms). Nevertheless, there are limitations to the conceptual framework and the degree to which the current empirical work is able to build from and directly respond to its inferences.

Via the conceptual framework the thesis implicitly generates something of a hypothesis: *quasi-marketised public services which are organised at different points against the underlying axes of allocation and production dimensions will be differently responsive to the respective priorities and preferences of either the state, service users or providers*. In turn it is anticipated that such configurations will have important (and detectable) implications for the quantity, quality and distribution of services and consequently on any outcomes achieved through such provision. The challenge of testing such expectations sits on three fronts. Firstly, there are limitations in the availability and clarity of metrics in support of the underlying dimensions. Secondly, the framework may underplay the salience of other factors associated with the quality of public service provision, yet which do not feature within, or directly relate to, the particular formulation of the quasi-market. Finally, and related to the preceding points, the limitations associated with available DWP administrative data constrains the degree to which conclusions about the causal relationship between altered quasi-market formulation and altered achievement of social outcomes can be drawn. Reflections are offered on each of these issues in turn.

9.2.1.1 Challenges in operationalising and scaling the underlying dimensions of allocation and production

In operationalising the allocation and production dimensions within the thesis, the programme markers associated with alternative quasi-marketised formulations are positioned notionally. The positioning of the Work Programme against each of the dimensions was not a clear-cut or straightforward consideration. The placement decision was made by considering a hypothetical case at the poles of each axis and then making an assessment as to the relative closeness of the Work Programme formulation vis-à-vis these extremes. The marker was then positioned figuratively with polarity and distance inferred through detailed programmatic knowledge. For some axes this was considerably more straightforward and unambiguous than others.

Within the allocation dimension the upper ‘financing’ component (Figure 3.5) directly positions the Work Programme at the extreme leftmost position of this axis as the scheme is fully congruent with a collectively financed arrangement. There is no scope for the Work Programme to draw in the private financial resources of individuals (through co-payments, for example).

Even in potential situations of hybridity on the financing axis there are reasonably straightforward and transparent routes available to the positioning of hypothetical programmes. For example, where direct funding for the provision of services through taxation has the potential to be supplemented via direct service user charging or insurance contributions, the positioning could be tied with reasonable accuracy to the relative contributions from each source. If overall financing from general taxation were equivalent to 75 percent of total programmatic spending with (in aggregate) a 25 percent contribution from service users then a programme marker could be placed three-quarters of the way along the axis, towards the right-hand side. This is illustrated by the green marker with an asterisk in the upper portion of **Error! Reference source not found..**

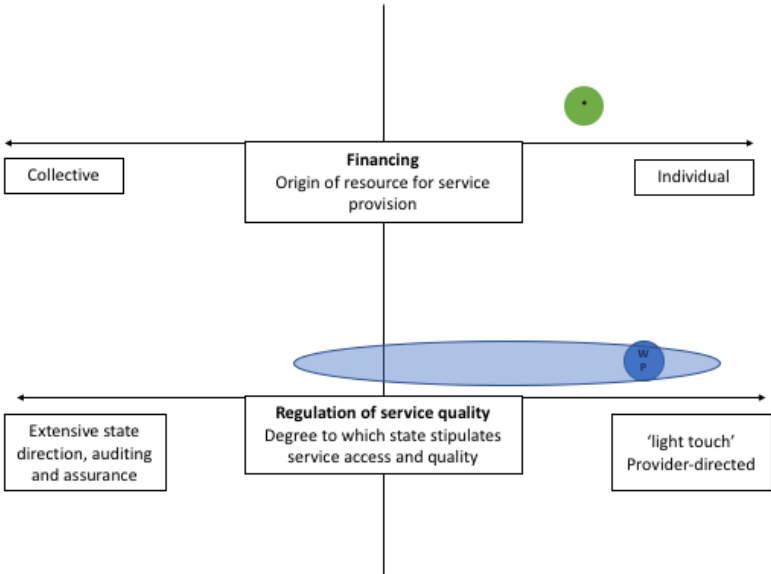


Figure 9.1 Notional positioning of programme markers against allocation axes

Contrastingly, the ‘regulation of service quality’ and the degree to which this is upheld via regulatory functions performed by the public sector is considerably more ambiguous. Grout and Stevens (2003) suggest that whether it is theoretically preferable for the public sector or private sector to own the assets underlying service provision and/or produce a service depends on the ability to contract and the relationship between cost reduction and quality. Complicating this, the welfare consequences of reductions in service quality and the challenge (and hence cost) of stipulating acceptable service quality in a contract varies considerably across different public service domains (Grout and Stevens, 2003). The question as to what form of contract and regulatory regime gives service providers the optimal incentives with respect to cost and quality remains. There is little by way of formalised tools or frameworks for tracing the relative intensity and efficacy of regulatory regimes applied to outsourced public service provision.

The positioning of the Work Programme against this axis within the thesis was a purely notional exercise. The marker is situated towards the right hand ‘light-touch’ provider-directed pole though in actuality there is a wide sphere within which the marker could

reasonably be placed. The blue shaded oval area in the lower portion of **Error! Reference source not found.** represents this high scope for ambiguity by capturing the range of potentially viable positions for the programme key. The programme position identified in Chapter 3 – though not explicit – was informed in large part by the narrative of the Freud report and the consequent commissioning strategy (DWP, 2008) which captured a clear ambition for the DWP to reduce its contract monitoring burden. Instead the responsibility for contract management and supply chain management in the Work Programme is subsumed by Prime providers. This commitment to a ‘light touch’ regime by the department is coupled with the wider perception that the DWP does not – and indeed because of the variability of minimum service guarantees cannot – keep a tight rein on upholding service quality and access (Finn, 2012 and NAO, 2012).

The challenge of developing standardised approaches through which to consider the extent and intensity of regulation applied by the state to upholding service quality is to be pursued through future research subsequent to the thesis. As a starting point, within a single policy domain or stream of public service provision one – reasonable but crude – proxy measure might be captured by considering the volume of resource (in financial terms and/or in terms of staff hours) dedicated by the public sector to auditing and contract management work, particularly where this is focused on upholding access and minimum service expectations. In some cases, such scrutiny functions may be undertaken by an ‘independent’ arms-length body, such as the Care Quality Commission in the situation of English health and social care regulation. The degree of quality assurance activities may then be conceived as a ratio, with the numerator capturing the amount spent on this quality auditing and assurance function and the denominator containing the total financial value of outsourced provision.

A ratio capturing the relative effort applied to quality assurance vis-à-vis contractual service provision may provide something of a baseline or starting point for the understanding of regulation in different forms of quasi-marketised arrangement. The ratio gives potential anchor points at the poles of the ‘regulation of service quality’ spectrum. An extreme position, the full embodiment of ‘extensive state direction, auditing and assurance’ might equate to a value close to 1, that is, the scrutiny effort has the same financial value as the outsourced contract provision. At the other extreme, nil effort to regulate service access and quality on the part of the public sector would take a value of 0. Tracking this ratio over time may give a sense of whether quality scrutiny is being escalated or scaled back. Importantly, such conclusions are only likely to be viable in situations where contracting structure, scale of contracts and the nature of outsourced provision are broadly comparable over time. Without this, comparability may be severely compromised since the ability to regulate quality varies considerably. Hipp and Warner also suggest that service monitoring systems for outsourced provision may oscillate between ‘relational contracting’ with close collaboration based on trust between independent providers and public sector purchasers, and more formal criteria which “increase the distance between private

providers and government overseers” (2008, p. 88) with potential implications for quality and access in the long-term.

Similar challenges apply to the axes underlying the production dimension. There are bodies of theoretical work which may helpfully underpin future operationalisation here. For example, on the axis of market openness and competitiveness Baumol’s (1982, p. 2) work moves beyond a rough continuum “with relative efficiency in resource allocation increasing monotonically as the number of firms expands” (i.e. an assumption that a plentiful array of provider organisations is a proxy for competition) to provide a guide to the relevant characteristics of contestability within monopolies and oligopolistic arrangements. Contestability may be signalled by the absence of cost discrimination against entrants, absence of profits beyond a normal rate, and the perceived vulnerability to hit-and-run entry (Baumol, 1982).

Other empirical studies may provide useful metrics through which to investigate the axis of user choice and voice. Here, importantly, Hipp and Warner (2008) make a distinction between the presence of service user choice (for example, within voucher-based training provision) and challenges to such choice having traction for provider behaviours (due to information asymmetries and preference misalignment). There is potentially a disconnect between the *designed* position of programmes against particular axes and the functioning of such a quasi-marketised formulation. Voucher systems are not tautologous to a user-oriented quasi-market arrangement. The suggestion from Hipp and Warner (2008) is that effective user choice may need to be buttressed by a strong performance monitoring and information-sharing function on the part of the public sector. We then rub against a further future conceptual and empirical challenge: that the axes may not be appropriately treated as independent entities or fields for analysis and description. There may be co-dependencies or interactions across the alternative quasi-market axes such that user voice can only be ‘strong’ in situations where access and quality regulation also take a particular form. There is then considerable future work required in order to develop the measurement and appropriate benchmarks underpinning the axes of allocation and production such that they can be operationalised in testable hypotheses.

9.2.1.2 Salience of factors beyond the quasi-market associated with the quality of public service provision

In the comparison of two quasi-market varieties offered in Chapter 8 the discussion centres on the variation in the characteristics of each scheme (Work Programme and Work Choice) as directly related to quasi-market forms. In an ideal comparative study, the impact evaluation would be constructed around two programmatic forms in which the only difference between the schemes is in their quasi-market formulation. All other programmatic characteristics which might moderate the quantity, quality and distribution of social outcomes achieved would then be held constant. In practice such a ‘pure’ comparison is not viable and there are salient points of deviation which are not fully accounted for in each programme’s relative forms of quasi-market. In

particular, and as acknowledged in Chapter 8, the generosity of the funding available per participant is not directly comparable. Work Choice is understood to be a more generously funded scheme. Because of the lack of financial information available and the uncertainty relating to cross-subsidy between payment groups within the Work Programme it was not possible to control for variation in spending generosity in the analysis. The legal form of dominant provider organisations within each programme is also understood to vary. Voluntary sector providers are understood to play a stronger role in Work Choice and it may be expected that social mission may have implications for provision which extends beyond the quasi-market form. A lack of transparency on the composition of supply chains within the Work Programme restricts any ability to control for the legal form and/or social motivation of service providers directly engaging with participants under the scheme.

There is also variation in the substance of programmatic content across Work Programme and Work Choice. The analysis does, in effect, conflate variation in programmatic content with variation in quasi-market form. However, in the Work Programme, due to the extreme form of payment-by-results and black box commissioning, there are no programmatic stipulations and hence quasi-market formulation and programme are one and the same. Thus, in this situation the blurring of programme implications and quasi-market implications for the patterning of employment outcomes is inevitable. Nevertheless, care must be taken when drawing causal conclusions in the extent to which variation in the outcomes achieved under each of the two schemes are solely attributable to the configuration of their respective quasi-markets.

9.2.1.3 Constrained conclusions due to availability of data and programme implementation

The strength of the conclusions that can be drawn in the thesis are also constrained by the structure of programme implementation and data accessible to the study. A key limitation both to this study and to the wider evaluation of the Work Programme is that the scheme was implemented in a way that makes impact evaluation approaches particularly challenging. There was no pilot programme study, no clearly defined comparator group and everyone eligible for support was immediately referred to the programme. Although the Work Programme replaced a range of preceding welfare-to-work schemes the absence of a temporally or geographically phased implementation approach has made it particularly challenging to construct an appropriate and robust quantitative impact evaluation. Because of this, combined with the challenge of isolating changes in the configuration of quasi-market arrangements from other salient aspects of programming and support (identified above) the degree to which causal claims can be tied to the quasi-market configuration must necessarily be subject to heavy caveats.

The administrative data available to the thesis relate only to the Work Programme and to Work Choice. Individual-level information for those participating in forerunner

welfare-to-work schemes was not included as part of the Memorandum of Understanding between the DWP and the University. A wider appraisal across a range of alternatively configured DWP programmes was therefore not possible. This means that a systematic consideration across a broader range of alternatively configured quasi-market arrangements has yet to be pursued. Moreover, the variation in the two quasi-marketised programme formulations under investigation took the form of a ‘natural’ experiment since the variation was not induced formally by the research team. These programmes vary on more than one quasi-market dimension and this means that the variation in outcomes cannot straightforwardly and unambiguously be attributed to the variation in a single independent variable. The ultimate findings must be couched accordingly, yet there are still pertinent and unique empirical contributions arising from the thesis.

9.3 Empirical contributions

Under the marketised governance of public services it is expected that that opportunism, or ‘gaming’, will be a particular challenge. The international literature also makes clear that the specificities of programme design and payment structures can play a key role in either facilitating or buttressing against these undesirable provider behaviours. The implication is that through careful contractual specification and regulation it may be possible to overcome, or at least mitigate, these behaviours and to ensure that providers act in support of the full range of programmatic objectives. The challenge is for policymakers – through design savvy – to configure and implement a suite of steering tools that retain the perceived innovative and efficient impetus of provider-directed quasi-markets whilst ensuring that a complex set of public sector objectives are met.

Crucially, detailed empirical work investigating these alternative accountability design tools is rare internationally and the detailed systematic quantitative analyses of the sort presented here is particularly scarce. In the case of the UK Work Programme, the detail and extent of empirical understanding of the impacts of the scheme and its design configurations have been severely hampered by the DWP’s failure to commission any form of quantitative impact evaluation, but even then impact evaluations of this sort do not link to targeted conceptual explorations of governance configurations and mechanisms as in this thesis. In using DWP’s administrative data the thesis is therefore of considerable value in extending our scholarly understanding of the implications of this particular and acute quasi-market design configuration.

The Work Programme’s designers responded to the anticipated quasi-market risks of opportunism and gaming with, in many ways, a set of novel and untested design features in the pursuit of the government’s objectives for the programme: to support *more* participants into work, *faster*, *for longer* and whilst *reducing gaps in performance outcomes between the easier- and harder-to-help*. There are three particularly bold marketised accountability design elements at play within Work Programme: a differential payment structure in a bid to override tendencies for creaming and parking of participants; large regional contracting areas at which

minimum performance is assessed and market-share is shifted in the attempt to stimulate competition and higher aggregate performance; and extensive sustainment payments are intended to ensure longer-term sustainable employment outcomes for participants.

The empirical work in this thesis assesses whether the Work Programme's crafted design structure is sufficient and appropriate in steering the private power market to the achievement of the full gamut of government policy objectives. The key empirical contributions stem from this three-fold combination of conceptually guided empirical analyses, the project's unique academic access to participant-level administrative data held by the DWP, and the research's bespoke and sophisticated range of quantitative analyses tailored specifically to investigate those design tools and allied promises made within Work Programme's private power market.

Via the complex cleaning and linkage of different DWP administrative datasets it has been possible – at the level of the individual – to connect the outcomes data with a considerably larger and broader array of 'employability' related explanatory variables than ever previously studied. The result is that this thesis, unlike any previous independent academic analysis of the Work Programme, uses the richest possible set of explanatory variables and official programme outcome information sourced from the administrative data.

Moreover, the analysis pushes far beyond the insights which have been offered in the official commissioned evaluation work as it centres on these particular design 'innovations' which mark a strident departure from previous UK and international approaches to steering incentives in quasi-marketised systems. Each of the empirical chapters makes a distinct and empirically original contribution to the academic understanding of these design tools, which have not been implemented in such extreme forms before and hence which are appraised for the first time here.

Chapter 5 investigates whether the differential payment structure has calibrated provider incentives within the private power market such that all programme participants have the same likelihood of entering and sustaining paid work, regardless of their characteristics and circumstances. The chapter offers an original and rich programme-wide quantitative assessment of the extent of claimant variation within payment groups and, as a result, the extent to which the Work Programme's current payment-group based differential payment system may be designing *in* rather than, as intended, designing *out* provider incentives to cream and park claimants. Irrespective of its initial apparent subtlety compared to previous employment programmes, this chapter's detailed original analyses reveal that the Work Programme's differentiated payment model has failed woefully to deliver its objectives around differentiated universalism. Instead of neutralising the effect of well-known indicators of labour market disadvantage, in the aftermath of the Work Programme's payment-by-results system there are persistent, sizeable and significant effects remaining across a range of participant characteristics and contexts.

Chapter 6 investigates the frequently neglected geography of incentives and performance in the Work Programme, assessing potential tensions between the spatially extensive regional Contract Package Areas and the far smaller local labour market contexts within which providers and service users exist and act. In doing so the chapter offers the first ever systematic consideration of spatial variation in Work Programme performance, despite the key spatial dimensions and spatial risks of welfare-to-work programmes such as the Work Programme.

Specifically, Chapter 6 explores whether the geographical container of regional Contract Package Areas is an appropriate spatial scale at which to build key elements of welfare-to-work programmes including upholding minimum performance levels and incentivising competition between providers. The flexibilities of the private power market, in conjunction with the general absence of contextual considerations in the design of key programme steering mechanisms, has left the Work Programme susceptible to spatial variation in performance. The original analytical work presented in Chapter 6 reveals highly variegated programme job outcome performance across Britain. As with previous welfare-to-work programmes, the Work Programme does appear to be 'jeopardised by the geography of unemployment' since more deprived local authority areas, with proportionally larger long-term unemployed populations, are systematically experiencing significantly *worse* performance than more economically buoyant areas. The programme's structure of competition and contractual minima appear to have done nothing to mitigate these spatial inequalities, leaving participants who are living in deprived areas particularly vulnerable to experiencing poor job outcomes. The failure to calibrate for locally varied labour market contexts works against the programmatic ambition to reduce *gaps in performance outcomes between the easier- and harder-to-help* neighbourhoods and areas. As a consequence, the findings presented in Chapter 6 suggest that the undifferentiated geographical container of regional Contract Package Areas does not appear to be functioning as an appropriate geographical unit to uphold minimum performance standards or incentivise providers for the pursuit of programmatic objectives. More broadly, the spatial perspective and findings outlined in Chapter 6 highlight the need for welfare-to-work design and analysis to pay greater attention to the key but frequently overlooked geographical dimensions of policy, practice and performance.

Chapter 7 analyses the Work Programme's extensive use of sustainment payments and the structure of financial incentives which extend well beyond an initial period of employment. The author's unique academic access to a set of Real Time Information on the month-to-month earnings of a sample of Work Programme participants facilitates the first of its kind analysis of earning trajectories using sequence analysis. This approach identifies clusters of common earning trajectories experienced by programme participants and offers a considerably more subtle longitudinal analysis of employment and earning trajectories than is available using conventional statistical techniques and currently employed in welfare-to-work analyses. It is the first of its

kind in relation to Work Programme outcomes in the UK context. The majority of Work Programme participants do not feature in the employee earning records across the tracking window as most do not experience any time in paid work across the programme (Meager et al., 2014). This is of clear policy concern, but already known. Amongst those who do experience time in paid employment, the analysis identifies for the first time in the literature four distinct participant clusters of longitudinal earning trajectories: ‘slow and unstable’; ‘stuck in low-pay’; ‘general earnings success’ and ‘gold standard’.

Concerningly, those earnings trajectory types which would be deemed ‘successful’ by policy makers only account for 13.7 percent of Work Programme participant experiences. The analyses presented in Chapter 7 reveal a disconnect between the intended incentive structure and the experience of participants, close to one-third of whom are likely to be experiencing low-paid and unstable employment trajectories in circumstances where they do ‘succeed’ in moving into paid work. The Work Programme’s complex sustainment payment structure does not seem to be incentivising providers to develop provision and supports that enable entry to employment to be successfully converted to sustainable, well-paid employment trajectories. For many Work Programme participants, even for ‘success stories’ the low-pay no-pay cycle none-too-successfully persists.

Chapter 8 offers a more synoptic reflection on the Work Programme’s private power market. It returns to the analytic framework developed in Chapter 2 and uses this to compare the configuration of two alternative quasi-marketised employment support programmes which have run in tandem in the UK context since 2011 with overlap in their cohorts: Work Programme and Work Choice. The analytic framework enables important differences in the type of quasi-market adopted by the two schemes to be teased out – Work Programme as a private power market compared to Work Choice as a softer hybrid provider-directed market configuration. A novel quasi-experimental analysis is then used to investigate the implications of the alternate market formulations for those with health conditions and disabilities by comparing the employment and earning outcomes for a matched group of participants on the two programmes. This is the first known study to use a quantitative impact evaluation method to consider the potential implications of alternative quasi-market formulations. It is, inevitably therefore, the first of its kind in relation to the Work Programme (and Work Choice). This is of particular value given both the persistent critique of poor programme performance for people with health conditions and disabilities combined with the UK government’s commitment to halve the disability employment gap (Conservative Party, 2017).

The findings suggest that the promises of innovation and performance improvement allied to the provider-directed Work Programme are not met. Employment and earnings outcomes are significantly and sizeably lower for the Work Programme than for Work Choice. The hybrid market position of Work Choice – which leans *towards* a provider-directed arrangement but retains important levers for both the state- and

user- preferences – emerges as an important mediator of programme participant experiences. The performance gap between the two is striking. Discovering its drivers more precisely for this key cohort emerges as a priority for research and policy.

9.4 Policy contributions

As discussed in Chapter 4, a key ambition for the research is to connect the original empirical insights with lessons which will be of direct relevance to policymakers grappling with the design of accountability tools within quasi-marketised public service arrangements both in the UK and beyond. These ‘incremental’ adjustments to current design levers are discussed first before considering the more seismic devolutionary shifts that are developing across the British employment support landscape.

The key policy lesson emerging from the first empirical chapter is that the Work Programme’s current differential payment system seems more likely to design *in* rather than design *out* ‘creaming’ and ‘parking’, given that the payment groups which underpin the differential pricing system are so crude and the variation in participant characteristics *within* these payment groups is so sizeable.

An instinctive corrective here is to calibrate payment levels more closely with modelled likelihoods of moving into sustained employment (Carter and Whitworth, 2015). In response, a statistical approach to differential payment setting is explored using the DWP’s own administrative data. This approach conditions the value of job outcome payments on the individualised probabilities of achieving sustained employment outcomes. The statistically derived programme streams are shown to be a viable and more effective way to design a set of alternative and empirically grounded payment groups, offering greater predictive power and value-for-money than is the case in the current Work Programme design. Advancing UK policy design around richer statistically informed profiling and payment design remains a priority in the coming decade. Further data than is currently available within DWP’s administrative records will be a key part of this. Yet even within the constraints of current administrative data the enhanced analysis, considerations, and comparative assessment set out in this chapter indicate that it is viable for UK policy makers to do better and offers a proposed way forwards.

Chapter 6 raises concerns that the current large scale Contract Package Areas (CPAs) are an inappropriate tool for comparing performance across multiple Prime providers (beyond a single CPA) and for upholding contractual minimum performance levels. Two main policy implications arise from its analyses.

Firstly, the extent of spatial variation discovered within CPAs, and in particular the extent of poor performance within those smaller local authority geographies, highlights Work Programme’s blind spots towards the spatial dimensions of high risk and weak realisation of programmatic objectives. Importantly, the empirical analyses presented in Chapter 6 highlight that the unit of performance assessment and

incentives will need to be substantially reduced in geographic extent in order to respond to these spatial concerns. That the UK's new Work and Health Programme, the far smaller successor scheme to the Work Programme, shifts to even larger contractual geographies than operated in the Work Programme, and fails to offer any spatially sensitive accountability levers, is of significant concern.

Secondly, since the current contractual minimum performance levels are standardised across Britain, and this unadjusted national standard is applied indiscriminately, there is a risk that Primes are judged as 'underperforming' when their lower aggregate performance at the level of CPA may be informed by lower prevailing levels of demand across a particular region. An intervening improvement would be to consider the development of contextually sensitive programme performance data. As a first step, this could be applied at the current level of CPAs (as is offered through the proposed 'contextual value added' performance metric at the close of the chapter). More substantively, however, if policy makers seek to enhance the performance of the programme in depressed local labour market areas in order to 'close the gap' then – continuing within the logic of a private power market – providers will need to be financially incentivised (or penalised for failure) to work against the grain of pre-existing spatial inequalities at the level of local labour markets.

The key policy conclusion emerging from Chapter 7 is that the Work Programme's current emphasis on sustainment payments within the structure of financial incentives has not broken the low-pay no-pay cycle for many programme participants. The overwhelming experience of programme participants is one of not entering employment at all, a clear remaining area of policy concern. Yet of those who do feature in the HMRC earning data (i.e. people who have at least one month of employment-related earnings within the tracking window), also of concern is that close to one-third experience an earning trajectory which is categorised as slow and unstable or where participants only receive very low pay across the period (equivalent to earning less than 16-hours per week at NMW). For a programme so heavily emphasising sustainment in its rhetoric and payments profile this is of significant concern.

At this stage there is no immediate or intuitive 'fix' to this payment-based accountability mechanism as conclusions cannot be drawn from the current analysis as to whether these patterns are associated with a failure in the transmission of incentives such that providers have been unable to *translate* the complex set of payment values to meaningfully alter the shape of provision at the frontline, *or* that despite the powerful incentives for sustained employment there are insufficient available in-work support services with a track record of success. Alongside welfare-to-work programmes, improving in-work support is emerging as a key policy weakness within Universal Credit's drive for in-work earnings progression and is likely to be core area of UK policy and research focus in the coming years.

9.4.1 Stepping out of the private power market?

Taking a wider view, the policy recommendations flowing from the first two empirical chapters in particular very much keep *within* the logic of the private power market. The recommendations do not seek to disturb this particular quasi-market formulation but rather to refine and enhance key marketised levers of price (statistically informed price setting) and performance monitoring for market share manipulation purposes (contextually sensitive assessments of programme performance in order to better reward/penalise providers who are over/under performing).

The disappointing performance of the Work Programme for those with health conditions and disabilities, and the varying performance of Work Programme and Work Choice outlined for the first time in the analysis of Chapter 8, may have more fundamental ramifications for current and future policy design.

The recommendations that emerge from the final empirical chapter specifically suggest that for the priority cohort of participants with health conditions and disabilities at least the private power market configuration as conceived in the Work Programme has not delivered the best possible service and outcomes for these participants. What emerges here is the potential value in softening the extreme form of provider-directed quasi-market and interlacing and overlaying alternative non-market accountability types in the pursuit of a softened, hybridised form of quasi-market arrangement in order to better attune provider behaviours to good quality and sufficiently intensive provision. Each of the empirical chapters, but perhaps Chapter 8 most unequivocally, indicate that future UK programmes would do well to bolster the traction of state- and user- priorities and preferences through procedural and/or democratic forms of accountability.

Taken together the findings suggest the need for policymakers to give serious consideration to moving *beyond* the refinement of those marketised accountability levers which have already been heavily (and disappointingly) relied upon within the Work Programme and instead to look to new roles and configurations of wider accountability mechanism beyond such marketised levers.

9.5 Burned, bruised or broken? Future policy and research directions

As noted at the start of this chapter Work Programme is now in its dying days. In the aftermath of the Work Programme the follow-on scheme – the Work and Health Programme – is considerably smaller and is more tightly focused on people who have health conditions or disabilities and, in minor part, on those who are very long-term unemployed. The programme replaces both the Work Programme and Work Choice which had a combined budget in of £540.8 million in 2015/16 (£416.4 million Work Programme, £124.4 million Work Choice); the Work and Health Programme will by contrast have a budget in the order of £130 million per year in 2019/20 (Powell, 2018), around 20% of the financial envelope without adjusting downwards for inflationary reductions in its real purchasing power.

Whilst this might superficially be read as a sign that the DWP has been ‘burned’ by the Work Programme experience and is recoiling from its bold private power market experiment, hearsay amongst mandarins would suggest that the scaling back of British quasi-marketised employment support has instead been driven principally by the challenging spending envelope set for DWP by the Treasury rather than any great programmatic learning on behalf of DWP. Considered by some to be of relevance was the lack of a robust quantitative impact estimate for the Work Programme which made it understandably challenging for DWP to demonstrate value for money of the Work Programme to HM Treasury’s satisfaction.

However, although scaled down the Work and Health Programme softens somewhat, but does not fundamentally scale back from, the preceding quasi-market logic and characteristics of the Work Programme’s private power market. The black box, lightly regulated, provider-directed market logic prevails. Andrew Thomas, Director of Contracted Services at the DWP, explains that within the black box programme specification *providers* will set the “Customer Service Standards” (Thomas, 2017, p. 6). Although the weighting of payment to outcomes takes a slightly softer formulation – 30 percent of payment will be made through guaranteed delivery fees with 70 percent on results – the financial incentives are overwhelmingly ‘job’ focused. There is no specified modular service component (which was seen as the justification for guaranteed payments within Work Choice). Service payments are not directly tied to any specific minimum service assurance. And contract areas become not smaller but instead far larger than in Work Programme, thought by many to reflect the far smaller financial envelope available and the desire by DWP to remain (at least partially) attractive to the market.

Indeed, a new and additional quasi-marketised accountability tool is set to be implemented through Work and Health Programme: an accelerator payment model, whereby the value of outcome payments made to providers increases in step with the proportion of participants who achieve successful outcomes. Though this approach has tentatively been welcomed as a potential antidote to issues of creaming and parking, concerns have already been raised that such a pricing arrangement is notoriously difficult to model and therefore may not overcome risks of gaming and neglect (Wilson, 2017).

While the DWP trudges on with the mantle of experimentation in its familiar mould of centrally contracted Prime providers in spatially vast private power markets (indeed the tongue in cheek subtitle for an event outlining the new Work and Health Programme was subtitled ‘the Sequel’), there has nevertheless been a ‘cooling’ attitude to large-scale payment by results schemes amongst Whitehall civil servants more broadly (Brown, 2013). Central government officials have acknowledged that their attitude was “bullish on payment-by-results at the start” (Mair, 2013, paragraph 5), where by ‘payment-by-results’ civil servants are typically referring to the heavy use of outcomes-based payment mechanisms within a lightly regulated provider-led quasi-market. There is a growing acknowledgement within Central government that there is

a need for a more measured and pragmatic approach. There is therefore a question as to whether DWP's approach to quasi-market formulations as seen within Work Programme and now Work and Health Programme is 'bruised' – that is, a promising approach that has been oversold and plagued by poor design, but potentially fixable (and in that case preferable) with the right design – or whether it is instead 'broken' and we should retreat from ever further adoption of marketised modes of coordination and look elsewhere for solutions to these internationally shared and on-going employment challenges.

While central government considers these questions, it is at the local level in the UK that striking new experiments are being made in the pursuit of revised and improved accountability regimes surrounding employment support provision. As described by this author (Whitworth and Carter, 2018), “amidst persistently disappointing provision and outcomes performance for ‘harder-to-help’ claimants ... within large national quasi-marketised contracted-out provision, and an increasing recognition that the status quo cannot continue amidst the ever-tightening fiscal squeeze of austerity, employment support's devolutionary moment has come”. Within the historically highly centralised policy domain in the British context, city-regions – combinations of contiguous local authority government areas – are emerging as key new scales of governance in the English context.

The emergence of this new commissioning scale for welfare-to-work provision is not simply a straightforward shrinking of the DWP's private power markets however. Rather, leading city-regions articulate a progressive vision of locally integrated employment support 'ecosystems' that strategically and operationally coordinate a range of currently siloed and disconnected key wraparound support services and budgets that need to be brought together and marshalled in order to enable the delivery of whole-person holistic keyworker support packages (Whitworth and Carter, 2018). Integration is seen to be key in order to respond to individuals' support needs and this requires connections to be made across personalised employment advice, skills, health, housing, financial advice, and family support.

Whilst the DWP approach is focused squarely on an off-the-shelf quasi-marketised approach that spurs competition between providers based on a classic principal-agent relationship with a single Prime provider, city-regions in contrast appear to be pursuing an approach that rather than rethinking or bolstering wider accountability tools to stave off the worst of the negative outcomes arising from provider cost-minimisation tactics within this market logic are instead turning to a qualitatively different logic and model that seeks at heart to guide local interactions through a relational dynamic. Though often still based with contracted provider models these projects seek more strongly both to avoid the predatory competitive environment of marketised relations and also to pursue the cultivation of functioning horizontal network accountability in which to situate providers (and Jobcentre Plus).

City-region frontrunners – Greater Manchester, London, Sheffield City Region and West Midlands – have each pioneered alternative locally-developed integration mechanisms through which to bring together stakeholders from relevant multi-agency support systems. Greater Manchester’s ‘Working Well’ local integration model is out front in the gradual development of its local employment ecosystem, emphasising and consistently growing relational connectivity locally (Whitworth and Carter, 2018). Within Greater Manchester, for example, large-scale devolution employment pilot programmes, the budget for Work and Health Programme, and to a notable extent mainstream Jobcentre Plus activities exist within Greater Manchester’s expanding relational network approach, with Local Integration Boards (operational co-case management of wraparound services) and ecosystem development and refinement at its heart. Whilst much innovative activity and process change is taking place, and whilst there is good reason to welcome its potential in the face of on-going performance critique of DWP’s centralised model (whether contracted-out or Jobcentre Plus), these new city-region approaches to governance – what can be considered the practice of *doing network governance* locally – have yet to be subject to academic critique or formal impact evaluation.

These localised approaches are important UK developments and represent a neat comparative opportunity to investigate the alternative governance strategies by which these city-regions are pursuing network accountability more fully than the UK’s highly centralised employment system has historically afforded as a means to tackle shared employment objectives. As these city-region activities expand and embed a set of key questions and priorities are emerging. In particular, there are questions as to *how* these local approaches inject resources (both in-kind services and cash), introduce and manage often complex governance frameworks (for service co-ordination and performance oversight) and instil new logics (relationships, trust and care).

A particular challenge will not only be the development of effective working partnerships across multiple partners locally and across scales (including, importantly, DWP) but, from a governance perspective, finding new ways to more effectively overlay the full range of accountability dimensions required, including network accountability. There may be tensions between each of these respective logics, cultures, aims and needs. Robust evaluation throughout will be important both to assess *whether* this new governance approach has beneficial implications in a process sense for support experiences as well as positive impacts on health and employment outcomes for the people such programmes seek to support.

This is not only a new frontier for empirical and policy research in the UK context but is also a frontier of international conceptual literature, the leading edge of which coalesces around the nature, types and realisation of governance hybridities. The UK is in many ways far behind European comparators in experiences of locally integrated employment support and it will be important to learn from those experiences. With a fuller understanding of the variation in governance types, and further research on how this variation plays out for service experience and outcomes, the need and hope is to

more effectively place market accountability levers within a more subtle and balanced governance arrangement such that markets come to be the servant to, rather than the master of, commissioners and service users.

References

- Abbott, A., and Forrest, J. (1986). Optimal Matching Methods for Historical Sequences. *The Journal of Interdisciplinary History*, 16(3), 471–494.
- Abbott, A., and Tsay, A. (2000). Sequence Analysis and Optimal Matching Methods in Sociology: Review and Prospect. *Sociological Methods & Research*, 29(1), 3–33.
- Abel, G. A., Barclay, M. E., and Payne, R. A. (2016). Adjusted indices of multiple deprivation to enable comparisons within and between constituent countries of the UK including an illustration using mortality rates. *BMJ Open*, 6(11), e012750. <https://doi.org/10.1136/bmjopen-2016-012750>
- Albertson, K., Fox, C., O’Leary, C., and Painter, G. (2018). *Payment by results and social impact bonds - Outcome-based payment systems in the UK*. Bristol: Policy Press. Retrieved from <https://policypress.co.uk/payment-by-results-and-social-impact-bonds>
- Alcock, P., Beatty, C., Fothergill, S., and Yeandle, S. (2003). *Work to Welfare: how men become detached from the labour market*. Cambridge: Cambridge University Press.
- Anyadike-Danes, M., and McVicar, D. (2005). You’ll never walk alone: Childhood influences and male career path clusters. *Labour Economics*, 12(4), 511–530.
- Arestis, P., and Sawyer, M. (2001). The Economic Analysis Underlying the ‘Third Way’. *New Political Economy*, 6(2), 255–278.
- Armstrong, D., Byrne, Y., Cummings, C., and Gallen, B. (2010). *The commissioning strategy: Provider survey on early implementation* (DWP Research Report No. 704). London: Department for Work and Pensions. Retrieved from <https://www.gov.uk/government/publications/the-commissioning-strategy-provider-survey-on-early-implementation-rr704>
- Arulampalam, W., Gregg, P., and Gregory, M. (2001). Unemployment Scarring. *The Economic Journal*, 111(475), 577–584.
- Ashworth, K., and Liu, W. C. (2001). *Jobseeker’s allowance: Transitions to Work and Early Returns to JSA* (DWP, In-house report 80).
- Atkinson, R., and Kintrea, K. (2001). Disentangling Area Effects: Evidence from Deprived and Non-deprived Neighbourhoods. *Urban Studies (Routledge)*, 38(12), 2277–2298.
- Australian Government. (2012). *Job Services Australia Star Ratings Methodology From July 2012 To June 2015*. Australian Government: Department of Jobs and Small Business. Retrieved from <https://www.jobs.gov.au/job-services-australia-provider-performance-star-ratings>
- Barbier, J.-C. (2004). *Activating social protection and the Employment insurance*. Berlin: WZB. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=CA0184D399CF97A3C04118E0C8DF402A?doi=10.1.1.623.7503&rep=rep1&type=pdf>
- Bartlett, W., and Le Grand, J. (1993). The Theory of Quasi-Markets. In J. Le Grand and W. Bartlett (Eds.), *Quasi-Markets and Social Policy* (pp. 13–34). Basingstoke: Macmillan.
- Baumol, W. (1982). Contestable Markets: An Uprising in the Theory of Industry Structure. *The American Economic Review*, 72(1), 1–15.
- Beale, I., Bloss, C., and Thomas, A. (2008). *The longer-term impact of the New Deal for Young People* (Working Paper No. 23). Norwich: Department for Work and Pensions. Retrieved from <http://dera.ioe.ac.uk/7485/1/WP23.pdf>
- Beatty, C., and Fothergill, S. (2005). The diversion from ‘unemployment’ to ‘sickness’ across British regions and districts. *Regional Studies*, 39(7), 837–854.
- Bennett, F., and Millar, J. (2005). Making work pay? *Benefits*, 13(1), 28–33.

- Black, D., Smith, J., Plesca, M., and Shannon, S. (2003). *Profiling UI Claimants to Allocate Reemployment Services: Evidence and Recommendations for States*. Washington, DC: U.S. Dept. of Labor, Employment and Training Administration. Retrieved from <https://catalog.princeton.edu/catalog/5226453>
- Blair, T., and Schroeder, G. (1998). *Europe: The Third Way/Die Neue Mitte*. Retrieved from <http://library.fes.de/pdf-files/bueros/suedafrika/02828.pdf>
- Blanchard, O. J., and Summers, L. H. (1987). Hysteresis in unemployment. *European Economic Review*, 31(1), 288–295.
- Boeri, T., Layard, R., and Nickell, S. (n.d.). Welfare-to-work and the fight against long-term unemployment: Report to Prime Ministers Blair and D'Alema. Retrieved from http://dera.ioe.ac.uk/4484/1/copy_0_blair_dalema_report.pdf
- Bonoli, G. (2010). *The Political Economy of Active Labour Market Policy* (RECWOWE Working Paper No. REC-WP 01/10). Edinburgh: Dissemination and Dialogue Centre.
- Bonoli, G. (2013). *The Origins of Active Social Policy: Labour Market and Childcare Policies in a Comparative Perspective* by Giuliano Bonoli. Oxford: Oxford University Press.
- Booth, A. L., Francesconi, M., and Frank, J. (2002). Temporary Jobs: Stepping Stones Or Dead Ends? *Economic Journal*, 112(480), F189.
- Bovaird, T., and Löffler, E. (2009). *Public Management and Governance, Second Edition*. Abingdon: Taylor & Francis.
- Bredgaard, T., and Larsen, F. (2007). Implementing public employment policy: what happens when non-public agencies take over? *The International Journal of Sociology and Social Policy*; 27(7/8), 287–300.
- Bredgaard, T., and Larsen, F. (2008). Quasi-Markets in Employment Policy: Do They Deliver on Promises? *Social Policy and Society*, 7(3), 341–352.
- Brewer, M., Clark, T., and Wakefield, M. (2005). Social security in the UK under New Labour: what did the Third Way mean for welfare reform? *Fiscal Studies*, 23(4), 505–537.
- Brewer, M., and Shephard, A. (2004). *Has Labour made work pay?* Joseph Rowntree Foundation. Retrieved from <https://www.ifs.org.uk/publications/3155>
- Brown, A. (2013). Institute for Government Blog: Working to make government more effective. Retrieved from <https://www.instituteforgovernment.org.uk/blog/disappointment-results>
- Bruttel, O. (2004). Contracting-out the Public Employment Service and the consequence for hard-to-place jobseekers: Experiences from Australia, the Netherlands and the UK. Presented at the Second Annual ESPAnet Conference, Oxford. Retrieved from <http://www.forschungsnetzwerk.at/downloadpub/ppr.11A.OB.pdf.pdf>
- Bryson, A., Dorsett, R., and Purdon, S. (2002). *The use of Propensity Score Matching in the Evaluation of Active Labour Market Policies* (Working Paper No. 4). London: Department for Work and Pensions. Retrieved from http://www.psi.org.uk/site/publication_detail/1520/
- Bryson, A., and Kasparova, D. (2003). *Profiling benefit claimants in Britain: A feasibility study* (No. 196). Norwich: Department for Work and Pensions. Retrieved from <http://eprints.lse.ac.uk/4991/>
- Brzinsky-Fay, C. (2007). Lost in Transition? Labour Market Entry Sequences of School Leavers in Europe. *European Sociological Review*, 23(4), 409–422.
- Brzinsky-Fay, C., Kohler, U., and Luniak, M. (2006). Sequence analysis with Stata. *The Stata Journal*, 6(4), 435–460.
- Buck, N. (2001). Identifying Neighbourhood Effects on Social Exclusion. *Urban Studies (Routledge)*, 38(12), 2251–2275.

- Byrne, D. (1999). *Social Exclusion*. Maidenhead: Open University Press.
- Cabinet Office. (2011). *Open Public Services White Paper*. London: The Stationery Office. Retrieved from <https://www.gov.uk/government/publications/open-public-services-white-paper>
- Caliendo, M., and Kopeinig, S. (2005). *Some Practical Guidance for the Implementation of Propensity Score Matching* (No. IZA DP No. 1588). Bonn: Institute for the Study of Labor (IZA). Retrieved from <ftp://ftp.iza.org/dps/dp1588.pdf>
- Caliński, T., and Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, 3(1), 1–27.
- Carmel, E., Papadopoulos, T., (2003). The new governance of social security in Britain. In J. Millar (Ed.), *Understanding Social Security: Issues for policy and practice* (pp. 31–52). Bristol: The Policy Press.
- Carpenter, H. (2006). *Repeat Jobseeker's Allowance spells* (Research Report No 394). Norwich: Department for Work and Pensions. Retrieved from <http://webarchive.nationalarchives.gov.uk/20090605235823/http://www.dwp.gov.uk/asd/asd5/rports2005-2006/rrep394.pdf>
- Carter, E. (Forthcoming). More than marketised? Exploring the governance and accountability mechanisms at play in Social Impact Bonds. *Journal of Economic Policy Reform*.
- Carter, E. (2018). Making markets in employment support: does the variety of quasi-market matter for people with disabilities and health conditions. In C. Needham, E. Heins, and J. Rees (Eds.), *Social Policy Review 30: Analysis and debate in social policy, 2018*. Bristol: Policy Press.
- Carter, E., and Whitworth, A. (2015). Creaming and Parking in Quasi-Marketised Welfare-to-Work Schemes: Designed Out Of or Designed In to the UK Work Programme? *Journal of Social Policy*, 44(2), 277–296.
- Carter, E., and Whitworth, A. (2017). Work Activation Regimes and Well-being of Unemployed People: Rhetoric, Risk and Reality of Quasi-Marketization in the UK Work Programme. *Social Policy & Administration*, 51(5), 796–816.
- CESI. (2013). *Briefing paper: Measuring Work Programme performance*. London: Centre for Economic and Social Inclusion.
- Clark, A. E., and Oswald, A. J. (1994). Unhappiness and Unemployment. *The Economic Journal*, 104(424), 648–659.
- Clasen, J., and Clegg, D. (2004). Does the Third Way Work? The Left and Labour Market Policy Reform in Britain, France, and Germany. In *Welfare State Change: Towards a Third Way?* Oxford: Oxford University Press.
- Coase, R. H. (1937). The Nature of the Firm. *Economica*, 4(16), 386–405.
- Coleman, N., and Lanceley, L. (2011). *Lone Parent Obligations: supporting the journey into work* (Research Report No. 736). Sheffield: Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/214510/rrep736.pdf
- Commercial Directorate Correspondence Team. (2016). Freedom of Information Act - Request for Information. Retrieved from https://www.whatdotheyknow.com/request/366570/response/908548/attach/3/Reply%20Letter%204011%20final.pdf?cookie_passthrough=1
- Conolly, A., Day, N., Hales, J., Kinsella, R., and O'Connor, W. (2010). *The operation of choice in multiple provider Employment Zones* (Research Report No. 650). London: Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/214420/rrep650.pdf

- Conservative Party. (2017). *Forward together: The Conservative manifesto, Our plan for a stronger Britain and a prosperous future*. Retrieved from www.conservatives.com/manifesto
- Considine, M. (2000). Selling the Unemployed: the Performance of Bureaucracies, Firms and Non-Profits in the New Australian 'Market' for Unemployment Assistance. *Social Policy & Administration*, 34(3), 274–295.
- Considine, M. (2001). *Enterprising States: The Public Management of Welfare-to-Work*. Cambridge: Cambridge University Press.
- Considine, M. (2005). The reform that never ends: quasi-markets and employment services in Australia. In E. Sol and M. Westerveld (Eds.), *Contractualism in Employment Services: A New Form of Welfare State Governance* (pp. 41–72). The Hague: Kluwer Law International.
- Considine, M., Lewis, J. and O'Sullivan, S. (2011). Quasi-Markets and Service Delivery Flexibility Following a Decade of Employment Assistance Reform in Australia. *Journal of Social Policy*, 40(4), 811–833.
- Crafts, N. (2005). Regional GDP in Britain, 1871–1911: Some Estimates. *Scottish Journal of Political Economy*, 52(1), 54–64.
- Creswell, J. (2009). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches* (Third). London: Sage.
- Creswell, J., and Plano Clark, V. (2011). *Designing and conducting mixed methods research* (2nd ed.). London: Thousand Oaks: Sage.
- Crisp, R., and Powell, R. (2017). Young people and UK labour market policy: A critique of 'employability' as a tool for understanding youth unemployment. *Urban Studies*, 54(8), 1784–1807.
- Cummins, S., Curtis, S., Diez-Roux, A. V., and Macintyre, S. (2007). Understanding and representing 'place' in health research: A relational approach. *Social Science & Medicine*, 65(9), 1825–1838.
- Daguerre, A., and Etherington, D. (2009). *Active labour market policies in international context: what works best? Lessons for the UK* (Working Paper No. 59). Department for Work and Pensions.
- Damm, C. (2012). *The third sector delivering employment services: an evidence review* (Working Paper No. 70). Birmingham: Third Sector Research Centre. Retrieved from <https://www.birmingham.ac.uk/generic/tsrc/documents/tsrc/working-papers/working-paper-70.pdf>
- Davies, B., and Raikes, L. (2014). *Alright for some? Fixing the Work Programme, locally*. IPPR North.
- de Graaf, W., and Sirovátka, T. (2012). Governance reforms and their impacts on the effects of activation policies. *The International Journal of Sociology and Social Policy*; 32(5/6), 353–363.
- Deakin, N., and Walsh, K. (1996). The enabling state: The role of markets and contracts. *Public Administration*, 74(1), 33–47.
- Dean, H. (2003). Re-conceptualising Welfare-To-Work for People with Multiple Problems and Needs. *Journal of Social Policy*, 32(3), 441–459.
- Deeming, C. (2015). Foundations of the Workfare State – Reflections on the Political Transformation of the Welfare State in Britain. *Social Policy & Administration*, 49(7), 862–886.
- Dehejia, R. H., and Wahba, S. (2002). Propensity Score-Matching Methods for Nonexperimental Causal Studies. *Review of Economics & Statistics*, 84(1), 151.
- De Marrais, K., and Lapan, S. (2004). *Foundations for Research: Methods of Inquiry in Education and the Social Sciences*. London: Lawrence Erlbaum Associates.
- Dingeldey, I. (2007). Between workfare and enablement – The different paths to transformation of the welfare state: A comparative analysis of activating labour market policies. *European Journal of Political Research*, 46(6), 823–851.

Disability Rights UK. (2016). *Statement on the DWP Green Paper: 'Improving lives'*. Retrieved from <https://www.disabilityrightsuk.org/statement-dwp-green--%E2%80%9Cimproving-lives%E2%80%9D>

Dockery, A. M., and Stromback, T. (2001). Developing public employment services: Preliminary assessment of the Australian experiment. *International Labour Review*, 140(4), 429.

Dolan, P., Peasgood, T., and White, M. (2008). Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. *Journal of Economic Psychology*, 29(1), 94–122.

Dorling, D. (2001). How Much does Place Matter? *Environment and Planning A: Economy and Space*, 33(8), 1335–1369.

Dorsett, R. (2004). *The New Deal for Young People: effect of the options on the labour market status of young men*. London: Policy Studies Institute.

Dorsett, R., and Lucchino, P. (2012). *Visualising the school-to-work transition: an analysis using optimal matching*. NIESR. Retrieved from <https://www.niesr.ac.uk/sites/default/files/publications/Visualising%20the%20school%20to%20work%20transition%2C%20an%20analysis%20using%20optimal%20matching.pdf>

Dorsett, R., and Lucchino, P. (2014). Explaining patterns in the school-to-work transition: An analysis using optimal matching. *Advances in Life Course Research*, 22, 1–14.

Dorsett, R., and Lucchino, P. (2016). *The Work Programme: factors associated with differences in the relative effectiveness of prime providers* (DWP ad hoc research report No. 26). London: Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/542084/ad-hoc-report-26-work-programme-relative-effectiveness-of-prime-providers.pdf

Driver, S., and Martell, L. (2000). Left, Right and the third way. *Policy and Politics*, 28(2), 147–161.

DWP. (2003). *The UK Employment Action Plan*. London: Department for Work and Pensions.

DWP. (2005). *Work and Pensions Longitudinal Study - Safeguards* (Information Directorate). Department for Work and Pensions. Retrieved from <https://www.gov.uk/government/statistics/work-and-pensions-longitudinal-study>

DWP. (2008). *DWP Commissioning Strategy*. Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/238764/7330.pdf

DWP. (2009). *Work Choice specification: Invitation to tender stage*. London: Department for Work and Pensions.

DWP. (2010a). *Ad-hoc analysis of households who have never worked*. London: Department for Work and Pensions.

DWP. (2010b). *The Work Programme Invitation to Tender: Specification and Supporting Information*. Department for Work and Pensions.

DWP. (2010c). *Work and Pensions Longitudinal Study - List of Uses* (Information Directorate). Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/202924/WPLS_Uses.pdf

DWP. (2010d). *Work Choice Provider Guidance: Section 02 Referral - Sources and Procedures*. London: Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/696754/work-choice-provider-guidance-section-2.pdf

- DWP. (2011a). Jobseeker's Allowance (JSA) Eligibility. Department for Work and Pensions. Retrieved from <https://www.gov.uk/jobseekers-allowance/eligibility>
- DWP. (2011b). *Work Programme Provider Guidance Chapter 1*. London: Department for Work and Pensions.
- DWP. (2012). *The Work Programme*. Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/49884/the-work-programme.pdf
- DWP. (2013a). Benefits sanctions - ending the 'something for nothing' culture. Retrieved from <https://www.gov.uk/government/news/benefit-sanctions-ending-the-something-for-nothing-culture>
- DWP. (2013b). *First Release: This Statistical Summary aims to give a structured overview of Work Programme Official Statistics produced by the Department for Work and Pensions*. Newcastle. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/210226/work-programme-stats-summary-june-2013_v2_020713.pdf
- DWP. (2013c). *Support for the very long term unemployed trailblazer: longer term analysis of benefit impacts*. Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/265931/svltu-dec-13-adhoc.pdf
- DWP. (2017a). *Work Programme National Statistics: Data up to June 2017*. Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/646239/work-programme-statistics-to-june-2017.pdf
- DWP. (2017b). *Work Programme Provider Guidance Chapter 11: Market Share Shift*. Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/611981/work-programme-provider-guidance-chapter-11.pdf
- DWP. (n.d.). *Work Programme National Statistics Background Information Note*. Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/692572/work-programme-statistics-background-information.pdf
- DWP. (2015). Duncan Smith: critics proved wrong on Work Programme [Press Release]. Retrieved 27 May 2018, from <https://www.gov.uk/government/news/duncan-smith-critics-proved-wrong-on-work-programme>
- Dwyer, P. (2004). Creeping Conditionality in the UK: From Welfare Rights to Conditional Entitlements? *The Canadian Journal of Sociology / Cahiers Canadiens de Sociologie*, 29(2), 265–287.
- Dwyer, P., and Wright, S. (2014). Universal Credit, ubiquitous conditionality and its implications for social citizenship. *Journal of Poverty and Social Justice*, 22(1), 27–35.
- Eichorst, W., and Konle-Seidl, R. (2008). *Contingent Convergence: A Comparative Analysis of Activation Policies* (IZA Discussion Paper No. 3905). Bonn: Institute for the Study of Labor (IZA). Retrieved from <https://www.econstor.eu/bitstream/10419/35570/1/589764675.pdf>
- Ellen, I. G., and Turner, M. A. (1997). Does neighborhood matter? Assessing recent evidence. *Housing Policy Debate*, 8(4), 833–866.
- Emmerson, C., Johnson, P., and Joyce, R. (2016). *IFS Green Budget 2016*. London: Institute for Fiscal Studies.
- Emmerson, C., Johnson, P., and Miller, H. (2012). *The IFS Green Budget: February 2012*. London: Institute for Fiscal Studies.
- English, L. M., and Skellern, M. (2005). Public–private partnerships and public sector management reform: a comparative perspective. *International Journal of Public Policy*, 1(1–2), 1–21.

- ERSA. (2013). *Work Programme Performance Report*. London: Employment Related Services Association.
- Escudero, V. (2015). *Are active labour market policies effective in activating and integrating low-skilled individuals? An international comparison* (Working Paper No. 3). Paris: International Labour Office (ILO).
- European Commission. (1993). *Growth, competitiveness, and employment. The challenges and ways forward into the 21st century, White Paper*. Luxembourg: European Commission.
- European Commission. (2015). *Identification of latest trends and current developments in methods to profile jobseekers in European Public Employment Services: Final report*. Brussels: Directorate-General for Employment, Social Affairs and Inclusion, European Commission.
- Evans, M., and Millar, J. (2006). Welfare to Work in the United Kingdom. In *International Perspectives on Welfare to Work Policy*. New York: Haworth Press.
- Everitt, B. S., Landau, S., and Leese, M. (2001). *Cluster Analysis*. London: Arnold.
- Finn, D. (2000). From full employment to employability: a new deal for Britain's unemployed? *International Journal of Manpower; Bradford*, 21(5), 384–399.
- Finn, D. (2008). *The British 'welfare market': Lessons from contracting out welfare to work programmes in Australia and the Netherlands*. York: Joseph Rowntree Foundation.
- Finn, D. (2009). The 'Welfare Market' and the Flexible New Deal: Lessons from Other Countries. *Local Economy*, 24(1), 38–45.
- Finn, D. (2010a). *Outcome Based Commissioning: Lessons from contracting out employment and skills programmes in Australia and the USA*. London: UK Commission for Employment and Skills.
- Finn, D. (2010b). Outsourcing Employment Programmes: Contract Design and Differential Prices Welfare-to-Work in the New Era: From Policy to Implementation - Steering by New Instruments. *European Journal of Social Security*, 12, 289–302.
- Finn, D. (2011a). *Job Services Australia: Design and Implementation Lessons for the British Context* (No. 752). Sheffield: Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/214527/rrep752.pdf
- Finn, D. (2011b). *The design of the Work Programme in international context: Report for the National Audit Office*. Retrieved from https://www.nao.org.uk/wp-content/uploads/2012/01/10121701_work_programme_design.pdf
- Finn, D. (2012). *Sub-Contracting In Public Employment Services: The Design and Delivery of Outcome Based and Black Box Contracts*. DG Employment, Social Affairs and Inclusion, European Commission.
- Foster, S., Colechin, J., Bivand, P., and Foster, R. (2014). *Employment support for unemployed older people*. London: Centre for Economic and Social Inclusion. Retrieved from <https://www2.learningandwork.org.uk/publications/employment-support-unemployed-older-people>
- Fournier, V., and Grey, C. (1999). Too Much, Too Little and Too Often: A Critique of du Gay's Analysis of Enterprise. *Organization*, 6(1), 107–128.
- Freud, D. (2007). *Reducing Dependency, Increasing Opportunity: Options for the future of welfare to work*. London: Department for Work and Pensions.
- 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. *Int. J. Soc. Welf. International Journal of Social Welfare*, 23(S1), S71–S86.
- Gardiner, B., Martin, R., Sunley, P., and Tyler, P. (2013). Spatially unbalanced growth in the British

economy. *Journal of Economic Geography*, 13(6), 889–928.

Gash, T., Panchamia, N., Sims, S., and Hotson, L. (2013). *Making public service markets work: Professionalising government's approach to commissioning and market stewardship*. London: Institute for Government.

Gertler, P., Martinez, S., Premand, P., Rawlings, L., and Vermeersch, C. (2016). *Impact Evaluation in Practice*. Washington, DC: World Bank Publications.

Gilbert, N. (2002). *Transformation of the welfare state: The silent surrender of public responsibility*. Oxford: Oxford University Press.

Gingrich, J. (2011). *Making markets welfare state: The politics varying market reforms*. Cambridge: Cambridge University Press.

Goulden, C. (2010). *Cycles of poverty, unemployment and low pay*. York: Joseph Rowntree Foundation.

Green, A., and Owen, D. (2006). *The geography of poor skills and access to work*. York: Joseph Rowntree Foundation

Greener, I. (2008). Markets in the public sector: when do they work, and what do we do when they don't? *Policy and Politics; London*, 36(1), 93–108.

Gregg, P. (2001). The Impact of Youth Unemployment on Adult Unemployment in the NCDS. *The Economic Journal*, 111(475), F626–F653.

Griggs, J., and Bennett, F. (2009). *Rights and Responsibilities in the Social Security System* (Social Security Advisory Committee Occasional Paper No. 6). Social Security Advisory Committee. Retrieved from <https://www.gov.uk/government/publications/ssac-occasional-paper-6-rights-and-responsibilities-in-the-social-security-system>

Griggs, J., Whitworth, A., Walker, R., McLennan, D., and Noble, M. (2008). *Person or place-based policies to tackle disadvantage? Not knowing what works*. York: Joseph Rowntree Foundation.

Grout, P. and Stevens, M. (2003). The Assessment: Financing and Managing Public Services. *Oxford Review of Economic Policy*, 19(2), 215–234.

Grover, C. (2007). The Freud Report on the Future of Welfare to Work: Some Critical Reflections Commentary & Issues. *Critical Social Policy*, 27, 534–545.

Haddon, C. (2012). *Making policy in opposition: the Work Programme, 2007-2010*. London: Institute for Government.

Hales, J., Taylor, R., Mandy, W., and Miller, M. (2003). *Evaluation of Employment Zones: report on a cohort survey of long-term unemployed people in the Zones and a matched set of comparison areas* (Research Report No. 176). London: Department for Work and Pensions.

Halpin, B. (2012). *Sequence Analysis of Life-course Data: A Comparison of Distance Measures* (Department of Sociology Working Paper Series No. WP2012-02). Limerick: University of Limerick. Retrieved from <https://ulsites.ul.ie/sociology/sites/default/files/sociology-sequence-analysis-of-Life-course-data.pdf>

Halpin, B. (2016). *Cluster Analysis Stopping Rules in Stata* (Working Paper No. WP2016-01). Limerick: Department of Sociology, University of Limerick. Retrieved from https://ulsites.ul.ie/sociology/sites/default/files/wp2016-01_0.pdf

Halpin, B. (2017). SADI: Sequence analysis tools for Stata. *The Stata Journal*, 17(3), 546–572.

Halpin, B., and Cban, T. W. (1998). Class Careers as Sequences: An Optimal Matching Analysis of Work-Life Histories. *European Sociological Review*, 14(2), 111–130.

Harkness, S., Gregg, P., and MacMillan, L. (2012). *Poverty: the role of Institutions, Behaviours and Culture*. York: Joseph Rowntree Foundation.

- Hart, O., Shleifer, A., and Vishny, R. W. (1997). The Proper Scope of Government: Theory and an Application to Prisons. *Quarterly Journal of Economics*, 112(4), 1127–1161.
- Heckman, J. J., Heinrich, C., and Smith, J. (2002). The Performance of Performance Standards. *The Journal of Human Resources*, 37(4), 778–811.
- Heckman, J. J., Lalonde, R. J., and Smith, J. A. (1999). Chapter 31 - The Economics and Econometrics of Active Labor Market Programs. In O. C. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics* (Vol. 3, pp. 1865–2097). Cambridge, MA: Elsevier.
- Heidenreich, M., and Graziano, P. R. (2014). Lost in activation? The governance of activation policies in Europe. *International Journal of Social Welfare*, 23(S1), S1–S5.
- Heinrich, C., Maffioli, A., and Vázquez, G. (2010). *A Primer for Applying Propensity-Score Matching* (Technical Notes No. No. IDB-TN-161). Washington, DC: Inter-American Development Bank. Retrieved from <https://publications.iadb.org/handle/11319/1681>
- Heins, E., and Bennett, H. (2016). ‘Best of Both Worlds’? A Comparison of Third Sector Providers in Health Care and Welfare-to-Work Markets in Britain. *Social Policy & Administration*, 50(1), 39–58.
- Henman, P., and Fenger, M. (2006). *Administering Welfare Reform: International Transformations in Welfare Governance*. Bristol: Policy Press.
- Hepple, B. (2013). Back to the Future: Employment Law under the Coalition Government. *Industrial Law Journal*, 42(3), 203–223.
- Hills, J. (1998). *Thatcherism, new Labour and the welfare state* (Centre for Analysis of Social Exclusion No. CASEpaper, 13). London: London School of Economics and Political Science. Retrieved from <http://eprints.lse.ac.uk/5553/>
- Hipp, L. and Warner, M. (2008). Market Forces for the Unemployed? Training vouchers in Germany and the USA. *Social Policy & Administration*, 42(1), 77–101.
- Hirsch, D., and Millar, J. (2004). *Labour’s welfare reform: progress to date*. York: Joseph Rowntree Foundation.
- Hirschman, A. O. (1970). *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.
- HMRC. (2014). *The Real Time Information Pilot*. Her Majesty’s Revenue and Customs. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/224619/RTI_Pilot_Report_130711-2.doc
- Holmes, E. (2014). *Work 2.0. Helping the hardest to help: targeted assistance, incentives and the Work Programme*. London: Policy Exchange.
- Holmstrom, B., and Milgrom, P. (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, & Organization*, 7, 24–52.
- Hood, C. (1991). A Public Management for All Seasons? *Public Administration*, 69(1), 3–19.
- Inclusion. (2011). *Work Programme results: perform or bust?* Centre for Economic and Social Inclusion. Retrieved from https://www2.learningandwork.org.uk/sites/default/files/response_downloads/perform_or_bust.pdf
- Jacobs, A. (1998). Seeing difference: Market health reform in Europe. *Journal of Health Politics, Policy & Law*, 23(1), 1.
- Jantz, B., and Jann, W. (2013). Mapping accountability changes in labour market administrations: from concentrated to shared accountability? *International Review of Administrative Sciences*, 79(2), 227–248.

- Jantz, B., Klenk, T., Larsen, F., and Wiggan, J. (2015). Marketization and Varieties of Accountability Relationships in Employment Services: Comparing Denmark, Germany, and Great Britain. *Administration & Society*, 50(3), 321–345.
- Jessop, B. (1993). Towards a Schumpeterian Workfare State? Preliminary Remarks on Post-Fordist Political Economy. *Studies in Political Economy*, 40(1), 7–39.
- Jessop, B. (1998). The rise of governance and the risks of failure: the case of economic development. *International Social Science Journal*, 50(155), 29–45.
- Kemp, P., Bradshaw, J., Dornan, P., Finch, N., and Mayhew, E. (2004). *Routes out of poverty: A research review*. York: Joseph Rowntree Foundation.
- Kettl, D. F. (1993). *Sharing Power: Public Governance and Private Markets*. Brookings Institution Press.
- Knuth, M. (2014). Broken Hierarchies, Quasi-markets and Supported Networks - A Governance Experiment in the Second Tier of Germany's Public Employment Service. *Social Policy & Administration*, 48(2), 240–261.
- Koning, P., and Heinrich, C. (2013). Cream-Skimming, Parking and Other Intended and Unintended Effects of High-Powered, Performance-Based Contracts. *Journal of Policy Analysis and Management*, 32(3), 461–483.
- Kooiman, J. (2003). *Governing As Governance*. Thousand Oaks, Calif: SAGE Publications Ltd.
- Köppe, S. (2017). Britain's new housing precariat: housing wealth pathways out of homeownership. *International Journal of Housing Policy*, 17(2), 177–200.
- Koppell, J. G. S. (2005). Pathologies of Accountability: ICANN and the Challenge of 'Multiple Accountabilities Disorder'. *Public Administration Review*, 65(1), 94–108.
- Lane, P., Foster, R., Gardiner, L., Lanceley, L., and Purvis, A. (2013). *Work Programme Evaluation: Procurement, supply chains and implementation of the commissioning model* (No. 832). Sheffield: Department for Work and Pensions.
- Layard, R. (2004). *Good Jobs and Bad Jobs* (Working Paper). London: Centre for Economic Performance, London School of Economics and Political Science. Retrieved from <http://eprints.lse.ac.uk/28757/1/OP019.pdf>
- Layard, R., Nickell, S., and Jackman, R. (1991). *Unemployment: Macroeconomic Performance and the Labour Market*. Oxford: Oxford University Press.
- Le Grand, J. (1991). Quasi-Markets and Social Policy. *The Economic Journal*, 101(408), 1256–1267.
- Le Grand, J., and Bartlett, W. (1993). Introduction. In J. Le Grand and W. Bartlett (Eds.), *Quasi-Markets and Social Policy* (pp. 1–12). Basingstoke: Macmillan.
- Leckie, G., and Goldstein, H. (2016). *The evolution of school league tables in England 1992-2016: 'contextual value added', 'expected progress' and 'progress 8'* (Bristol Working Papers in Education Series No. 02/2016). Bristol: University of Bristol. Retrieved from <https://www.bristol.ac.uk/media-library/sites/education/documents/bristol-working-papers-in-education/The%20evolution%20of%20school%20league%20tables%20in%20England%201992-2016.pdf>
- Leitch, S. (2006). *Leitch Review of Skills: Prosperity for All in the Global Economy - World Class Skills*. London: The Stationery Office.
- Lester, P. (2013). *Review of Work Programme Minimum Performance Levels*. London: Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/209112/paul-lester-review.pdf

- Leuven, E., and Sianesi, B. (2003). *PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing*. Retrieved from <https://econpapers.repec.org/software/bocbocode/S432001.htm>
- Levy, J. (2004). Activation through Thick and Thin: Progressive Approaches to Labour Market Activation. In N. Ellison, L. Bauld, and M. Powell (Eds.), *Social Policy Review 16*. Bristol: Policy Press.
- Lindsay, C., McQUAID, R. W., and Dutton, M. (2007). New Approaches to Employability in the UK: Combining ‘Human Capital Development’ and ‘Work First’ Strategies? *Journal of Social Policy*, 36(4), 539–560.
- Lipp, R. (2005, January). *Job Seeker Profiling: The Australian Experience*. Presented at the EU-Profiling seminar, Nuremberg. Retrieved from http://doku.iab.de/veranstaltungen/2005/profiling2005_Lipp_Australia.pdf
- Lister, R. (1997). A Differentiated Universalism. In *Citizenship: Feminist Perspectives* (pp. 66–90). Palgrave, London.
- Lister, R. (1998). Citizenship and Difference: Towards a Differentiated Universalism. *European Journal of Social Theory*, 1(1), 71–90.
- Lister, R., and Bennett, F. (2010). The new ‘champion of progressive ideals’? *Renewal: A Journal of Labour Politics; London*, 18(1/2), 84–109.
- Lødemel, I., and Moreira, A. (2014). *Activation Or Workfare?: Governance and Neo-liberal Convergence*. Oxford: Oxford University Press.
- Lødemel, I., and Trickey, H. (2001). *‘An offer you can’t refuse’*. Bristol: Policy Press.
- Lowery, D. (1998). Consumer Sovereignty and Quasi-Market Failure. *Journal of Public Administration Research and Theory: J-PART*, 8(2), 137–172.
- Loxha, A., and Morgandi, M. (2014). *Profiling the Unemployed: A Review of OECD Experiences and Implications for Emerging Economies* (Social Protection and labor discussion paper No. 1424). Washington, DC: World Bank Group.
- Lundsgaard, J. (2002). *Competition and Efficiency in Publicly Funded Services*. Paris: OECD.
- LWI. (2016a). *Disability and health reform plans are welcome – but we need to do a lot more and do it better*. London: Learning and Work Institute. Retrieved from <http://www.learningandwork.org.uk/2016/10/31/disability-and-health-reform-plans-are-welcome-we-need-do-lot-more-and-do-it-better/>
- LWI. (2016b). *Halving the Gap: Making the Work and Health Programme work for disabled people*. London: Learning and Work Institute. Retrieved from http://www.learningandwork.org.uk/sites/niace_en/files/files/Work%20and%20Health_Halving_the_Gap_FINAL_1.pdf
- Macdonald, R., Shildrick, T., and Furlong, A. (2014). In search of ‘intergenerational cultures of worklessness’: Hunting the Yeti and shooting zombies. *Critical Social Policy*, 34(2), 199–220.
- Mair, V. (2013). Government departments are ‘cooling’ on payment-by-results, conference participants suggest. *Civil Society*. Retrieved from <https://www.civilsociety.co.uk/news/government-departments-are--cooling--on-payment-by-results--conference-participants-suggest.html>
- Marshall, T. (1950). *Citizenship and social class and other essays*. Cambridge: Cambridge University Press.
- Martin, J. P. (2015). Activation and active labour market policies in OECD countries: stylised facts and evidence on their effectiveness. *IZA Journal of Labor Policy*, 4, 4.
- Martin, R., Pike, A., Tyler, P., and Gardiner, B. (2016). Spatially Rebalancing the UK Economy:

Towards a New Policy Model? *Regional Studies*, 50(2), 342–357.

Mason, T. (2010). Shaw Trust relieved at DWP decision on Work Choice contracts. *Civil Society*. Retrieved from <https://www.civilsociety.co.uk/news/shaw-trust-relieved-at-dwp-decision-on-work-choice-contracts.html>

Matty, S. (2013). *Predicting likelihood of long-term unemployment: the development of a UK jobseekers' classification instrument* (No. Working paper No 116). Sheffield: Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/210303/WP116.pdf

McConnell, S., Burwick, A., Perez-Johnson, I., and Winston, P. (2003). *Privatization in Practice: Case Studies of Contracting for TANF Case Management: Final report*. Washington, DC: Mathematica.

McIvor, N. (n.d.). *The Work and Pensions Longitudinal Study (WPLS)*. Department for Work and Pensions. Retrieved from <http://www.eurim.org.uk/activities/pi/060223talk.pdf>

McKnight, A. (2002). Low-paid Work: Drip feeding the Poor. In J. Hills, J. LeGrand, and Piachaud (Eds.), *Understanding Social Exclusion*. Oxford: Oxford University Press.

McKnight, A. (2015). *The Coalition's Record on Employment: Policy, Spending and Outcomes 2010-2015* (Working Paper No. 15). London: London School of Economics and Political Science. Retrieved from <http://sticerd.lse.ac.uk/dps/case/spcc/WP15.pdf>

McQuaid, R. W., Green, A., and Danson, M. (2005). Introducing Employability. *Urban Studies*, 42(2), 191–195.

McQuaid, R. W., and Lindsay, C. (2005). The Concept of Employability. *Urban Studies*, 42(2), 197–219.

McVicar, D., and Anyadike-Danes, M. (2002). Predicting successful and unsuccessful transitions from school to work by using sequence methods. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 165(2), 317–334.

Meager, N., Newton, B., Sainsbury, R., Corden, A., and Irvine, A. (2014). *Work Programme Evaluation: the participant experience report* (No. Research Report No 892). London: Department for Work and Pensions. Retrieved from <https://www.gov.uk/government/publications/work-programme-evaluation-participant-experience>

Meagher, G., and Goodwin, S. (2015). Introduction: Capturing marketisation in Australian social policy. In G. Meagher and S. Goodwin (Eds.), *Markets, Rights and Power in Australian Social Policy* (pp. 1–27). Sydney: Sydney University Press.

Milligan, G. W., and Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, 50(2), 159–179.

Morphy, L., Henderson, R., and Fraser, C. (2012). *The Programme's Not Working: Experiences of homeless people on the Work Programme*. Homeless Link; St Mungo's; Crisis. Retrieved from <https://www.homeless.org.uk/sites/default/files/site-attachments/The%20Programme's%20Not%20Working%202012.pdf>

Mulheirn, I. (2011). *Will the Work Programme Work? Examining the future viability of the Work Programme*. London: Social Market Foundation.

Mulheirn, I., Foley, B., Menne, V., and Prendergrast, J. (2009). *Vicious Cycles: Sustained employment and welfare reform for the next decade*. London: Social Market Foundation.

NAO. (2007). *Sustainable employment: supporting people to stay in work and advance*. London: National Audit Office.

NAO. (2012). *Introduction of the Work Programme*. London: National Audit Office.

NAO. (2014). *The Work Programme* (Report by the Comptroller and Auditor General). London:

National Audit Office.

NAO. (2015). *Outcome-based payment schemes: government's use of payment by results*. London: National Audit Office.

Newman, J. (2001). *Modernizing Governance: New Labour, Policy and Society*. London: Sage.

Newton, B., Meager, N., Bertram, C., Corden, A., George, A., Lalani, M., ... Weston, K. (2012). *Work Programme evaluation: Findings from the first phase of qualitative research on programme delivery* (No. 821). London: Department for Work and Pensions. Retrieved from <https://www.gov.uk/government/publications/work-programme-evaluation-findings-from-the-first-phase-of-qualitative-research-on-programme-delivery-rr821>

Oakley, M. (2015). *Employment support for a high-wage economy*. York: Joseph Rowntree Foundation.

O'Connell, P., McGuinness, S., and Kelly, E. (2010). *A Statistical Profiling Model of Long-Term Unemployment Risk in Ireland* (No. Working Paper No. 345). Dublin: The Economic and Social Research Institute. Retrieved from <https://www.esri.ie/publications/a-statistical-profiling-model-of-long-term-unemployment-risk-in-ireland/>

OECD. (1994). *The OECD jobs study: facts, analysis, strategies*. Paris: OECD.

OECD. (2001). *Innovations in Labour Market Policies The Australian Way: The Australian Way*. Paris: OECD Publishing.

OECD. (2013). *OECD Employment Outlook 2013*. Paris: OECD Publishing.

OECD. (2014). *Connecting people with jobs: Activation policies in the United Kingdom*. Paris: OECD Publishing.

OECD. (2018). *Working age population: OECD Labour Force Statistics*. Paris: OECD Publishing.

Olson, M. (1965). *The Logic of Collective Action: Public Goods and the Theory of Groups, Second Printing with New Preface and Appendix*. Cambridge, MA: Harvard University Press.

Osborne, D., and Gaebler, T. (1992). *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector*. New York: Plume.

PAC, Public Accounts Committee. (2008). *Sustainable employment: supporting people to stay in work and advance* (No. Thirteenth Report of Session 2007–08). London: The Stationery Office.

PAC, Public Accounts Committee. (2012). *Department for Work and Pensions: The Introduction of the Work Programme* (No. Eighty-fifth report of session 2010–12). London: The Stationery Office.

PAC, Public Accounts Committee. (2013). *Department for Work and Pensions: Work Programme Outcome Statistics* (No. Thirty-third report of session 2012–13). London: The Stationery Office.

Pattie, C., Whitworth, A., and Johnston, R. (2015). Does campaign contact influence individuals' vote choices? an alternative approach. *European Political Science*, 14(3), 279–297.

Paul, K. I., and Moser, K. (2009). Unemployment impairs mental health: Meta-analyses. *Journal of Vocational Behavior*, 74(3), 264–282.

Payne, C., and Payne, J. (2000). *Early Identification of the Long-term Unemployed* (No. PSI Research Discussion Paper 4). London: Policy Studies Institute. Retrieved from

Payne, R., and Abel, G. (2012). *UK indices of multiple deprivation - a way to make comparisons across constituent countries easier* (Health Statistics Quarterly 53). Office for National Statistics. Retrieved from <http://webarchive.nationalarchives.gov.uk/20151014023659/http://www.ons.gov.uk/ons/rel/hsq/health-statistics-quarterly/no--53--spring-2012/index.html>

Peck, J. (1998). Workfare in the sun: politics, representation, and method in U.S. welfare-to-work

- strategies. *Political Geography*, 17(5), 535–566.
- Peck, J. (2001). *Workfare States*. London: Guilford Press.
- Peck, J., and Jones, M. (1995). Training and Enterprise Councils: Schumpeterian Workfare State, or What? *Environment and Planning A: Economy and Space*, 27, 1361–1396.
- Peck, J., and Theodore, N. (2000). Commentary. ‘Work first’: workfare and the regulation of contingent labour markets. *Cambridge Journal of Economics*, 24(1), 119–138.
- Pierre, J., and Peters, B. G. (2000). *Governance, Politics, and the State*. Macmillan.
- Pollitt. (2003). *The Essential Public Manager*. Maidenhead: Open University Press.
- Pollock, G., Antcliff, V., and Ralphs, R. (2002). Work orders: Analysing employment histories using sequence data. *International Journal of Social Research Methodology*, 5(2), 91–105.
- Powell, A. (2018). *Work and Health Programme* (Briefing Paper No. 7845). London: House of Commons Library.
- Powell, M. (2000). New Labour and the third way in the British welfare state: a new and distinctive approach? *Critical Social Policy*, 20(1), 39–60.
- Powell, M. (2015). Making Markets in the English National Health Service. *Social Policy & Administration*, 49(1), 109–127.
- Prätorius, R. (2003). Debating Governance: Whatever That Means. *Journal of Public Administration Research & Theory*, 13(2), 235.
- Purvis, A., Foster, S., Lanceley, L., and Wilson, T. (2014). *Fit for purpose: Transforming employment support for disabled people and those with health conditions*. London: Centre for Economic and Social Inclusion.
- Purvis, A., Foster, S., Lane, P., Aston, J., and Davies, M. (2013). *Evaluation of the Work Choice Specialist Disability Employment Programme: Findings from the 2011 Early Implementation and 2012 Steady State Waves of the research* (No. 846). London: Department for Work and Pensions. Retrieved from <https://www.gov.uk/government/publications/evaluation-of-the-work-choice-specialist-disability-employment-programme-rr846>
- Quintini, G., and Manfredi, T. (2009). *Going Separate Ways? School-to-Work Transitions in the United States and Europe*. Paris: OECD.
- Quirke, D. (2015). *All’s Fair in the Work Programme? Understanding the equalities impact of the Work Programme in London: an extended London Councils Briefing*. London: London Councils. Retrieved from https://www.londoncouncils.gov.uk/sites/default/files/Policy%20themes/Tracking%20welfare%20ref%20orms/150420_WPEqualitiesAnalysis_v6b_combined.pdf
- Rafferty, A., and Wiggan, J. (2011). Choice and Welfare Reform: Lone Parents’ Decision Making around Paid Work and Family Life. *Journal of Social Policy*, 40(2), 275–293.
- Rees, J., Taylor, R., and Damm, C. (2013). *Does sector matter? Understanding the experiences of providers in the Work Programme* (Working Paper No. 92). Birmingham: Third Sector Research Centre.
- Rees, J., Whitworth, A., and Carter, E. (2013). *Support for all in the UK Work Programme? Differential payments, same old problem...* (No. Working Paper 115). Birmingham: Third Sector Research Centre.
- Rees, J., Whitworth, A., and Carter, E. (2014). Support for All in the UK Work Programme? Differential Payments, Same Old Problem. *Social Policy & Administration*, 48(2), 221–239.
- Rhodes, R. a. W. (1996). The New Governance: Governing without Government1. *Political Studies*, 44(4), 652–667.

- Riccucci, N. M. (2010). Envisioning Public Administration as a Scholarly Field in 2020: Rethinking Epistemic Traditions. *Public Administration Review*, 70, S304–S306.
- Riley, T., Bivand, P., and Wilson, T. (2014). *Making the Work Programme work for ESA claimants: analysis of minimum performance levels and payment models*. London: Centre for Economic and Social Inclusion. Retrieved from <https://www2.learningandwork.org.uk/publications/making-work-programme-work-esa-claimants-analysis-minimum-performance-levels-and-paymen>
- Romzek, B. S., LeRoux, K., and Blackmar, J. M. (2012). A Preliminary Theory of Informal Accountability among Network Organizational Actors. *Public Administration Review*, 72(3), 442–453.
- Rosenbaum, P. R., and Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1), 41–55.
- Rubin, D. B. (1980). Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment. *Journal of the American Statistical Association*, 75(371),
- Rubin, D. B. (1986). Statistics and Causal Inference: Comment: Which Ifs Have Causal Answers. *Journal of the American Statistical Association*, 81(396), 961–962.
- Sainsbury, R. (2014). Introduction: Universal Credit: the story so far... *Journal of Poverty and Social Justice*, 22(1), 11–13.
- Sainsbury, R. (2017). Activation in the UK: The Frontline and the ‘Black Box’ of Employment Service Provision. In R. van Berkel, D. Caswell, P. Kupka, and F. Larsen (Eds.), *Frontline Delivery of Welfare-to-Work Policies in Europe: Activating the Unemployed* (pp. 53–69). Abingdon: Routledge.
- SASC. (2017). *In-work progression and Universal Credit* (Occasional Paper No. 19). London: Social Security Advisory Committee. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/657842/ssac-occasional-paper-19-in-work-progression-and-universal-credit.pdf
- Savas, E. (2000). *Privatization and public-private partnerships*. New York: Chatham House.
- Scherer, S. (2001). Early Career Patterns: A Comparison of Great Britain and West Germany. *European Sociological Review*, 17(2), 119–144.
- Schutz, P., Chambless, C., and DeCuir, J. (2004). Multimethods Research. In K. deMarrais and S. Lapan (Eds.), *Foundations for Research: Methods of Inquiry in Education and the Social Sciences* (pp. 267–282). London: Lawrence Erlbaum Associates.
- Scottish Parliament. (2013). *The Impact of Welfare Reform on Scotland. 2nd report of the Welfare Reform Committee* (No. 303).
- Sheldrick, B. (2000). The Contradictions of Welfare to Work: Social Security Reform in Britain. *Studies in Political Economy*, 62(1), 99–121.
- Shildrick, T. (2012). Low pay, no pay churning: the hidden story of work and worklessness. *Poverty*, (142), 6–9.
- Shildrick, T., Macdonald, R., Webster, C., and Garthwaite, K. (2012). *Poverty and insecurity: Life in low-pay, no-pay Britain*. Bristol: Policy Press.
- Simmons, D. (2011). Child Poverty Action Group. Retrieved from <http://www.cpag.org.uk/content/escalating-conditionality>
- Slater, T. (2012). The Myth of ‘Broken Britain’: Welfare Reform and the Production of Ignorance. *Antipode*, 46(4), 948–969.
- Slife, B., and Williams, R. (1995). *What’s behind the research? Discovering hidden assumptions in the behavioural sciences*. London: Thousand Oaks: Sage.

- Sol, E., and Westerveld, M. (2005). *Contractualism in Employment Services. A New form of Welfare State Governance*. The Hague: Kluwer Law International.
- Soukup, T. (2011). Profiling: Predicting Long-Term Unemployment at the Individual Level. *Central European Journal of Public Policy*, (Vol. 4/1), 118–143.
- StataCorp. (2013). *Stata Multivariate Statistics Reference Manual Release 13*. College Station, Texas: Stata Press.
- Stoker, G. (1998). Governance as theory: five propositions. *International Social Science Journal*, 50(155), 17–28.
- Strickland, P. (1996). *Jobseeker's Allowance* (House of Commons Library Research Paper No. 96/5). Retrieved from <http://researchbriefings.files.parliament.uk/documents/RP96-5/RP96-5.pdf>
- Struyven, L., and Steurs, G. (2005). Design and redesign of a quasi-market for the reintegration of jobseekers: empirical evidence from Australia and the Netherlands. *Journal of European Social Policy*, 15(3), 211–229.
- Suleiman, R. (2014). *Stepping Stones: The role of the voluntary sector in future welfare to work schemes*. London: NCVO. Retrieved from http://www.ncvo.org.uk/images/documents/about_us/media-centre/stepping-stones-vol-sec-and-welfare-to-work-work-schemes-jul-2014.pdf
- Sunley, P., Martin, R., and Nativel, C. (2001). Mapping the New Deal: Local Disparities in the Performance of Welfare-to-Work. *Transactions of the Institute of British Geographers*, 26(4), 484–512.
- Sunley, P., Martin, R., and Nativel, C. (2006). *Putting Workfare in Place: Local Labour Markets and the New Deal*. Oxford: Blackwell.
- Sutherland, H., Evans, M., Hancock, R., and Zantomio, F. (2008). *The impact of benefit and tax uprating on incomes and poverty*. York: Joseph Rowntree Foundation. Retrieved from
- Tarr, A., and Finn, D. (2012). *Implementing Universal Credit: will the reforms improve the service for users?* York: Joseph Rowntree Foundation.
- Tenbenschel, T. (2005). Multiple modes of governance. *Public Management Review*, 7(2), 267–288.
- Theodore, N. (2007). New Labour at work: long-term unemployment and the geography of opportunity. *Cambridge Journal of Economics*, 31(6), 927–939.
- Theodore, N., and Peck, J. (2001). Searching for best practice in welfare-to-work: The means, the method and the message. *Policy and Politics; London*, 29(1), 81–94.
- Thomas, A. (2017). *The Work and Health Programme*. Presented at the Understanding the Work and Health Programme: The Sequel, Learning and Work Institute, Birmingham. Retrieved from <https://www2.learningandwork.org.uk/WorkHealthJan2017>
- Thompson, A., Trennell, P., Hope, M., and McPhillips, A. (2011). *Work Choice Evaluation: Commissioning and Transition of Clients to the Programme* (In-House Research No. 6). Sheffield: Department for Work and Pensions.
- Thompson, S. (2015). *The low-pay, no-pay cycle*. York: Joseph Rowntree Foundation.
- Tomlinson, M., and Walker, R. (2010). *Recurrent poverty: the impact of family and labour market changes*. York: Joseph Rowntree Foundation.
- Tonge, J. (1999). New packaging, old deal? New Labour and employment policy innovation. *Critical Social Policy*, 19(2), 217–232.
- Torfig, J. (1999). Workfare With Welfare: Recent Reforms of the Danish Welfare State. *Journal of European Social Policy*, 9(1), 5–28.
- Treib, O., Bähr, H., and Falkner, G. (2007). Modes of governance: towards a conceptual clarification. *Journal of European Public Policy*, 14(1), 1–20.

- Trickey, H., and Walker, R. (2001). Steps to compulsion within British labour market policies. In Lødemel, I., and Trickey, H. (Eds.) *'An offer you can't refuse': Workfare in international perspective*. Bristol: Policy Press.
- Turok, I., and Webster, D. (1998). The New Deal: Jeopardised by the geography of unemployment? *Local Economy*, 12(4), 309–328.
- UK Statistics Authority. (2013). *Final Report on Statistics Relating to the DWP Work Programme and Pre-Work Programme* (Monitoring Review). London: UK Statistics Authority. Retrieved from <https://www.statisticsauthority.gov.uk/publication/final-report-on-statistics-relating-to-the-dwp-work-programme-and-pre-work-programme/>
- van Berkel, R., and Borghi, V. (2007). New modes of governance in activation policies. *The International Journal of Sociology and Social Policy*, 27(7/8), 277–286.
- van Berkel, R., de Graaf, W. de, and Sirovátka, T. (2012). Governance of the activation policies in Europe: Introduction. *International Journal of Sociology and Social Policy*, 32(5/6), 260–272.
- van Berkel, R., Sager, F., and Ehrler, F. (2012). The diversity of activation markets in Europe. *International Journal of Sociology and Social Policy*, 32(5/6), 273–285.
- van Berkel, R., and van der Aa, P. (2005). The marketization of activation services: a modern panacea? Some lessons from the Dutch experience. *Journal of European Social Policy*, 15(4), 329–343.
- van Ewijk, H. (2009). *European Social Policy and Social Work: Citizenship-Based Social Work*. Routledge.
- van Ham, M., Manley, D., Bailey, N., Simpson, L., and MacLennan, D. (Eds.). (2012). *Neighbourhood Effects Research: New Perspectives*. Dordrecht: Springer.
- Vegeris, S., Adams, L., Oldfield, K., Bertram, K., Davidson, R., Durante, L., and Vowden, K. (2011). *Flexible New Deal Evaluation: Customer Survey and Qualitative Research Findings*. London: Department for Work and Pensions. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/214533/rrep758.pdf
- Waddell, G., and Burton, A. K. (2006). *Is work good for your health and well-being?* London: The Stationery Office.
- Walker, R., and Kellard, K. (2001). *Staying in Work: Policy Overview* (Research Report No. 265). Norwich: Department for Education and Employment. Retrieved from <http://webarchive.nationalarchives.gov.uk/20130402140845/https://www.education.gov.uk/publications/eOrderingDownload/RR265.PDF>
- Weishaupt, T. (2011). *From the Manpower Revolution to the Activation Paradigm: Explaining Institutional Continuity and Change in an Integrating Europe*. Amsterdam: Amsterdam University Press.
- White, S. (2000). Review Article: Social Rights and Social Contract—Political Theory and the New Welfare Politics. *British Journal of Political Science*, 30(3), 507–532.
- Whitworth, A. (2012). Lone Parents and Welfare-to-Work in England: A Spatial Analysis of Outcomes and Drivers. *Social Policy & Administration*, 47(7), 826–845.
- Whitworth, A. (2013). *Single parents' experiences of employment support from Jobcentre Plus and the Work Programme*. London: Gingerbread.
- Whitworth, A., and 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), 274–289.

- Whitworth, A., and Griggs, J. (2013). Lone Parents and Welfare-to-work Conditionality: Necessary, Just, Effective? *Ethics and Social Welfare*, 7(2), 124–140.
- Wiggan, J. (2012). Telling Stories of 21st Century Welfare: The UK Coalition Government and the Neo-Liberal Discourse of Worklessness and Dependency Themed Issue: Social Policy in an Age of Austerity. *Critical Social Policy*, 32, 383–405.
- Wiggan, J. (2015a). Varieties of marketisation in the UK: examining divergence in activation markets between Great Britain and Northern Ireland 2008–2014. *Policy Studies*, 36(2), 115–132.
- Wiggan, J. (2015b). What variety of employment service quasi-market? Ireland’s JobPath as a private power market. In Z. Irving, M. Fenger, and J. Hudson (Eds.), *Social Policy Review 27: Analysis and debate in social policy, 2015* (pp. 151–168).
- Williamson, O. (1975). *Markets and hierarchies: analysis and antitrust implications: a study in the economics of internal organization*. New York: The Free Press.
- Williamson, O. E. (1979). Transaction-Cost Economics: The Governance of Contractual Relations. *Journal of Law & Economics*, 22, 233–262.
- Wilson, T. (2017). *Work and Health Programme: We don’t know nothing*. Presented at the Understanding the Work and Health Programme, Learning and Work Institute, Birmingham. Retrieved from <https://www2.learningandwork.org.uk/thesequelpresentations>
- Wilson, J. P. (2013). *The Routledge Encyclopaedia of UK Education, Training and Employment: From the Earliest Statutes to the Present Day*. Abingdon: Routledge.
- WPSC. (2009). *DWP’s Commissioning Strategy and the Flexible New Deal* (No. Second Report of Session 2008–09). London: Work and Pensions Select Committee, House of Commons.
- WPSC, Work and Pensions Select Committee (2013). *Can the Work Programme Work for All User Groups?* (No. First Report of Session 2013–14). London: House of Commons.
- WPSC, Work and Pensions Select Committee (2015). *Welfare-to-work* (No. Second Report of Session 2015–16). London: House of Commons.
- WPSC, Work and Pensions Select Committee (2016). *In-work progression in Universal Credit* (No. Tenth Report of Session 2015–16). London: House of Commons.
- WPSC, Work and Pensions Select Committee (2011). *Work Programme: Providers and Contracting Arrangements* (No. Fourth Report of Session 2010–12). London: House of Commons.
- Wu, C.-F., Cancian, M., and Meyer, D. R. (2008). Standing Still or Moving Up? Evidence from Wisconsin on the Long-Term Employment and Earnings of TANF Participants. *Social Work Research*, 32(2), 89–103.
- Zehavi, A. (2012). Welfare State Politics in Privatization of Delivery: Linking Program Constituencies to Left and Right. *Comparative Political Studies*, 45(2), 194–219.
- Zijl, M., van den Berg, G., and Heyma, A. (2004). *Stepping Stones for the Unemployed: The Effect of Temporary Jobs on the Duration until Regular Work* (Discussion paper No. 1241). Bonn: Institute for the Study of Labor (IZA).
- Ziliak, S. T., and McCloskey, D. N. (2008). *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives*. Ann Arbor: University of Michigan Press.

Appendix 1

Employment and benefit spell data cleaning for the production of variables on benefit claiming and employment histories

To mitigate the limitations of the historic employment data used in the analysis a process of cleaning and adjusting of the data takes place before it is used. This process is heavily informed by the methods used by in-house DWP analysts and is described in ...

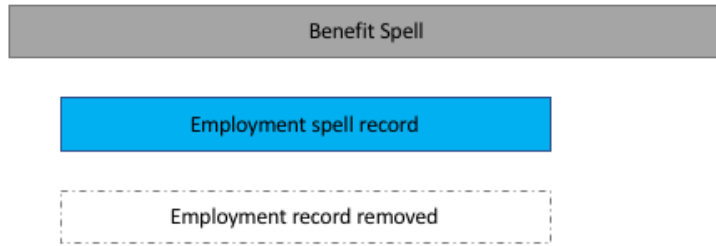
The cleaning of the P45/46 employment spell data begins by removing any inappropriate data, adjusting unreliable dates, and removing duplicate records. The following types of employment spell record are removed:

- Records with no personal identifier;
- Records where the employment start date is later than the employment end date;
- Records with end dates before the beginning of the data series (a cut off set 5-years prior to participant attachment date);
- Occupational pension records;
- Known benefit spells that are duplicated in DWP administration data; and
- Records starting in the future;

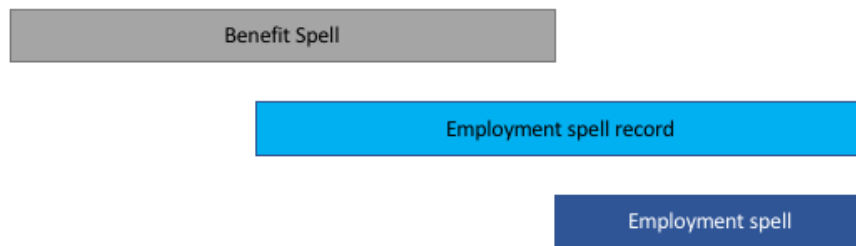
Start and end dates are often very approximate in the P45/46 records. Where HMRC do not know the date on which an employment spell started, they assign a start date of the 6th April in the year that they become aware of the employment spell. This may not be the actual year in which the spell began. A similar process occurs when HMRC do not know the date on which an employment spell ended.

The employment spell data is examined in conjunction with DWP benefit records, which are reliably recorded on administration systems. Employment spell start and end dates are then adjusted using the start or end date of the nearest benefit record under the assumption that people move directly from benefits to employment or vice versa. Overlapping spells are handled as follows:

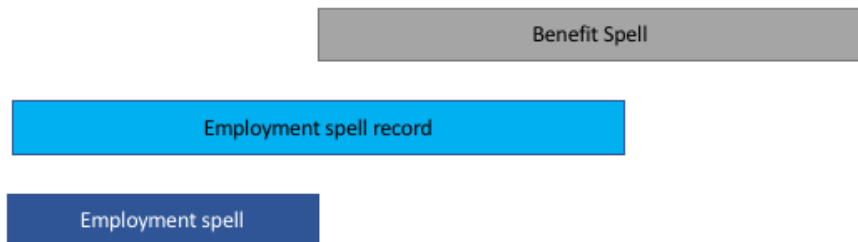
1. When an out-of-work benefit spell wholly surrounds an employment spell the employment spell is removed entirely.



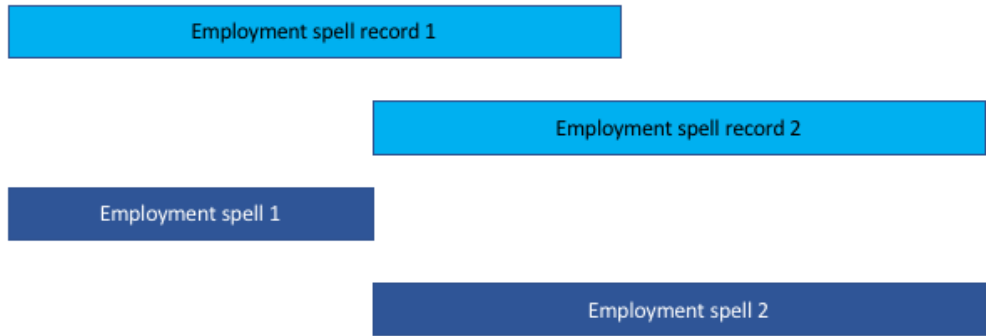
2. When a benefit spell overlaps the start of an employment spell the employment spell start is adjusted to the end of the benefit spell.



3. When a benefit spell overlaps the end of an employment spell – the employment spell end is adjusted to the start of the benefit spell



4. When an overlap occurs across multiple employment records the earlier spell record is curtailed (starts are understood to be more accurately recorded than end dates) so that when adding up ‘time in employment’ across multiple spells, this does not exceed the maximum extent of calendar days in which a person has active employment records:



The cleaned records are then used to calculate the amount of time in paid employment prior to participating in the programme.

Appendix 2

Preparation of administrative data

<i>Variable</i>	<i>Definition</i>	<i>Source</i>	<i>Values</i>	<i>Cleaning and coding procedure</i>
<i>Ageband</i>	Participant age at point of programme referral, broken into age brackets.	LMS	1 18 - 24; 2 25 - 34; 3 35 - 44; 4 45 - 54; 5 55 - 59; 6 60 +	Age is the age at point of referral to the Work Programme calculated using participant date of birth information from the LMS dataset. If the dob field is missing information or appears incorrectly populated, for example, calculated age means that a participant's age at referral date is beyond the valid range for programme participation, the age variable is replaced with 'age' variable from older and/or alternative records for the same participant. In some cases this is also missing, ultimately the final number missing the age variable is 96 cases and these are dropped from the analysis.
<i>AnyIB</i>	Any spell claiming Incapacity Benefit (or other disability-related out-of-work benefit, ESA) within 5 year period preceding participation in the Work Programme	NBD	1 yes has IB record; 0 no	Where cases are missing this means there has been an unsuccessful connection from the WPAD basefile to the NBD dataset. There are other datasets which could be used to provide benhistory information but these were not available through the assigned DWP data access business case. 2,309 cases in the WPAD basefile were not successfully matched to NBD and as a result were removed from the analysis (this is 0.15% of the sample)
<i>Attdat</i>	Date that participant's programme attachment was recorded within the PRaP system	WPAD	Dates from 05 June 2011 onwards	A small proportion of cases (<2%) in the WPAD base file were referred to, but not attached to the Work Programme. These cases are removed from the empirical analysis developed here, which uses data only from those participants who were successfully attached to the programme
<i>Benhist</i>	Type of out-of-work benefit claimed at point of Work Programme (or Work Choice) referral as captured through alternative eligibility and	LMS OPPORTUNITIES (also available in WPAD)	1 JSA; 2 JSA ex-IB; 3 contribution based ESA and esa credit; 4 ESA income related; 5 ESA ex-IB; 6 IB/IS; (missing values removed)	This is constructed from the 'OTPTYP' variable which captures the alternative eligibility and payment group options for programme participants. At the point of developing this variable 69 cases in the WPAD core dataset were missing OTPTYPE information. However, all of those cases with 'missing' OTPTYP information had a duplicate record in WPAD (i.e. multiple rows per unique id), and so the more fully populated record was kept. No cases were dropped as a result of missing information.

	referral routes into programme				
<i>care_lv</i>	Prior to programme participation JCP staff raised 'care leaver' as a specific Work Programme barrier, that is, someone who may be disadvantaged by an experience of being a 'looked after child' within the local authority care system.	WPAD	1 marker present; 0 marker not present; missing		Only 96 cases have a marker against this record.
<i>carer</i>	Participant is known by JCP to have caring responsibilities	LMS	1 marker present; 0 marker not present, missing values are presumed to be non-care leaver		79 % of cases have missing values and these are presumed to be 'non carers'. 6,768 cases are recorded as having caring responsibilities.
<i>CPA</i>	Contract Package Area in which programme participant resides and where Work Programme services are received	WPAD	Numerical codes from 1 – 18.		All cases in WPAD have viable, non-missing CPA information.
<i>Contract</i>	Prime contract to which participant is referred	WPAD	Numerical codes		Very few cases in WPAD have missing contract information. Where cases are missing contract information they are also missing other key sociodemographic details (incomplete records) and such records are removed from the analysis.
<i>Cumuldis</i>	An indicator of the degree to which day-to-day activities are adversely impacted by the participant's health condition or disability.	WPAD	Integers from 0 - 10		This is constructed as a composite indicator by bringing together 10 different markers on the ways in which everyday activities are adversely impacted by a participant's disability or health condition. It is produced by aggregating the responses to questions outlined in Annex XX to produce a score (out of 10) where a higher value indicates a higher number of challenges.

<i>daysesa2</i>	Number of days recorded in receipt of ESA or other disability-related out-of-work benefit in 2-years preceding programme participation	NBD	Continuous variable from 0 - 730	The detailed procedure for cleaning benefit spell records is described in Appendix 1.
<i>daysesa5</i>	Number of days recorded in receipt of ESA or other disability-related out-of-work benefit in 5-years preceding programme participation	NBD	Continuous variable from 0 - 1826	
<i>daysjsa2</i>	Number of days recorded in receipt of JSA in 2-years preceding programme participation	NBD	Continuous variable from 0 - 730	
<i>daysjsa5</i>	Number of days recorded in receipt of JSA in 5-years preceding programme participation	NBD	Continuous variable from 0 - 1826	
<i>depch</i>	Programme participant has a dependent child or children aged under 16	LMS	1 has dependent child or children (aged under 16); 0 does not have dependent children aged under 16	There are considerable inconsistencies in the fields which supported the development of this variable: 'has children' and 'date of birth of youngest child'. Where cases had a valid date for dob of youngest child but were missing information on the number of children, the cases were 'given' dependent children. Where cases had a record for dependent children, but the dob of youngest child suggested that all children were now aged over 16, cases were recoded to 0.
<i>depch_5</i>	Programme participant has a dependent child aged 5 years or younger at point of programme participation	LMS	1 has youngest child aged 5 or below; 0 does not have dependent children or youngest child is aged over 5	This variable is principally constructed from the field 'date of birth of youngest child'. Where this field was missing, but within the recent 'spells' (as held by NBD) a participant's record indicated that a child had entered the family within the preceding 5 years, the marker '1' was given.

<i>disadva</i>	Programme participant is understood to be in a disadvantaged labour market position due to ex-offender status or through use of drugs and/or alcohol	WPAD originally LMS	0 not disadvantaged in this way; 1 'Ex-Offender'; 2 'Misuser of Drugs'; 3 'Ex-Offender and Misuser of Drugs'; 4 'Misuser of Alcohol'; 5 'Ex-Offender and Misuser of Alcohol'; 6 'Misuser of Drugs and Misuser of Alcohol'; 7 'Ex-Offender, Misuser of Drugs and Misuser of Alcohol'	This field is populated by bringing together the 'disadvantage' markers provided on the WPAD dataset and which have been populated by Jobcentre plus staff in LMS. There are no issues with missing data in this field. Where staff do not provide information against these particular markers of disadvantage it is assumed that the issue is not present '0'. Even though the classification of categories may seem highly specific (which would usually prevent such granular analysis in survey work) the large number of cases available to this study mean that there is not an issue with small cell counts. There are no combinations of disadvantage (cells) with fewer than 1800 cases.
<i>ethnicity</i>	Programme participant's ethnic group	LMS	1 white British; 2 white Irish; 3 other white; 4 = white and black mixed; 5 white and Asian; 6 other mixed; 7 Indian; 8 Pakistan; 9 Bangladesh; 10 Other Asian; 11 Black Caribbean; 12 Black African; 13 other Black; 14 Chinese; 15 Other; 16 prefer not to say/none	The coding for different ethnic groups within the LMS dataset has changed over time and these alternative codings were reconciled to produce a single set of codes. Some participants elect not to provide this data; rather than dropping these cases an additional category is set up for those where ethnicity information has not been provided.
<i>empdays2</i>	Number of days recorded as in paid employment in 2-years preceding programme participation	NBD and P45/46	Continuous variable from 0 - 730 (though no-one at very high end of the scale due to programme eligibility criteria)	
<i>empdays5</i>	Number of days recorded as in paid employment in 5-years preceding programme participation	NBD and P45/46	Continuous variable from 0 - 1826 (though no-one at very high end of the scale due to programme eligibility criteria)	
<i>empspell</i>	Number of distinct spells of employment recorded in 5-	NBD and P45/46	0 no employment spell; integers to 20	

	years preceding programme participation.				
<i>esa_spell</i>	Number of distinct ESA spells (or spells claiming other disability-related out-of-work benefits) within 5-year period preceding WP participation	NBD and P45/46	Integer		
<i>ex_army</i>	Prior to programme participation JCP staff raised 'ex armed forces' as a specific Work Programme barrier	WPAD	1 marker present; 0 marker not present; missing		Only 158 cases are recorded as ex-armed forces personnel.
<i>ex_carer</i>	Participant is identified by JCP staff as a former 'carer'.	WPAD	1 ex-carer or 0 missing not applicable		There is very little documentation as to how this characteristic is identified and only 2073 cases (less than 0.13%) of the sample express this characteristic.
<i>ex_off</i>	Prior to programme participation JCP staff raised 'ex offender' as a specific Work Programme barrier	WPAD	1 marker present; 0 marker not present; missing		33,186 cases are recorded as ex-offenders within this field. This is a lower incidence than is recorded in the variables supporting the 'disadva' variable although there are considerable inconsistencies in the cross tabs between these two variables. The 'disadva' variable is used in further analysis because it is more finegrained in terms of identification of particular and compound disadvantages; and because the suite of pre-Work Programme disadvantage variables (of which ex-off is a part) appear to be particularly poorly populated.
<i>homeless</i>	Participant is identified by JCP staff as being homeless	LMS and WPAD	1 homeless; 0 not homeless		There are two variables which capture information on participants who are understood to be homeless. 'CLNPWAFG' is the 'person without accommodation flag' collected by JCP as a matter of course and a question on homelessness is also situated in the battery of 'WP barriers' markers, which are populated prior to programme referral (ORF_CIA_WP_BAR_3). There are some inconsistencies between these two markers and so a decision was made to combine these two indicators. Any current record of 'homelessness' was used to populate this variable as '1'; all other cases are coded '0'.
<i>ICD_group</i>	International Classification of Disease Code for participant (where provided)	LMS	0 no ICD code; 1 mental & behavioural problems excluding depression, anxiety, alcoholism, stress & drug abuse, which		In the original 'ICDCODE' variable missing and 0 values indicate participants without any diagnosis code on the system. Missing values are recoded to 0. The original variable sourced through LMS 'ICDCODE' contains 1062 separate codes for highly specific classifications of diseases; there are also a number of

are presented separately; 2 diseases of nervous system excluding epilepsy which is coded separately; 3 diseases of circulatory or respiratory system; 4= Diseases of the Musculoskeletal system and Connective Tissue excluding back pain and arthritis which is coded separately; 5=Injury, Poisoning and certain other consequences of external causes; 6 Other; 7 depression; 8 anxiety; 9 back pain; 10 alcoholism; 11 pain (all pain except back pain); 12 epilepsy; 13 arthritis; 14 stress; 15 drug abuse

duplicate codes within this series (for example code 235 and 236 both relate to Emphysema). A series of cross tabs and collapses were performed to identify the most frequent codes for Work Programme participants. The resultant coding brings a degree of granularity across different conditions without introducing fragmentation associated with small cell counts.

<i>Job Outcome date</i>	Date on which job outcome payment is triggered	WPAD	Date	Two separate fields containing job outcome dates are available in PrAP and DWP analysts provided information on the field which had been subject to validation. This is the variable which is used in the production of official programme statistics.
<i>Job Outcome binary</i>	Participant has achieved a payable Work Programme job outcome	WPAD	0 participant has not achieved job outcome; 1 participant has achieved a job outcome payment	Binary variable is constructed from job outcome dates; where job outcome date is missing, this value is set to 0; where a date for job outcome is recorded the code is set to 1
<i>language</i>	LMS record indicates that participant has problems with written and/or spoken English	LMS	0 no issue; 1 has language issue	1,248,949 cases (79.56%) of sample have missing value on language variable; missing is recoded to 0.
<i>lone_p</i>	Participant is a lone-parent	LMS	0 participant is lone parent; 1 participant is lone parent aged 21 or under; 2 participant is lone parent aged 22 or over at time of referral to programme	Previous research shows differential achievement of job outcomes by lone parents according to their age. To accommodate the divergence in the likely direction of effect sizes associated with this variable, additional granularity is set into the original JCP marker, by combining this with information on participant age. Cases with missing data are presumed not to be lone parents.
<i>lowqual</i>	Participant has low or no qualifications	LMS	0 participant does not have marker for low qualification; 1 participant answers	In the original field provided in LMS nearly 70% of Work Programme participants have missing data. Despite only having responses from 30 % of the sample, 123,040 programme participants (7.84%) have

			'yes' to having low or no qualifications; 2 participant answers 'no' to having low or no qualifications	a marker indicating a potential labour market challenge from low or no qualifications. The other LMS variable which contains information on qualification level 'HIGHQUAL' which contains detail on the highest qualification level attained by programme participants is even more sparsely populated. In order to retain this information a three part coding structure identified - those who were not asked or did not respond to this question; those who identified as having low-qualifications; and those who actively responded 'no' in response to low qualification questions.
<i>ment_h</i>	Prior to programme participation JCP staff raised 'mild to moderate mental health issues' as a specific Work Programme barrier	WPAD	1 marker present; 0 marker not present; missing	Only 1,013 cases (0.06% of records) have a marker against this field; 80% of cases are missing. This would suggest that the mental health marker is severely under populated and is therefore not used in further analysis.
<i>Noemp</i>	Within 5-year period preceding Work Programme participation participant has no record of paid employment	NBD and P45/46	1 participant has no record of paid work; 0 participant has record of paid work	The detailed procedure for combining employment and benefit spell records is described in section XX. Where cases are missing P45/46 employment records it is assumed that programme participants had no time in employment within this time period.
<i>Postcode</i>	Home address postcode of programme participant	LMS	Postcode e.g. OX2 6GG	Postcode information required considerable cleaning as there were inconsistencies in the formatting of postcode details within LMS. Subsequent to cleaning 4,415 cases (0.12% of sample) had postocde details which were not succesfully merged onto a national postcode look up file, which was used to connect postcodes to higher-level geographies (LSOA to LA). Cases which could not be connected to the postcode look up file could not be connected to local area context information and were therefore dropped from further analysis.
<i>Refdat</i>	Date that participant's programme referral was recorded within the PRaP system	WPAD	Dates from 01 June 2011 onwards	All cases in WPAD have viable, non-missing referral dates.
<i>refugee</i>	Participant is identified by JCP staff as being a refugee or	WPAD	0 None selected (default); 1 Refugee; 2 Other humanitarian protection; 3	Refugee and 'other humanitarian protection' are collapsed to a single binary marker (1); all other responses are coded 0. 19,941 cases are recorded as 1 against refugee status (1.27% of sample)

	having other humanitarian protection status		Refugee - No; 4 Prefer not to Say; 5 Not collected	
<i>sex</i>	Sex of programme participant as recorded in 'CCSEX' variable	WPAD	0 Female; 1 Male	419 cases have missing data against this variable (0.03 % of sample) and are removed from the analysis
<i>sr_disab</i>	Participant has a self-reported disability or long-term health condition	LMS	1 has disability; 0 does not have disability	Only 419 cases (0.03% of sample) have missing data for this variable. Missing values are presumed to be people who are without a self-reported disability and are recoded to 0
<i>substn</i>	Prior to programme participation JCP staff raised 'history of substance dependency' as a specific Work Programme barrier	1 marker present; 0 marker not present; missing	Only 756 cases (0.05% of records) have a marker against this field; 80% of cases are missing.	
<i>Sustainment</i>	Count of the number of sustainment payments made to Prime for participant's sustained period in employment	WPAD	0 where none; else integer up to maximum number of eligible sustainment payments	Those who are missing sustainment payment information are presumed not to have achieved sustained employment and code is set to 0

Appendix 3

Full logistic regression output underpinning Chapter 5.

	Odds Ratio	Standard Error	p-value
<i>JSA ex-IB</i>	1.645	0.027	0.000
<i>Contribution ESA</i>	0.854	0.039	0.001
<i>Income-related ESA</i>	0.558	0.022	0.000
<i>ESA ex-IB</i>	0.404	0.018	0.000
<i>IB/IS</i>	1.021	0.051	0.672
<i>Days of ESA in 2y</i>	1.000	0.000	0.000
<i>Any time on IB</i>	0.940	0.008	0.000
<i>Spells of ESA in 5y</i>	0.883	0.004	0.000
<i>Self-reported disability</i>	0.762	0.005	0.000
<i>Mental health</i>	0.568	0.027	0.000
<i>Nervous system</i>	0.706	0.042	0.000
<i>Circulatory or respiratory</i>	0.769	0.038	0.000
<i>Musculoskeletal</i>	0.829	0.040	0.000
<i>Injury</i>	1.490	0.065	0.000
<i>Other ICD class</i>	0.885	0.037	0.003
<i>Depression</i>	0.797	0.032	0.000
<i>Anxiety</i>	0.783	0.034	0.000
<i>Back pain</i>	0.797	0.036	0.000
<i>Alcoholism</i>	0.752	0.039	0.000
<i>Pain (excluding back pain)</i>	0.856	0.043	0.002
<i>Epilepsy</i>	0.562	0.034	0.000
<i>Arthritis</i>	0.656	0.038	0.000
<i>Stress</i>	1.019	0.052	0.720
<i>Drug abuse</i>	0.718	0.042	0.000
<i>Cumulative impact of disability</i>	0.934	0.003	0.000
<i>Age 25 - 34</i>	0.680	0.004	0.000
<i>Age 35 - 44</i>	0.628	0.004	0.000
<i>Age 45 - 54</i>	0.553	0.003	0.000
<i>Age 55 - 59</i>	0.399	0.004	0.000
<i>Age 60 +</i>	0.208	0.004	0.000
<i>Ex-offender</i>	0.586	0.006	0.000
<i>Misuser of drugs</i>	0.482	0.019	0.000
<i>Ex-offender and misuser of drugs</i>	0.346	0.015	0.000
<i>Misuser of alcohol</i>	0.541	0.020	0.000
<i>Ex-offender and misuser of alcohol</i>	0.471	0.031	0.000
<i>Misuser of drugs and alcohol</i>	0.565	0.043	0.000
<i>Ex-offender and misuser of drugs and alcohol</i>	0.404	0.031	0.000
<i>Low-qualifications</i>	0.689	0.005	0.000
<i>Not low-qualifications</i>	1.122	0.005	0.000
<i>Challenges with English language</i>	0.850	0.007	0.000
<i>Homeless</i>	0.696	0.024	0.000

<i>Other white ethnicity</i>	1.082	0.012	0.000
<i>Mixed ethnicity</i>	0.973	0.013	0.035
<i>Indian ethnicity</i>	1.127	0.018	0.000
<i>Pakistan ethnicity</i>	1.065	0.013	0.000
<i>Bangladesh ethnicity</i>	1.194	0.021	0.000
<i>Other Asian ethnicity</i>	1.228	0.024	0.000
<i>Black Caribbean ethnicity</i>	0.954	0.010	0.000
<i>Black African ethnicity</i>	1.175	0.012	0.000
<i>Other Black ethnicity</i>	1.051	0.022	0.016
<i>Chinese ethnicity</i>	1.217	0.053	0.000
<i>Other ethnicity</i>	1.144	0.015	0.000
<i>Ethnicity - prefer not to say</i>	1.062	0.009	0.000
<i>Refugee or other humanitarian protection</i>	1.228	0.020	0.000
<i>Days in employment</i>	1.001	0.000	0.000
<i>No employment record</i>	0.775	0.004	0.000
<i>Number of employment spells</i>	1.028	0.001	0.000
<i>Male*</i>	0.960	0.004	0.000
<i>Has children*</i>	1.354	0.013	0.000
<i>Male*has children</i>	1.016	0.012	0.158
<i>Woman with young child</i>	0.597	0.009	0.000
<i>Lone parent aged under 21</i>	0.671	0.062	0.000
<i>Lone parent aged over 21</i>	1.354	0.013	0.000
<i>Local authority JSA rate</i>	0.947	0.002	0.000
<i>IMD decile 2</i>	0.986	0.016	0.361
<i>IMD decile 3</i>	0.955	0.015	0.003
<i>IMD decile 4</i>	0.913	0.013	0.000
<i>IMD decile 5</i>	0.869	0.012	0.000
<i>IMD decile 6</i>	0.843	0.012	0.000
<i>IMD decile 7</i>	0.808	0.011	0.000
<i>IMD decile 8</i>	0.767	0.010	0.000
<i>IMD decile 9</i>	0.732	0.010	0.000
<i>IMD decile 10</i>	0.672	0.009	0.000
<i>Observations</i>	1,563,874		
<i>Pseudo-R²</i>	0.0958		