UNIVERSITY OF
LEADING
THE WAY
WESTMINSTER⌗

WestminsterResearch

http://www.westminster.ac.uk/research/westminsterresearch

Improving DWP assessment of the relative costs and benefits of employment programmes

David Greenberg
Genevieve Knight
Stefan Speckesser
Debra Hevenstone

A report of research carried out by the Policy Studies Institute on behalf of the Department for Work and Pensions

Improving DWP assessment of the relative costs and benefits of employment programmes

by David Greenberg, Genevieve Knight,
Stefan Speckesser and Debra Hevenstone

DWP Department for Work and Pensions

Department for Work and Pensions

Working Paper No 100

Improving DWP assessment of the relative costs and benefits of employment programmes

David Greenberg, Genevieve Knight, Stefan Speckesser and Debra Hevenstone

Contents

List of tables

List of figures

Acknowledgements

The Authors

David Greenberg is a Visiting Research Fellow at the Policy Studies Institute (PSI) and professor emeritus of economics at the University of Maryland Baltimore County (UMBC). Much of his research focuses on the evaluation of government programmes. With his co-authors, Professor Greenberg has the fourth edition of a widely used textbook on cost benefit analysis (CBA), several journal articles, a guide for conducting and using cost analyses of employment and training programmes, and a meta-analysis of findings from cost benefit analyses of American welfare-to-work programmes. He taught a course in CBA at UMBC for many years and taught similar courses elsewhere.

Genevieve Knight is a Principal Research Fellow with the PSI Her main field of research is the evaluation of public policy programmes within the labour market. Genevieve has worked on a broad range of evaluations since joining PSI in 1998. Since 2003, Genevieve has been working in the field of CBA.

Stefan Speckesser was a Senior Lecturer in Economics at the University of Westminster Business School. Stefan is now a Principle Research Fellow at the Institute for Employment Studies. Stefan is an empirical economist using a wide range of experimental and non-experimental methodologies. He has worked on numerous quantitative evaluation studies of welfare-to-work programmes.

Debra Hevenstone is a Research Fellow at the PSI. Debra has a PhD in Sociology and Public Policy from the University of Michigan, 2005, with an MA in Social Science from the University of Chicago, and a BA in Social Science from Bard College. Debra has experience in qualitative programme evaluation, advanced statistical methods, network analysis, and agent-based modelling as well as specialities in the fields of demography, social stratification, and labour economics.

Abbreviations

CBA	Cost benefit analysis
CBAs	Cost benefit analyses
DWP	Department for Work and Pensions
EZ	Employment Zone
ESA	Employment and Support Allowance
ESF	European Social Fund
IB	Incapacity Benefit
IS	Income Support
JSA	Jobseeker's Allowance
HMRC	HM Revenue and Customs
LMS	Labour Market System
NBD	National Benefits Database
ND	New Deal
ND25+	New Deal for individuals aged 25 plus
ND50+	New Deal for individuals aged 50 plus
NDDP	New Deal for Disabled People
NDLP	New Deal for Lone Parents
NDP	New Deal for Partners
NDYP	New Deal for Young People
PSI	Policy Studies Institute
WBLA	Work-Based Learning for Adults
WPLS	Work and Pensions Longitudinal Survey
WTC	Working Tax Credit

Glossary of terms

Cost Benefit Analysis	Cost Benefit Analysis attempts to quantify in monetary terms the value of as many of the consequences of employment and training programmes as possible to determine whether their benefits outweigh their costs.
Cost Benefit Framework	Cost Benefit Framework specifies how the Department for Work and Pensions' cost benefit analyses should be conducted and reported.
Cost-effectiveness	Cost-effectiveness pertains to programme costs and benefits relative to other programmes.
Displacement effects	Displacement effects occur if a firm expands at the expense of other firms because its employment costs are subsidised by the Government.
Employment and training programmes	Employment and training programmes provide services (e.g. job search, training, and subsidise jobs) to non-workers on benefits to help them move into paid employment.
Entry and deterrent effects	Entry effects occur where a programme is perceived as beneficial, but is available only to benefit recipients, and some persons may take up benefits in order to qualify. Deterrent effects can occur in the case of mandatory programmes for benefit recipients, some individuals who might otherwise have taken up benefits may decide not to do so in order to avoid the 'hassle' of participating in the programme.
Equilibrium wage effects	Equilibrium wage effects can theoretically occur when those in an employment and training programme search harder for jobs or their job skills increase and, as a result, their weeks or hours are greater than they otherwise would have been, then the resulting increase in labour supply will tend to put downward pressure on the equilibrium wage within the labour markets in which they work. If the programme tends to reduce job vacancies – for example, by making job search more efficient or imparting job skills that are in demand – this will also increase downward wage pressures. Thus, workers who are employed in the same labour markets as programme participants could receive lower wages than otherwise would be the case. Notice that if wage rates are lower than they would be without a programme, then this will tend to mitigate substitution effects resulting from the programme.

Fiscal substitution effects	Fiscal substitution effects can occur for 'wage subsidy programmes' that pay subsidies to private sector employers that hire members of specific disadvantaged target groups or for 'public sector employment' programmes that directly place targeted disadvantaged individuals into jobs at government agencies or non-profit institutions. In this case, the targeted workers may be hired instead of, or even replace, those who are not targeted and, hence, are more expensive to employ.
General equilibrium effects	General equilibrium effects are those where employment and training programmes may have effects on the well-being of those who are not enrolled in the programme and, because of this, on well-being of the economy. Such effects include substitution effects, displacement effects, equilibrium wage effects, multiplier effects, and entry effects.
Incapacity benefits	Incapacity benefits are cash benefits available to disabled persons who are not working or who are working only a few hours a week.
Impacts	Impacts are programme effects (e.g. on employment and benefit receipts) on programme participants.
Income Support	Income Support is a non-contributory, income-assessed benefit available to people who are not required to work.
Jobseeker's Allowance	Jobseeker's Allowance provides cash benefits to the unemployed.
Micro-simulation model	Micro-simulation models use data on individuals (i.e. micro-data) to predict how changes in government programmes will affect individuals and what the changes will cost the government.
Multiplier	Multiplier is a value greater than one that is used to adjust estimates of expenditures if it is expected that they generate benefits to the economy that exceed their immediate value.
New Deal for Disabled People	New Deal for Disabled People is a voluntary programme that provides counselling and services through Job Brokers to help the disabled enter employment.
New Deal for Lone Parents	New Deal for Lone Parents is a voluntary programme that provides counselling and services through Jobcentre Plus to help unemployed lone parents enter employment.
New Deal for Partners	New Deal for Partners offers advisory support to partners of claimants of Jobseeker's Allowance (JSA)/Incapacity Benefit/Income Support/Severe Disablement Allowance/Carer's Allowance.
ND25+	New Deal for individuals aged 25 plus is a mandatory programme for longer term JSA claimants (over 18 months) who are aged over 25 years.

ND50+	New Deal for individuals aged 50 plus is a voluntary programme with specialist advice for those aged over 50 years.
Net benefits	Net benefits are the benefits of a programme less the costs of the programme.
Quantiles	In this context, quantiles are subgroupings – points taken at regular intervals from the distribution of a continuous outcome variable, such as the earnings distribution.
Sensitivity tests	Sensitivity tests are used to determine how estimates of benefits and cost change when the assumptions on which they are based are changed.
Substitution effects	Substitution effects occur if participants in a programme take jobs that individuals who did not participate would otherwise have held.

Summary

This report was produced with the primary aim of informing cost benefit estimates within the Cost-Benefit Framework (CBF). The Department for Work and Pensions' (DWP's) CBF is a guidance document for the production of cost-benefit information. The research consisted of two components – a literature review and new empirical estimates from DWP administrative data.

General equilibrium in cost benefit analyses

This report reviews what is known about general equilibrium effects and discusses when they are likely to be important. We also present estimates and recommendations to account for these effects in cost benefit analyses (CBAs). These estimates can be used to guide sensitivity tests.

The key conclusions about treating general equilibrium effects in conducting CBAs of employment and training programmes are:

- base-case estimates should first be derived by assuming no general equilibrium effects;

- for substitution effects, in conducting sensitivity analyses of demand-side employment and training programmes that subsidise employers for employing certain categories of disadvantaged workers or place such workers directly into jobs, 60 per cent can reasonably be used as the upper bound value of the short-run substitution effect and 30 per cent as the lower bound value. However, in reporting the sensitivity results, it should be emphasised that findings from the survey approach imply that they are considerably smaller. It should be emphasised that even if as large as 60 per cent **if the economy is growing**, then these substitution effects would be expected to diminish over time. However, for supply-side programmes of job search assistance, training, and financial incentives that encourage or help individuals to seek or enter employment, use ten or 20 per cent as the upper bound and zero as the lower bound. For both, a caution should be added about the existence of substitution effects by briefly reviewing the evidence presented in Tables 2.2 and 2.3 of this report;

- for equilibrium wage effects, sensitivity tests of the base-case for **large-scale demand-side programmes** might be conducted by assuming that there is a small increase in wages of one or two per cent. Similar sensitivity tests for supply-side programmes do not seem appropriate because the evidence concerning these effects is too weak;

- for entry effects, the likely direction and importance of these effects should be discussed in reporting cost-benefit findings of employment and training programmes.

Subgroup impacts and the distribution of impacts

Subgroups can be defined for:

- exogenous variables (variables that are not affected by a programme such as gender and age);

- endogenous variables (variables affected by a programme such as hours worked); and

- continuous outcome variables such as earnings or duration of benefit for which impacts of a programme may vary for different subgroupings of that variable.

The key recommendations are:

- if there is a subgroup which is of particular interest, a subgroup impact would better assess how the programme meets the needs of this group;

- when the estimated programme impacts appear to differ among the subgroups, separate CBAs for each subgroup is a natural next step;

- important informative subgroups can be formed from programme delivery aspects or variations in cost;

- subgroups should be defined on the basis of **status prior to programme entry**;

- it is important to test whether programme impacts on different subgroups differ statistically from one another (for example, a t-test of the statistical significance of the difference between the subgroup impact sizes), not just whether the individual impacts are statistically significant (for example, a t-test of the statistical significance of the impact from zero);

- where subgroup impacts are produced and found, it is important to explore the delivery of services to these subgroups to help understand if the subgroup impacts are attributable to intrinsic differences between the groups or because they were treated differently under the programme;

- significance tests are rarely possible for programme operating costs, although there are exceptions and so this possibility should always be explored;

- subgroups to be investigated should be defined at the outset of the evaluation;

- to conduct a CBA using subgroup impacts requires that costs are collected for these subgroups.

Specifically for quantile subgroup impacts of the continuous outcomes like earnings or duration of employment or benefit claim:

- in general, the target of the programme needs to be carefully considered and the quantile treatment effects are most needed where a particular part of an outcome distribution (of earnings or benefit duration) may be the desired target. This is relevant where one might consider targeting/limiting programme access to a particular group in the future;

- quantile impacts might be most successfully estimated for programmes where participation has been randomly assigned.

Duration of benefits and employment, and wages

We would recommend to routinely include in impact evaluations the gross estimates for the average:

- number of days employed in the first 12 and 24 months following programme participation;

- number of days on benefit in the first 12 and 24 months following participation;

- duration of participants' first jobs following programme participation; and

- wage or earnings of participants' first jobs following the programme participation.

Chapter 5 and Appendix A show the estimates of the above variables for New Deal programmes. These estimates can be used in CBAs if more desirable net impact estimates do not exist.

Other recommendations are:

- Net impacts on earnings and benefit and employment duration of a programme should be carried out when a comparison group can be constructed.

- All evaluations should be required to provide information about the data processing rules applied for overlapping or inconsistent information as these rules can affect the measured impact values.

- Ideally, there should be an agreed standardised set of basic data processing rules for overlapping and inconsistent information in the DWP employment and benefits administrative data.

- While descriptive evidence on gross outcomes from the National Benefits Database, such as those estimated here can improve cost benefit estimates, estimating net programme impacts is more difficult since a comparison group must be identified. A comparison group can be non-participating benefit claimants with similar characteristics, or those not eligible for a programme due to living in a non-pilot area, for example.

- In estimating and reporting net impacts, it is important to explain the reasons for choosing a comparison group, in order to help establish the validity of this group as suitable and relevant for the evaluation design.

- It is recommended that in the establishment of a new policy, careful design for a pilot should take place and include a design for the evaluation which enables identification of a net impact. In the absence of a pilot, included in the design of the policy should be an evaluation design which enables net impacts and consideration of value for money for the policy programme to be established. In discussing a planned policy, an explanation of the considerations made in this design should be made, and such an explanation is also needed in the case where a decision is made not to establish an evaluation design for a policy, together with discussion of the implications this has for identifying the net impacts and value for money for the policy.

Multiple participation

This research explores whether evaluation studies can be readily redesigned to address multiple participation.

Several types of multiple participation are explored:

- a programme can offer a set of different programme options at one time (multiple programmes);

- an individual can be on more than one programme at a single point in time (programme combinations);

- an individual can participate in the same programme more than once (repetition – programme combinations);

- an individual can participate in different programmes subsequently over the life-course (participation dynamics of when people participate with the programme combinations they undertake).

For all types of multiple participation, a descriptive assessment of the programme complexity (with regard to the observed scale of what programme choices occur for participants and the description of the types of participant in each) is required as part of impact analyses. This assessment needs to be included in the CBA.

For better assessment of programme participation, we recommend that:

- more quantitative management information on (caseworker) selection for services is recorded into the Labour Market System (and an equivalent for private provision);

- that the quality of recording is better maintained and checked; and

- that this information is used for evaluation purposes.

The first two types of multiple participation require impact estimation to apply methods which account for the various alternative programme choices. Applying the Lechner (2001) framework, usually carried out using the Sianesi (2004) stata software, one can estimate the effect of one option as compared to one or several other options. In practice, there are two informative estimates which can be achieved:

1   an estimate of the average programme effect in one particular option compared to non-participation in any of the programmes; or

2   a comparison of participating in one particular option as compared to another option ('Pairwise comparison').

For multiple participation due to repetition of programme participation or participation dynamics, we additionally recommend discussion and estimation of the scale and timing of multiple participation observed. This will provide descriptive evidence of how important these issues might be.

In the absence of applying more recently developed techniques which account for multiple participation, where there is substantial multiple participation and the characteristics of those with second participation differ statistically (determined with t-tests or appropriate testing) from those with first participation only, then the impact estimate analyses should be constrained to the first participation. For the CBA, this means that the total additional time generated is for those who are on their first participation. If the characteristics are statistically similar, then proceed with caution to use the overall impact, clearly reporting the scale of multiple participation, and discussing the potential for bias.

1 Introduction

The Department for Work and Pensions (DWP) commissioned this study in 2008 with the primary aim of informing cost benefit estimates that are derived following the Cost-Benefit Framework (CBF). Since then, the department has produced a Social Cost Benefit Analysis (CBA) Framework that complements the CBF (see, Fujiwara, 2010). The social framework discusses evidence relating to wider impacts of employment and training programmes. It also makes recommendations for how some specific impacts should be quantified and monetised.

The secondary aim of this project was to explore the feasibility and value of making further improvements to estimates of net benefits across New Deal employment programmes. This research builds on the external review of the DWP CBF carried out by Greenberg and Knight in 2007 (Working Paper No. 40). Further detail of the extent of the research is now outlined.

1.1 Operational background and context of the research

DWP runs a range of employment programmes. The department has regularly undertaken comparative cost benefit analyses (CBAs) of large numbers of these programmes. These exercises estimate the absolute and relative cost effectiveness of the various interventions. The department also frequently undertakes CBAs for individual employment programmes at the appraisal and evaluation stages of the policy cycle.

To aid comparability and consistency in the assessment of its policies and programmes, the DWP has developed a CBF. This framework describes the steps that should be taken and the assumptions that should be made when conducting a CBA of an employment programme. The CBF ultimately attempts to ensure that public funds are spent efficiently and on programmes that generate the greatest net benefits to society.

At the inception of this project in late 2008, CBAs produced following the CBF occasionally utilised estimates of earnings in work and the duration of additional jobs (the latter occasionally being a necessary input for a CBA) based on somewhat ad hoc assumptions. This was due to a lack of resources required to obtain more robust estimates. Similarly, many of the possible wider benefits and costs of DWP employment programmes such as improved health outcomes, reductions in crime, and general equilibrium effects were not quantified and monetised in internal CBAs due to a lack of reliable evidence. These wider impacts are now incorporated in internal CBAs drawing upon evidence collected in this report as well as other sources (Fujiwara, 2010).

This report has helped inform the Social CBA Framework. It should also enable DWP to make further improvements to the reliability and usefulness of their CBA estimates by, for example, taking forward some of the recommendations from the external review of the CBF by Greenberg and Knight (2007). It uses some recent DWP datasets such as the Work and Pensions Longitudinal Survey and existing methodologies to explore opportunities to improve CBA practice and the quality of internal estimates.

The research aims to provide recommendations with regard to the evidence gaps and seeks to inform with regard to key questions in using CBF to guide analyses.

1.2   Scope of the research

The research consisted of two components: a literature review and new empirical estimates from DWP administrative data.

We considered the following areas in the literature review:

- general equilibrium effects;
- subgroup impacts and distribution of impacts;
- duration of benefits and employment and wages;
- multiple participation in programmes or other interventions.

The analysis of the administrative data provides:

- empirical estimates of gross duration of benefits and employment, and annual nominal taxable earnings.

These estimates have been produced for the range of New Deal programmes. It is emphasised that these are not net impact estimates.

1.3   Review research methods

We reviewed DWP documents passed by the project manager, which included the 2008 CBF guidance, 2008 versions of the 2005/06 employment programme CBAs, any later CBA results available, internal documents reporting on wider effects, and cost-effectiveness analyses. We note that the project was substantially delayed by data access issues, which led to a relatively long timeframe for the project. During this time, the materials such as the CBF guidance have altered. We have tried to acknowledge this.

The focus of the literature review was determined with reference to the research brief set out by DWP and through meetings with the DWP project leader. A prioritisation was made in key areas where development was preferred by DWP, and revisions occurred to the project focus in subsequent teleconference review meetings or by agreement in email communications. It is noted that the key areas for development were defined as those where an improvement would make a significant difference to the estimates reported, and where such a change could be obtained at relatively low cost. The literature review research provides a critical assessment, as well as an examination of the general cost-benefit literature on specific points.

1.4   Outline of this report

The general equilibrium literature review is contained in Chapter 2. Following this, Chapter 3 contains a discussion of subgroup impacts, including subgroups of the outcome distribution which are estimated with quantile regression. Then in Chapter 4, the methods for estimating net impacts on duration of benefits and employment and wages are considered. This is followed by empirical estimates for duration of benefits and employment, and earnings levels in Chapter 5. These estimates are limited to gross estimates for New Deal programmes. Multiple participation and how to account for it is covered in Chapter 6. This is done by examining the available data and what further simple estimates can provide information on multiple participation. In addition, we assess that evaluation study designs could be altered to ensure that this required information is collected in future, even if more rigorous impacts that account for multiple participation are not estimated.

The references are split into two reference lists so that key sections are easily found. In particular the references for the General Equilibrium review in Chapter 2 are separated, and follow the general references.

There are four appendices. The first, Appendix A, contains tables of the empirical gross estimates for New Deal programmes from Chapter 5. Appendix B has detailed data information related to Chapter 5. Appendix C contains more detailed discussion of the econometric methods for producing quantile treatment effects for subgroups of the outcome variable. Appendix D has descriptive information on the available DWP datasets, as sourced from DWP in 2008.

2   General equilibrium analyses

This chapter examines the literature for general equilibrium analyses and assesses how best to use this evidence to inform Department for Work and Pensions cost benefit analysis (CBA).

2.1   Introduction

The employment and training programmes that are assessed using the Cost Benefit Framework (CBF) guidance may have effects on the well-being of those who are not enrolled in the programme and, because of this, on the economy. Such effects include substitution effects, displacement effects, equilibrium wage effects, multiplier effects, and entry effects. The current CBF contains recommendations for treating substitution effects as part of sensitivity tests, but does not contain recommendations concerning the remainder of these effects.

This chapter describes what is known about these so-called 'general equilibrium effects'. It discusses when they are likely to be important and when they are not, and then examines the available empirical evidence. Most importantly, we inform on plausible ranges for substitution effects (note that these are confusingly called displacement effects in the United States (US) literature). To investigate a plausible range, we have conducted a limited literature review of sizes found to date. As will be seen, however, empirical evidence about the magnitudes of substitution effects and other general equilibrium effects is quite limited.

Estimation of general equilibrium effects is **not** performed within this project, as the returns to estimation are likely to be low relative to the time and resources required. We also do **not** consider monopoly rents – no existing research has been found that estimates monopoly rents and so selection of a more appropriate measure is considered outside the remit of this project.

We support the conclusion of Greenberg and Knight (2007) that, except in those rare instances when a reliable estimate exists, the best course of action is simply acknowledging the possible existence of general equilibrium effects and discussing whether they are likely to be important. The information and tables contained in this review can inform this discussion. In some cases, they can also be used to guide sensitivity tests to support a conclusion about the likely importance of these effects.

2.2   Definitions of types of general equilibrium effects

Before discussing the empirical evidence, we first define these various effects[1] and then examine some of the factors that influence them.

2.2.1   Substitution effects

By increasing job skills or increasing the number of persons seeking employment or their intensity of job search, employment and training programme may increase competition for available jobs. Hence, programme participants may end up in jobs that would otherwise have been held by non-

---

[1]   Different authors use the terms discussed below to mean different things. For example, US economists usually refer to what we are calling 'substitution effects' as 'displacement' effects. We attempt to use these terms consistently in accordance with the definitions provided in the text of this report.

participants (Johnson, 1972; Schiller, 1973).[2] Programme participants are, in effect, substituted for non-participants. If workers who are substituted against become unemployed or accept lower-wage jobs, their earnings are obviously less than they otherwise would be. The benefits of employment and training programmes on society as a whole will therefore be less than the benefits received by programme participants.

2.2.2   Fiscal substitution effects

A somewhat different sort of substitution effect, but with similar implications for policy, is sometimes called a 'fiscal substitution' effect (Johnson and Tomola, 1977). This effect can occur in the case of programmes (which are often called 'wage subsidy programmes') that pay subsidies to private sector employers that hire members of specific disadvantaged target groups and programmes (which are sometimes referred to as 'public sector employment') that directly place targeted disadvantaged individuals into jobs at government agencies or non-profit institutions. In this case, the targeted workers may be hired instead of, or even replace, those who are not targeted and, hence, more expensive to employ. It is also possible that some of the same persons who employers would have employed in the absence of a wage subsidy or public sector employment programme are employed under the programme and, thus, essentially replace themselves. In this instance, the employer receives a windfall, but substitution does not occur because there is no effect on non-participants in the programme. There is instead what Calmfors, Forslund, and Hemstrom (2002) term a 'deadweight effect'. Deadweight effects are especially likely if the programme is not tightly targeted on a particular disadvantaged group.

2.2.3   Displacement effects

These effects occur if some firms expand at the expense of others because their employment costs are subsidised by the government.

2.2.4   Equilibrium wage effects

If participants in an employment and training programme search harder for jobs or their job skills increase and, as a result, their weeks or hours are greater than they otherwise would have been, then the resulting increase in labour supply will tend to put downward pressure on the equilibrium wage within the labour markets in which they work. If the programme tends to reduce job vacancies – for example, by making job search more efficient or imparting job skills that are in demand – this will also increase downward wage pressures. Thus, workers who are employed in the same labour markets as programme participants could receive lower wages than otherwise would be the case. Notice that if wage rates are lower than they would be without a programme, then this will tend to mitigate substitution effects resulting from the programme.

2.2.5   Multiplier effects

In the absence of full-employment, expenditures on employment and training programmes can potentially stimulate the economy through so-called multiplier effects. That is, as individuals are hired to provide training or equipment is purchased to operate the programmes, the recipients of these expenditures will, in turn, spend their newly gained revenues, and this, in turn, will generate still more revenues that will be spent, and so forth. Each round of additional expenditures will, of course, result in increased hiring. It is important to emphasise that these multiplier effects are only germane when unemployment is substantial – that is, at least five or six per cent. Otherwise, programme expenditures will tend mainly to bid up prices and wages.

---

[2]   Indeed, as pointed out by Davidson and Woodbury (1993), because it is now more difficult for unemployed non-participants to find jobs, they may search for jobs less diligently.

2.2.6   Entry and deterrent effects

If employment and training programmes are perceived as beneficial, but are available only to benefit recipients, some persons may take up benefits in order to qualify. On the other hand, in the case of mandatory programmes for benefit recipients, some individuals who might otherwise have taken up benefits may decide not to do so in order to avoid the 'hassle' of participating in the programme. The latter negative effect on entry is sometimes known as a 'deterrent effect'. Manski and Garfinkel (1992) and Moffitt (1992, 1996), among others, argue that both positive and negative programme 'entry effects' such as these could be substantial.

Entry and deterrent effects are not equilibrium effects in the same way as substitution and displacement effects and equilibrium wage effects are because they do not work through labour markets. They instead affect enrolment in benefit programmes. As discussed in the next section, positive entry effects are only likely to occur in the long-run and, hence, are likely to be missed in evaluations and CBA of new training and employment programmes. Negative entry effects (i.e. deterrent effects) can occur in the short-run and, by definition, keep individuals from enrolling in benefit programmes in order to avoid participating in mandatory employment and training programmes. Thus, deterrent effects are a general equilibrium effect of employment and training programmes in the sense that they affect the well-being of individuals who are not enrolled in these programmes.

The importance of the equilibrium effects defined above and exactly how they manifest themselves depends on the specific characteristics of the employment and training programmes under consideration, the state of the relevant labour markets, and a number of other factors. These are considered next. Following that, we review empirical attempts to estimate the size of the effects.

2.3   Factors influencing equilibrium effects[3]

As just suggested, different types of employment and training programmes have different equilibrium effects. For purposes of discussion, we distinguish between 'supply-side programmes' and 'demand-side programmes'. The former provides job search assistance, training, and financial incentives that encourage individuals to seek employment and helps them to enter employment. The latter subsidises employers for employing certain categories of disadvantaged workers or places such workers directly into jobs, often within the public and non-profit sectors. As noted above, we refer to such programmes as 'wage subsidies programmes' and 'public sector employment', respectively.

2.3.1   Substitution effects

Regular substitution effects, as opposed to fiscal substitution effects, result from supply-side employment and training programmes. If supply-side programmes cause programme participants to search for jobs more diligently or more effectively or increases their job skills, then some of their increases in employment may come at the expense of non-participants.

The importance of substitution effects partially depends on the number of existing job vacancies. The fewer the number of job vacancies, the more difficult it is for unemployed non-participants in supply-side programmes to find jobs that are alternatives to those taken by participants. Thus, the magnitude of substitution effect is likely to reflect the state of the relevant local labour markets.

---

[3]   Useful theoretical frameworks for examining the general equilibrium effects of employment and training programmes can be found in Organisation for Economic Co-operation and Development (OECD) (1993); Calmfors (1994); and Calmfors, Forslund, and Hemstrom (2002).

If a local labour market in which former programme participants work is tight (i.e. the ratio of vacancies to job seekers is high), then alternative job opportunities are likely to be available to those outside the programme target group; but if it is loose, then the cost of substitution to those affected could be substantial. However, even if there is a substantial substitution effect, it is unlikely to be permanent. If the economy is expanding, substitution effects should diminish over time, as job opportunities become open and absorb those who were substituted against. Thus, the substitution effect is likely to be more important in the short-run than in the long-run.

Several arguments have been put forward to suggest that substitution may not seriously undermine the effectiveness of supply-side programmes. First, if macroeconomic policy can keep aggregate unemployment rates low, then it should be relatively easy for displaced ineligibles to find alternative job opportunities. Second, as Cohen (1969) and Johnson (1979) suggest, if programme participants are less likely to seek employment while in the programme (for example, if they are receiving training they consider valuable) than they otherwise would have been (a so-called 'locking-in effect', then this opens more jobs to ineligibles, at least temporarily. Third, if they improve job matching or increase the work-skills of participants, supply-side programmes should make it less expensive for employers to fill positions. As a result, employers may be willing to hire more workers at a given wage. Fourth, if wage rates decline as a consequence of increased competition in job markets that is attributable to supply-side programmes, then this will tend to increase employment and offset substitution effects. Fifth, as emphasised by Johnson (1979) and Katz (1994), if training programmes can impart skills that allow trainees to leave slack occupational labour markets for tight ones, they can decrease the competition for job vacancies in the slack markets, thereby making it easier for ineligibles in these markets to find jobs. Such a possibility could produce a result that is the exact opposite of a substitution effect – total employment could increase by more than the number of persons who are trained.[4]

2.3.2   Fiscal substitution effects

Fiscal substitution effects result from demand-side programmes[5]. Much of what was said above about substitution effects that result from supply-side programmes is also applicable to fiscal substitution effects. For example, locking-in effects, a relatively large number of job vacancies and expansions in the economy are all likely to reduce the size of fiscal substitution effects. Demand-side programmes differ from supply-side programmes, however, because they are not primarily intended to improve job matching or job skills, and this may cause the substitution effects resulting from these programmes to be more severe. However, to the extent that a wage subsidy or a public sector employment programme focuses on placing low-skilled workers into jobs, it will change the composition of the non-employed in the direction of increased skill levels. Thus, those who remain in the non-employed pool may have less difficulty in ultimately finding jobs on their own, than would programme participants.

---

[4]   Johnson (1979) suggests that whether this occurs depends on wages adjusting sluggishly in the slack labour market and trainees being readily absorbed into the tight market. Furthermore, as Johnson notes, it will occur less to the extent that the programme induces non-workers to enter the workforce or merely moves workers from one slack labour market to another. In addition, as Hamermesh (1972) and Hall (1979) point out, the training must be for jobs for which vacancies exists.

[5]   Much of the literature on demand-side programmes is concerned with temporary programmes aimed principally at reducing cyclical unemployment (for example, see Kopits, 1978; Johnson and Tomola, 1977; Knabe, 2008, and the references therein). Our major concern here, in contrast, is with permanent programmes aimed at increasing the employment of certain categories of disadvantaged workers, such as lone parents or the long-termed unemployed. Both types of demand-side programmes can result in fiscal substitution effects, however.

In contrast to supply-side programmes, demand-side programmes may generate substitution effects that can grow over time, especially if the programmes are viewed by employers as relatively permanent. The reason is that participants in these programmes are less expensive to employ than non-participants, but it may take employers time to make the adjustments necessary to replace regular employees with programme participants. This may be especially likely if, as is often the case, demand-side programmes stipulate that the programme participants who are hired perform work that will not otherwise be performed. If employers attempt to circumvent such constraints, it will take them some time to do so.

2.3.3   Displacement effects

Displacement effects are only likely to take place if a programme gives some employers a competitive advantage over other employers. Thus, in the case of employment and training programmes, the possibility of displacement effects are mainly restricted to demand-side programmes that pay subsidies to private sector employers to hire additional workers or hire certain categories of workers. Displacement occurs if employers that expand their employment as a result of the subsidy do so at the expense of employers at competing firms who find it necessary to contract their employment[6].

2.3.4   Equilibrium wage effects

Demand-side employment and training programmes could increase employer demand for workers and, hence, result in some upward pressure on wages. This might occur because the programmes make at least certain categories of workers less expensive to hire or because substitution is incomplete, perhaps due to programme rules, and, as a result, more jobs and job vacancies result.

Supply-side programmes, in contrast, tend to increase competition in labour markets, thereby resulting in downward pressure on wages. Potentially, this would tend to mitigate substitution effects. However, if they improve job matching or increase the work-skills of participants, supply-side programmes should make it less expensive for employers to fill positions. As a consequence, employer demand for workers will increase, at least partially offsetting downward pressure on wages that result from increased competition in the labour market. For any downward pressure on wages to actually result in wage rates that are much lower than they otherwise would be, however, three conditions must hold: (1) the minimum wage must not constrain downward movements in wage rates; (2) programme participants must account for a fairly large share of the workers in the relevant labour markets; and (3) programme effects on job search and employment must be fairly large.

Most supply-side programmes seem unlikely to bring about large equilibrium wage effects. Most participants who are employed because of the programmes tend to work in low-wage labour markets. Thus, at least to some degree, the minimum wage probably constrains reductions in equilibrium wages. Moreover, the programme target groups tend to be narrow (e.g. the disabled and persons with serious health problems, the long-term unemployed, lone parents, etc.). Thus, members of these groups usually account for only a fairly small proportion of the total supply population in any given labour market. Finally, programme impacts are typically fairly modest.

Thus, most supply-side programmes seem unlikely to bring about large equilibrium wage effects.

---

[6]   Knabe (2008) presents a theoretical comparison of displacement among different forms of wage subsidies that are intended to expand employment.

2.3.5   Multiplier effects

In general, multiplier effects are probably best ignored in cost-benefit studies of employment and training programmes unless it can be argued that, in the absence of these programmes, there would be no multiplier effects. For this to occur the expenditures that would have been made on the programmes would have to be used to reduce taxes and these tax reductions would then have to be saved by those who receive them. Or, alternatively, the reduction in expenditures could be used to reduce the government's deficit. If the revenues that would have been expended on the programmes are instead used for other purchases, either by the government or private individuals, the multiplier effects would continue to exist, although they could, perhaps, be somewhat larger or smaller than those generated by the employment and training programmes.

2.3.6   Entry and deterrent effects

To qualify for many employment and training programmes, individuals must often be on benefits of some kind. Because of this link between benefit programmes and employment and training programmes, entry onto benefits may be influenced either positively or negatively, depending upon whether participation in the programme is a mandatory requirement for receiving benefits or is instead voluntary.[7] A mandatory programme can have both positive and negative entry effects. On the one hand, if an employment programme appears attractive to individuals – for example, in terms of increasing their earnings potential – then they may do whatever is necessary to qualify for benefits (even leaving a low-wage job, or extending their time on benefits to qualify if there is a specified period of benefit receipt for eligibility) so that they can participate in the programme, a positive entry effect. On the other hand, if the programme appears unattractive – for example, it appears likely to take up one's time with little offsetting rewards – then individuals may be deterred from taking up benefits even if they otherwise qualify, a negative entry effect. In contrast to mandatory employment and training programmes, voluntary programmes should only result in positive entry effects onto benefits because whatever is deemed onerous about the training programme can be avoided by simply deciding not to participate in it.

In the case of a new employment and training programme, there may be little in the way of positive entry effects, because the payoff from participating in the programme is likely to be unknown. However, deterrent effects can occur early on in the case of a mandatory programme because the time requirement of participating in the programme and the tasks that would have to be performed would be readily apparent.

Entry and deterrent effects have obvious implications for how employment and training programmes affect caseload size. They also have implications for assessing the effects of new employment and training programmes on employment and earnings because these effects are usually measured by studying the existing pool of those on benefits; but entry and deterrent effects affect the composition of those who receive benefits over time and, therefore, may cause programme effects on employment and earnings to change over time. These changes will usually not be observed. In addition, deterrent effects may bias the estimates of employment and earnings effects in the case of both new and long existing employment and training programme because the programme's effects on those who are deterred from taking up benefits cannot be observed.

As just suggested, standard programme evaluation methods usually fail to capture entry and deterrent effects. They usually must be assessed by separate studies. The approaches that have been used in such studies are described later.

---

[7]   For a detailed analysis, see Moffitt, 1996.

2.4   Review of empirical evidence on equilibrium effects

2.4.1   Substitution effects

With only a few exceptions, empirical studies of substitution effects have used one of the following three approaches: (1) surveys in which employers or programme participants are asked whether the work performed by programme participants would have been performed in the absence of the programme; (2) econometric macroeconomic evaluations that estimate relationships between aggregate variables such as employment or unemployment and various measures of programme size such as programme expenditures or the number of participants; and (3) using general equilibrium models to predict programme effects. As discussed below, all three approaches are subject to important limitations.

The survey approach has been almost entirely limited to attempting to estimate fiscal substitution effects resulting from demand-side programmes – that is, public sector employment and wage subsidy programmes. In principle, the survey approach should capture deadweight effects, as well as substitution effects, although it usually would be difficult to distinguish between the two. This approach is obviously limited because neither employers nor programme participants may have a very accurate idea of whether the work that participants perform would have been performed if the programme did not exist, a hypothetical situation. Moreover, if there are restrictions on using programme participants to perform regular work, then employers may not want to admit that violations of these restrictions are occurring.

The macroeconomic approach attempts to exploit the fact that the size of employment and training programmes varies geographically and over time and thus programme effects on employment and unemployment should also vary. To illustrate with one of a number of alternative regression specifications that has been used, the number of employed non-participants can be regressed against a time lagged measure of the number of employed former programme participants and a set of control variables. A coefficient of -.5 would imply that for every 100 programme participants that obtain jobs as a result of the programme, 50 non-participants would be non-employed due to the programme. Similarly, a coefficient of -1 would imply full substitution so that no increase in employment results from the programme. In principle, the estimates made with this approach for wage subsidy programmes should incorporate displacement effects, as well as substitution effects. In most cases, it would be difficult to determine the individual contributions of these two effects. The major problem with the macroeconomic approach is reverse causality, which is recognised by virtually all the analysts who have used the approach. That is, although employment and unemployment may be affected by programme participation, the opposite may also be true. Government investments in employment and training programmes are likely to be larger in places and at times when unemployment is relatively high and employment is relatively low than when the opposite conditions exist. Various instrumental variables and lag structures have been used to try to mitigate the reverse causality problem, but it is not clear how successful these attempts are.

General equilibrium modelling has been used relatively infrequently because the effort required to build such models is considerable, and they are specific to a given programme of interest and not readily adapted to analysing other programmes. Both direct and indirect programme effects must be formally modelled and values must be assigned to each parameter in the model. To the extent possible, these values are taken from existing information and previous research. If they do not exist, the model builders can estimate them or make plausible assumptions about them. When uncertainty exists about some of the parameters, sensitivity tests are usually made by using alternative values. In addition to concern over whether the correct parameter values are used, a potential drawback of general equilibrium modelling is that key relationships may be overlooked or incorrectly modelled. Unfortunately, there is rarely any way to verify the accuracy of the model's predictions. Nonetheless, in examining substitution effects, they may be suggestive about likely orders of magnitude.

Findings from various studies of substitution effects that rely on the survey, econometric estimation, and general equilibrium modelling approaches are summarised in Tables 2.1, 2.2, and 2.3, respectively. The survey approach has been used only to estimate fiscal substitution effects for demand-side programmes and general equilibrium models have been limited to analyses of supply-side programmes. Econometric estimates of substitution effects have been made for both demand- and supply-side programmes. Thus, the studies in Table 2.2 have been grouped separately by these two types of programmes. A few econometric studies have utilised micro-data on individuals or firms, but most have relied on macro-data.

Table 2.1   Estimates of fiscal substitution effects for demand-side programmes from surveys

Country	Study	Type of programme	Size of substitution effect[1] %	Source[4]
Sweden	LO (1993 and 1994)	PSE	20-39	CFH
Sweden	Temo (1993-1995)	PSE	3-17	CFH
Sweden	NUTEK (1994)	PSE	30	CFH
Sweden	Hallstrom (1995)	PSE	20-25	CFH
Sweden	AMS (1997)	PSE	8-24	CFH
Sweden	AMS (1998a)	PSE	13-27	CFH
Sweden	AMS (1998b)	PSE	0-14	CFH
Sweden	Johansson (1999)	PSE	11-26	CFH
Netherlands	DeMunnik (1992)	PSE	15	OECD
US	National Commission for Manpower Policy (1978)	PSE	18	Authors
Sweden	Sehlstedt and Shroder (1989)	Wage subsidy	23 to 49	CFH
Sweden	AMS (1995)	Wage subsidy	36	CFH
Sweden	Anxo and Dahlin (1996)	Wage subsidy	69-84	CFH
Sweden	AMS (1997)	Wage subsidy	42-48	CFH
Sweden	AMS (1998a)	Wage subsidy	40-51	CFH
Sweden	AMS (1998b)	Wage subsidy	32-35	CFH
United Kingdom	Atkinson and Meager (1994)	Wage subsidy	Over 80	JVG
Netherlands	DeKoning et al. (1995)	Wage subsidy	Over 85	JVG
Netherlands	DeKoning et al. (1995)	Wage subsidy	Over 90	JVG
Belgium	Van der Linden (1995)	Wage subsidy	Over 90	JVG
France	Belleville (2001)	Wage subsidy	69-80	Authors
Ireland	Breen and Halpin (1989)	Wage subsidy	95[2]	OECD
Australia	Department of Employment, Education and Training (1989)	Wage subsidy	67-79[3]	OECD

Notes: PSE = Public Sector Employment programmes.

[1]   The estimated effect may include deadweight, as well as substitution effects.

[2]   A decomposition of the effects indicate the following: deadweight – 70 per cent; substitution – 21 per cent; displacement – four per cent.

[3]   Deadweight only.

[4]   This column indicates whether the estimates in the third column were obtained from Calmfors, Forslund, and Hemstrom, 2002 (CFH), from Jongen, Van Gameren, and Graafland 2003 (JVG), from the July 1993 issue of *Employment Outlook* published by OECD, or directly from the study listed in the second column (authors).

Table 2.2   Estimates of substitution effects from econometric studies

Country	Study	Type of programme[2]	Size of substitution effect[1] %	Source[3]
Demand-side programmes				
Sweden	Gramlich and Ysander (1981)	PSE	Road construction: 100%	
			Health and welfare: 0%	CFH
Sweden	Forslund (1996)	PSE	Work experience: 0%	
			Relief work: 84%	
			Youth programmes: 76%	CFH
Sweden	Forslund and Krueger (1997)	PSE	Construction: 69%	
			Health and welfare: 0%	CFH
Sweden	Lofgren and Wikstrom (1997)	PSE	Work experience: 0%	
			Relief work: 0%	
			Youth programmes: 94%	CFH
Sweden	Dahlberg and Forslund (1999)	PSE	66%	CFH
Sweden	Edin, Forslund and Holmlund (1999)	PSE	Youth programmes: 76%	CFH
Sweden	Dahlberg and Forslund (2005)	PSE	65%	Authors
Sweden	Holmlund (1995)	PSE	Youth: 80%	P
Sweden	Skedinger (1995)	PSE	Youth: 40-110%	CFH
Poland	Puhani (1999, 2002)	PSE	Inconclusive findings	Authors
US	Johnson and Tomola (1977)	PSE	Small initially, but about 100% after 1st year	Authors
US	National Planning Association (1974)	PSE	46%	JT
US	Wiseman (1976)	PSE	39%	JT
US	Farkas, Smith, and Stronmsdorfer (1983)[1]	PSE	Youth: 26%	Authors
Sweden	Lofgren and Wikstrom (1997)	Wage subsidy	0%	CFH
Sweden	Dahlberg and Forslund (1999)	Wage subsidy	65%	CFH
Finland	Kangashaju and Venctoklis (2003)	Wage subsidy	Over 100%	K
Denmark	Rotger and Arendt (2010)[1]	Wage subsidy	0%	Authors
Netherlands	DeKoning *et al.* (1992)	Wage subsidy	69-76%	OECD
Poland	Puhani (1999, 2002)	Wage subsidy	Inconclusive findings	Authors
US	Perloff and Wachter (1979)[1]	Wage subsidy	67%	Authors
US	Farkas, Smith, and Stronmsdorfer (1983)[1]	Wage subsidy	Youth: 47%	Authors
Sweden	Calmfors and Skedinger (1995)	Mixed wage subsidy and PSE	Possibly substantial effect	Authors

Table 2.2   Continued

Country	Study	Type of programme[2]	Size of substitution effect	Source[3]
Supply-side programmes				
Sweden	Forslund (1996)	Training	0%	CFH
Sweden	Lofgren and Wikstrom (1997)	Training	0%	CFH
Sweden	Dahlberg and Forslund (1999)	Training	0%	CFH
Sweden	Dahlberg and Forslund (2005)	Training	11%	Authors
Sweden	Calmfors and Skedinger (1995)	Training	Inconclusive but probably a small effect	Authors
E. Germany	Hagen and Steiner (2001)	Training	65%	Authors
UK	Adam *et al.* (2008)[1]	Job counselling and financial incentives	Findings consistent with the possibility of small short-run effect in some geographic areas	Authors
Mixed demand- and supply-side programmes				
Finland	Rantala (1995)	Training and PSE	Youth: very small effect	P
Finland	Pehkonen (1997)	Training and PSE	Youth: possibly substantial effect	Authors
Finland	Ericksson and Pehkonen (1998)	Training and PSE	Possibly substantial	Authors
UK	Blundell *et al.* (2004)	Job search and wage subsidy	Probably small or none	Authors

Notes:

[1] The Perloff and Wachter (1979); Farkas, Smith, and Stronmsdorfer (1983); Blundell *et al.* (2004); and Adam *et al.* (2008) studies are based on micro data on individuals. The Rotger and Arendt (2010) study is based on micro data on firms. The remaining studies are all based on aggregated macro-data.

[2] For studies based on macro-data, estimates for wage subsidy programmes may incorporate displacement effects, as well as substitution effects.

[3] This column indicates whether the estimates in the fourth column were obtained from Calmfors, Forslund, and Hemstrom, 2002 (CFH), from Pehkonen, 1997 (P), from Johnson and Tomola, 1977 (JT), the July 1993 issue of *Employment Outlook* published by OECD, or directly from the study listed in the second column (authors).

Table 2.3   Estimates of substitution effects for supply-side programmes from general equilibrium models

Country	Study	Type of programme	Size of substitution effect	Source[1]
US	Davidson and Woodbury (1993)	Financial incentives	30-60%	Authors
Canada	Lise, Seitz, and Smith (2004)	Financial incentives	Very large	Authors
Australia	Dixon and Rimmer	Financial incentive	60% in short-run, but diminishing over time	Authors
Australia	Dixon and Rimmer	Job search	70% in short-run, but diminishing over time	Authors

[1]   The estimates in the fourth column were all obtained directly from the studies listed in the second column.

Whenever possible, the magnitude of the substitution effect has been shown as the ratio of the number of jobs lost by non-participants due to a programme's substitution effect to the number of jobs gained by programme participants. Thus, a substitution effect of 30 per cent would imply that for every 100 programme participants that obtain jobs as a result of the programme, 30 non-participants would be non-employed due to the programme.[8] Precise estimates of substitution effect ratios cannot be obtained from some of the studies listed in Tables 2.2 and 2.3, either because the estimates are in a form that is not readily converted into a ratio or because the findings are of low statistical significance or are not very robust to alternative regression specifications. In these instances, orders of magnitude suggested by those conducting the studies are indicated.

As is evident from the tables, many of the findings for substitution effects are for Sweden, a country that, in comparison to other countries, has traditionally made relatively large expenditures in employment and training programmes as a percentage of GDP (Calmfors, Forslund, and Hemstrom, 2002). Thus, Swedish policy makers have particular reason to be concerned with substitution effects and they may be more readily detected in that country.

Only three of the studies listed in the tables pertain to the UK. Although two of these studies use an econometric approach to examine substitution effects, and therefore appear in Table 2.2, unlike most of the other studies in that table, they are based on data on individuals, rather than on macro-data.[9] The Blundell *et al.* (2004) study used a difference-in-differences methodology to determine whether the New Deal for Young People programme had an impact on the employment status of persons who were not eligible to participate in the programme. Similarly, the Adams *et al.* (2008) study used the difference-in-differences methodology to determine whether the Pathways to Work programme had an impact on the benefit status of persons who were ineligible to participate.

Tables 2.1-2.3 indicate that there is a great deal of variation in the estimates of substitution effects.

---

[8]   In principal, the size of the employment and training programme should not have very much of an effect on this ratio. For example, a wage subsidy programme could have a substitution effect of (say) 30 per cent regardless of whether it's funding accounts for a large or small share of gross domestic product (GDP). However, econometric studies that rely on macro-data (i.e. most of those summarised in Table 2.2) are less likely to be able to detect the substitution effects resulting from a small programme than a large programme.

[9]   The Farkas, Smith, and Stromsdorfer (1983) and Perloff and Wachter (1979) studies are also based on micro-data.

This variation is not surprising given that the estimates pertain to different time periods and geographic locations. Moreover, even within a given table, the methods used to obtain the estimates differ greatly. Furthermore, even when programmes are classified as being of the same type – for example, as a training programme or as a wage subsidy or public sector employment programme – they may differ in important respects. For example, one training programme may offer remedial education, another vocational education, and a third both.

Even given the variation among the estimates, however, a pattern is discernible. The estimates imply that demand-side programmes tend to result in much larger substitution effects than supply-side programmes. As can be seen in Table 2.2, this is true even when the same study uses the same methodology in estimating the substitution effects associated with both types of programmes.

Examining the estimates produced by each methodological approach, it can be seen that those relying on the survey approach (Table 2.1) suggest that wage subsidy programmes tend to cause substantial substitution effects (which are often over 60 per cent), while the substitution effects resulting from public sector employment programmes are more moderate (that is, 30 per cent or less). Estimates based on the econometric approach (Table 2.2) seem to indicate that supply-side programmes probably do not cause very large substitution effects, but demand-side programmes, regardless of whether they are wage subsidy or public sector employment programmes, probably do, especially if they are targeted at youth. More specifically, of the 30 estimates appearing in Table 2.2 for demand-side programmes, three are inconclusive, seven indicate that there is no substitution effect, and the remaining 20 estimates all suggest that the effect is large (typically over 50 per cent). In contrast, all but one of the seven estimates for supply-side programmes imply that the substitution effect resulting from such programmes is small, or at worst no more than moderate (that is, under 20 per cent). Contrary to the econometric estimates for supply-side programmes, estimates that rely on the general equilibrium models suggest that substitution effects for this type of programme are quite large (that is, over 50 per cent).[10] The reasons for this difference are not apparent.

What are the lessons from these findings for cost benefit analyses (CBAs) of employment and training programmes? It seems clear that the findings must be treated with great caution in using them in CBA. They vary greatly and the methodologies that underlie them are all subject to important weaknesses. Nevertheless, there appears to be considerable evidence that substitution effects are likely to be quite important when demand-side programmes are considered, suggesting that any CBAs of such programmes should be subject to sensitivity tests in which fairly large substitution effects are assumed. Results from this sensitivity test would then be compared to findings from the standard analysis in which no substitution effect is assumed.

We suggest that a value towards the higher end of the scale be used as a plausible upper bound for this purpose. It may also be useful to conduct a second sensitivity analysis with a lower bound

---

[10]   Bartik (2002) developed several models that are more or less in keeping with the general equilibrium approach to investigate the US welfare reform efforts around the turn of the century. These efforts resulted in substantial increases in the labour supply of lone parents. His models predict that the short-run substitution effects would be substantial, but fall over time. Jongen, Van Gameren and Graafland (2003) used a general equilibrium model to examine substitution and deadweight effects resulting from highly 'stylised' public sector employment, wage subsidy, and training programmes in the Netherlands, finding that these effects were quite large, especially for the first two programmes. Because these results are not based on actual programmes, they are not listed in Table 5.3.

estimate in order to bracket the likely importance of substitution effects.[11] For example, Table 2.2 reports seven econometric estimates of the substitution effects resulting from wage subsidy programmes. Two of these estimates imply a zero substitution effect; one indicates that it is 47 per cent, and the remaining four estimates imply that it is at least 65 per cent.[12] Similarly, five of the 13 survey-based estimates in Table 2.1 imply that the substitution effects resulting from wage subsidy programmes are in the 30 to 50 per cent range, while the remaining seven estimates imply that it is above 70 per cent. Thus, 60 per cent, or even a larger value, might be used as an upper bound value for the substitution effect in conducting sensitivity analyses of wage subsidy programmes and 30 per cent might be used as a lower-bound value. Based on the same evidence as that reviewed here, Fujiwara (2010) suggests conducting a single sensitivity test by using the mid-point between these two values, 45 per cent.

A similar approach can be used to obtain a substitution effect value that can be used in conducting sensitivity analyses of public sector employment programmes. Unfortunately, however, unlike the wage subsidy situation, substitution effect estimates based on the econometric and survey approaches diverge considerably. For example, there are 19 econometric estimates for public sector employment programmes in Table 2.2. Five of these estimates imply that there are no substitution effects; three imply that it is less than 50 per cent; and the remaining 11 estimates imply that it is at least 65 per cent. In contrast, the nine estimates relying on the survey approach that are reported in Table 2.1 imply that substitution effects resulting from public sector employment programmes are in the ten per cent to 30 per cent range. As discussed earlier, the econometric and survey approaches are each subject to important limitations. Consequently, it is difficult to determine which provides superior estimates of the substitution effect. Thus, we suggest that 60 per cent be used as an upper bound and 30 per cent as a lower bound for sensitivity analysis purposes in CBA of public sector employment programmes. Fujiwara (2010) again recommends using the mid-point of 45 per cent of these two bounds. We further suggest that in reporting the results of these sensitivity tests, it be stressed that the upper bound results are most consistent with substitution effect estimates based on the econometric approach and the lower bound findings reflect substitution effect estimates that rely on the survey approach.

If the economy is expanding, either 30 per cent or 60 per cent should probably be viewed as a short-run effect for demand-side programmes. As previously discussed, as job opportunities increase in a growing economy, workers who are substituted against should be able to find jobs and substitution effects should diminish over time. However, it is important to note that this expectation is conjectural, as the topic of how substitution effects resulting from demand-side programmes change over time has been little studied empirically. In general, the estimates of the substitution effects resulting from demand-side programmes that are provided by survey and econometric studies seem to pertain more to the short-run than to the long-run.

---

[11]   In determining this value, it would be helpful to compare the quality of the individual studies listed in Tables 2.1-2.3 and put the greatest weight on substitution effect estimates from the higher quality studies. However, as indicated in the tables, the estimates from many of the listed studies were obtained from secondary sources. These sources typically simply reported study findings without assessing study quality. Moreover, a number of these studies were written in foreign languages such as Swedish and thus were inaccessible to us.

[12]   Some individual studies provide a range of estimates, rather than a single point estimate. For purposes of the discussion in the text, we use the mid-point of the range. For example, the range in the DeKoning *et al.* (1992) study in Table 2.2 is 69 to 76 per cent. Thus, we use a value of 73 per cent.

- In conducting sensitivity analyses of wage subsidy and public sector employment programmes, 60 per cent is a reasonable upper bound value and 30 per cent is a reasonable lower bound value to use for the short-run substitution effect.

- In reporting sensitivity results for demand-side programmes, it should be emphasised that if the economy is growing, substitution effects would be expected to diminish over time.

- In reporting the sensitivity results for public sector employment programmes, it should be emphasised that while some econometric estimates suggest that the substitution effect might be as large as 60 per cent, or even larger, findings from the survey approach imply that they are 30 per cent or considerably smaller.

Most Department for Work and Pensions programmes are supply-side programmes. Unfortunately, there are fewer substitution effect estimates than in the case of demand-side programmes on which to base conclusions. Moreover, the evidence that exist concerning the size of substitution effects that result from supply-side programmes is not very clear. All but one of the seven estimates from econometric studies of such programmes suggest that the substitution effect is zero or quite small, while all four of the estimates from general equilibrium models imply that it is quite large. Although six of the seven econometric estimates are for training programmes and three of the four estimates from general equilibrium models are for financial incentive programmes, it is not obvious that substitution effects should differ in size between these two programme types. By using both approaches to estimate substitution effects for the same programme, perhaps a start can be made at reconciling differences between the findings resulting from the two approaches. Until then, the substitution effect value that should be used in conducting a sensitivity test of cost-benefit estimates for supply-side programmes is not obvious.

One might argue, however, that because so few studies have been based on general equilibrium models and the models themselves have not been verified and are thus still speculative, the findings that rely on them should be ignored in conducting sensitivity tests. If this argument is accepted, then a very small value of, say, ten per cent or (as Fujiwara 2010 suggests) 20 per cent, which reflects findings based on the econometric approach, might be used as an upper bound for sensitivity analysis of supply-side programmes, while zero could be reasonably used as the lower bound. An alternative and perhaps better approach might be to forego sensitivity analyses altogether and simply caution readers about the possible existence of substitution effects by briefly reviewing the evidence presented in this report concerning them.

Because it is difficult to draw plausible conclusions from existing evidence concerning the size of substitution effects resulting from supply-side programmes, it may be best to forego sensitivity analyses for substitution effects in conducting CBAs of such programmes. Instead, readers of the studies can simply be cautioned about the possible existence of substitution effects by briefly reviewing the evidence presented in Tables 2.2 and 2.3 of this report. Alternatively, sensitivity analysis could be conducted using ten or 20 per cent as the upper bound value of the substitution effect and zero as the lower bound value, while, cautioning readers that findings based on general equilibrium models, which imply much larger substitution effects, are being ignored.

As in the case of demand-side programmes, substitution effects for supply-side programmes would be expected to diminish over time if the economy is expanding and those who are substituted again eventually obtain employment. In the case of supply-side programmes, empirical evidence exists that is consistent with this conjecture. For example, Adam *et al.* (2008) found that individuals who

received Jobseeker's Allowance in several Pathways to Work sites and, hence, competed for jobs with incapacity claimants who were mandated onto the Pathways programme, were 3.5 per cent less likely to move off benefits within six months, evidence consistent with the possible existence of a substitution effect. After 12 months, this effect was less than half as large. Based on a general equilibrium model, Dixon and Rimmer (2006) estimated that the displacement effects of financial incentive and job search programmes would be very large after the programmes were introduced, about 60 and 70 per cent, respectively; but disappear five years later in the case of the financial incentive programme and eight years later in the case of the job search programme. Their model, however, does not allow for exogenous job growth. The reduction in the substitution effect is instead predicted to occur because of downward pressure on wages caused by increases in labour supply resulting from the programmes.

2.4.2   Displacement effects

Separate empirical estimates of displacement effects do not appear to exist. However, displacement effects are mainly limited to wage subsidy programmes, and as indicated in the previous section, econometric estimates of the substitution effects engendered by wage subsidy programmes that are based on macro-data may incorporate displacement effects, as well as substitution effects. Thus, in conducting CBAs of wage subsidy programmes, sensitivity tests that assume substitution ratios of between 30 and 60 per cent, as recommended in the previous section, may be viewed as accounting for both displacement effects and substitution effects.

2.4.3   Equilibrium wage effects

It was suggested earlier that demand-side employment and training programmes might tend to put upward pressure on wages. Conceptually, as also suggested earlier, the effects of supply-side programmes on wages appear to be somewhat ambiguous, but seem more likely to be negative than positive. Unfortunately for our purposes, most empirical attempts to investigate the equilibrium wage effects of employment and training programmes do not differentiate between demand- and supply-side programmes, but instead estimate a combined effect for the two types of programmes. For example, one well-known study estimated the effect on the wage growth of average compensation per employee of total national expenditures on employment and training programmes in 19 OECD countries (OECD, 1993). The resulting estimated effect was positive and statistically significant in ten countries, negative and significant in two countries, and small and statistically insignificant in seven countries (including the UK).

However, Calmfors, Forslund, and Hemstrom (2002) summarise findings from three Swedish studies that did provide separate estimates for demand- and supply-side programmes. Two of these studies, as expected, found that demand-side programmes had positive effects on wages in the short-run, while the remaining study found no short-run effect for these programmes[13]. One of the studies found that supply-side programmes had a negative short-run effect on wages, as anticipated, one found a negative effect with one regression specification and no effect with another, and the third study did not find an effect.

To summarise, there is some indication that, in general, employment and training programmes put upward pressure on wages, although, there is at least a bit of evidence from Sweden that supply-side programmes may reduce wages. However, many studies of equilibrium wage effects have not obtained statistically significant results and few studies with statistically significant results have

---

[13]   Calmfors, Forslund, and Hemstrom (2002) also report on finding from a study that examined a public sector employment programme, but not supply-side programmes. This study, as anticipated, found that the programme it examined put upward pressure on wages.

found very large effects. Thus, in cost-benefit studies of employment and training employment programmes, it seems most reasonable to derive base-case estimates by assuming that there is no effect on wages. Sensitivity tests of the base-case for large-scale demand-side programmes might then be conducted by assuming that there is a small increase in wages of (say) one per cent or two. Similar sensitivity tests for supply-side programmes do not seem warranted. The evidence about even the direction of the wage effect resulting from these programmes is simply too weak.

> In cost-benefit studies of employment and training employment programmes, base-case estimates should be first derived by assuming no effect on wages. Sensitivity tests of the base-case for large-scale demand-side programmes might then be conducted by assuming a small increase in wages of perhaps one or two per cent. Similar sensitivity tests for supply-side programmes are inappropriate because of weak
> evidence concerning these effects.

2.4.4   Entry effects

Our assessment of the theoretical arguments concerning entry effects is that their importance is somewhat speculative. Empirical evidence about the magnitude of entry effects is quite limited. With just a few exceptions, this evidence is all based on aggregate macro-data. They typically rely on time series comparisons of caseload size across geographical areas that vary in terms of programme participation. Moreover, the evidence is all limited to welfare-to-work programmes targeted on the US and Canadian welfare populations, often lone parents. Because they typically rely on macro-data, rather than micro-data, studies of the entry effects of these programmes have usually been conducted separately from analyses of the impacts of these programmes on employment, although both studies are sometimes performed as part of the same evaluation.

Most of the findings from these studies are consistent with what one might anticipate: voluntary programmes for welfare recipients appear to encourage modest entry onto the welfare rolls by providing services that may otherwise be difficult to obtain, while mandatory programmes seem to discourage entry modestly by raising the cost of receiving welfare (that is, they make it more burdensome to receive welfare). Thus, the effects for voluntary programmes appear to partially defeat one of the major purposes of such programmes, which is to reduce welfare caseloads.

Table 2.4 presents findings from ten empirical studies of training programmes for welfare recipients that provide some evidence on entry effects. All the listed programmes were located in the US.[14] Most of these studies are based on aggregate caseload data. They rely on regression adjusted

---

[14]   A brief summary of the programmes listed in Table 2.4 follows. The Washington State Family Independence Program (FIP) provided financial incentives to Aid to Families with Dependent Children Program (AFDC) recipients to encourage education, training, and work. The Oregon Demonstration Program (ODP) provided group job search, basic skills education, vocational training, and counselling. The Massachusetts Employment and Training Choices Program (ET) provided a variety of training opportunities to welfare recipients including on-the-job training, work experience, occupational training, and basic education. Job placement was also emphasised. The Ohio Community Work Experience Program (OHCWEP) provided a variety of services to AFDC recipients including job search, work experience, supported work, and on-the-job training, The West Virginia Community Work Experience Program (WVA) required AFDC recipients to participate in work experience (i.e. they were assigned to government and non-profit agencies in exchange for their grant). The California Greater Avenues for Independence Program (GAIN) was a demonstration programme that tested the effects of basic education, skills training, work experience, and job search on AFDC recipients.

comparisons across study sites with and without the programme of interest and across time. Results from these studies tend to be sensitive to specification changes and to changes in methodology. Thus, as indicated in Table 2.4, there are two instances in which different researchers obtained somewhat conflicting findings. The first six studies listed in the table are of voluntary programmes and the remaining four are of mandatory programmes. The last three columns in the table report the study findings. The estimates in the first of these three columns pertain to the net effect of the studied programmes on the size of welfare caseloads. That is, they incorporate the effects of these programmes on entry onto, as well as exit from, the caseload. The second from last column indicates the estimated direction of exit effects.[15] Caseload exit would occur if a welfare-to-work programme successfully places current welfare recipients into jobs with sufficiently high earnings so they no longer qualify for benefits. The last column presents estimates of programme effects on applications for welfare benefits, a measure of entry effects. As can be seen, there are far fewer separate estimates of entry and exit effects than of overall net effects.

In order to make the results presented in Table 2.4 as comparable as possible across studies, the reported estimates are averaged over the time period examined in the study, which varies between one and 12 years after the programme began. This averaging obscures some information pertaining to the timing of the caseload effects. For example, most of the studies tend to suggest that programme effects on caseloads initially grow and then either level off or diminish.

Table 2.4 indicates that the net effect of mandatory welfare-to-work programmes on caseload size is negative, implying that they increase exits from the welfare rolls or deter entry onto the rolls or both. This is consistent with one goal of these programmes, which is to reduce caseloads. In general, mandatory programmes should increase the exit of current recipients from the rolls, while also deterring entry onto the welfare rolls, unless they provide training that individuals feel will increase their future earnings.

---

[15]   Estimates for exit effects were reported in different units across the studies. Therefore, we only indicate the estimated direction of the effect.

Table 2.4   Effects of training programmes on welfare caseloads

Welfare programme	Study	Data and evaluation design	Months since covered by estimates beginning of programme	Kind of welfare case	Net effect on caseload	Direction of effects on exits from welfare (%)	Effect on welfare applications (%)
Voluntary programmes							
FIP	Wolf (1990)	Comparison of matched treatment/control sites using monthly time series	10	1-parent 2-parent	2.0 28.0	n.e. n.e.	n.e. n.e.
FIP	Marcotte (1992)	Comparison of matched treatment/control sites using monthly time series	42	1-parent 2-parent	5.0 to 6.0 19.0 to 62.0	n.e. n.e.	n.e. n.e.
FIP	Wissoker and Watts (1994)	Comparison of treatment/control sites using monthly time series data. All state welfare offices included in analysis	60	1-parent 2-parent	6.9 37.3	n.e. n.e.	no effect no effect
ODP	Johnson et al. (1992)	Comparison of matched treatment/control sites using monthly time series	12	1-parent	6.6	decrease	positive[a]
ET	O'Neill (1990)	Quarterly time series used to compare pre- and post-programme periods	52	1-parent 2-parent	0.0 to 4.0 -7.0 to -1.0	n.e. n.e.	n.e. n.e.
ET	O'Neill (1990)	Compared MA caseloads with the remaining states using CPS data	52	1-parent	negligible	n.e. n.e.	n.e. n.e.
ET	Garasky (1990)	Quarterly time series used to compare pre- and post-programme periods	36	1-parent	-11.2 to -8.8	n.e. n.e.	n.e. n.e.
Mandatory programmes							
OHCWEP	Schiller and Brasher (1992)	Pooled quarterly time series/county cross section	58	1-parent 2-parent	-7.0 -15.4	large increase[c]	5.2[c]
Ohio Work Program	Chang (1996)	Pooled monthly time series/county cross section	142	2-parent	-0.4	n.e.	n.e.
WVA	Chang (1996)	Pooled monthly time series/county cross section	144	2-parent	-8.5	[b]	-20.0
GAIN	Phillips (1993)	Pooled monthly time series/county cross section	54	1-parent 2-parent	n.e. n.e.	n.e. n.e.	-4.0 -4.5

[a] Programme effect on entry positive, but not measured as a per cent of applicants.

[b] Some estimates indicated that programme exits increased and other estimates implied that they decreased.

[c] Combined effects for 1-parent and 2-parent families.

n.e. = not estimated

Key to programmes: FIP=The Washington State Family Independence Program, ODP=The Oregon Demonstration Program, ET=The Massachusetts Employment and Training Choices Program, OHCWEP=The Ohio Community Work Experience Program, WVA=The West Virginia Community Work Experience Program, GAIN=The California Greater Avenues for Independence Program.

The findings for caseload size for voluntary programmes are mixed. The ODP and the Washington State FIP appear to have increased the size of the AFDC caseload,[16] while the Massachusetts ET Choices Program seemed to have decreased the size of the caseload.[17] This ambiguity is not surprising. As discussed above, voluntary programmes should have a positive entry effect, but they could cause exits from the existing caseload to either decrease (as welfare recipients stay on the rolls to participate in training activities) or increase (as recipients become more job ready).

Numerous studies have utilised individual-level data to estimate the effects of welfare-to-work programmes on the receipt of welfare. Because many of these studies cover only individuals who are on welfare at the time the programmes began, the effects on welfare receipt are often interpreted as effects on welfare exit. Only three studies have directly examined exit effects from welfare using aggregate caseload data. Findings from Johnson, Klepinger, and Dong's (1992) study of a voluntary programme in Oregon indicate that exits decreased, perhaps because welfare recipients remained on the rolls longer in order to complete their training (a 'locking-in effect'). Estimates from Schiller and Brasher's (1992) study of a mandatory programme in Ohio, suggest, as anticipated, that exits from AFDC increased, although this increase in exits was partially offset by an increase in welfare applications. Findings from Chang's (1996) evaluation of a mandatory programme in West Virginia were ambiguous.

Five aggregate-level time series studies have directly examined entry effects by looking at whether applications for welfare increase or decrease as a result of a welfare-to-work programme, although the value of the empirical results is reduced by their sensitivity to model specification changes. All five studies compare welfare application rates in sites that have a welfare-to-work programme with application rates in sites that do not have the programme. Three of the five studies are consistent with expectations that voluntary programmes for welfare recipients encourage entry onto the welfare rolls and mandatory programmes discourage entry. Of the two studies of voluntary programmes, Johnson, Klepinger, and Dong (1992) find that a voluntary programme in Oregon had a positive entry effect, but Wissoker and Watts (1994) do not find a positive entry effect for a voluntary programme in the state of Washington. Of the three studies of mandatory programmes, two indicate, as anticipated, that entry effects were negative (Chang 1996, and Phillips 1993), while the remaining study finds no evidence of a negative entry effect, along with a much larger increase in exits (Schiller and Brasher, 1992).

In addition to the five studies just discussed, there are two studies of entry effects that are not listed in Table 2.4. The first of these (Card and Robins, 2005) is not listed because it does not examine effects on welfare applications. Instead, it looks at the effects of a demonstration programme, Canada's Self-Sufficiency Project (SSP), on persons in British Columbia and New Brunswick who had already applied for welfare. SSP provided generous financial incentive payments to these persons if they remained on the welfare rolls for over a year after the programme began and then took a job in which they worked at least 30 hours a week. The supplement was equal to one-half the difference between a 'target' earnings level (initially $37,000 in 2007 Canadian dollars in British Columbia and $30,000 in New Brunswick) and an individual's earnings. Because the income individuals could receive if they worked at least 30 hours was much larger under SSP than under welfare, the programme provided a strong monetary incentive to leave welfare and work full time. The analysis is based on a random assignment experiment that was specifically designed to determine the extent

---

16   The very large positive estimate for 2-parent families under the Washington State FIP is probably because this programme greatly liberalised the standards these families had to meet to qualify for AFDC. The rules were not similarly liberalised for 1-parent families.

17   The findings of O'Neill (1990) suggest that after the first year, the voluntary ET Choices Program may have had a small positive effect on AFDC caseloads.

to which welfare applicants extended their welfare stay in order to become eligible for the financial incentives – that is, to enter SSP. The results indicate that there was a three percentage point rise in the fraction of welfare applicants who remained on the welfare rolls for at least a year. The study authors conjecture that because the entry effect on those who already applied for welfare was so modest and the cost and stigma of newly applying for welfare are appreciable, it is unlikely that the SSP incentives would encourage more than a negligible number of additional welfare applications. The SSP experiment arguably provides the best available information on the entry effects of programmes intended to encourage employment.

Moffitt (1996) illustrated entry effects for voluntary and mandatory programmes using a micro-simulation model. This study was excluded from Table 2.4 because it is not based on real programme data. Moffitt's estimates, which are often sizable, suggest that a mandatory programme for AFDC recipients with a heavy participation time requirement would reduce entry into welfare, while a voluntary programme would increase entry. Much of the latter effect, however, is attributable to an assumption that training provided by the programme would reduce the stigma attached to welfare receipt, rather than from the training itself.

It does not appear very feasible to directly account for entry effects in conducting CBAs of employment and training programmes for benefit recipients. Although there is evidence that voluntary programmes encourage entry, while mandatory programmes deter entry, the magnitude of entry effects seems likely to be driven by the specific details of the design of the programme being evaluated. Moreover, the empirical evidence on entry effects is quite sketchy and all of it is limited to programmes located in North America. However, the likely direction and importance of entry effects should be discussed in reporting cost-benefit findings of employment and training programmes for benefit recipients.

> Although empirical information about entry effects is insufficient for sensitivity tests, their likely direction and importance should be discussed in reporting cost-benefit findings of employment and training programmes for benefit recipients.

2.5   Conclusions

In this section, we briefly list our key conclusions about how various general equilibrium effects should be treated in conducting CBAs of employment and training programmes.

- In conducting sensitivity analyses of demand-side employment and training programmes, 60 per cent can reasonably be used as an upper bound value of the substitution effect and 30 per cent as a lower bound value. In principle, this sensitivity test should take account of displacement effects that result from wage subsidy programmes, as well as substitution effects resulting from such programmes. However, in reporting the sensitivity results for public sector programmes, it should be emphasised that while econometric estimates suggest the substitution effect might be as large as 60 per cent, or even larger, findings from the survey approach imply that they could be smaller than 30 per cent. Moreover, even if the substitution effect was initially as large as 60 per cent, it would be expected to diminish over time if the economy is growing.

- Sensitivity analysis of supply-side programmes might be conducted using 10 or 20 per cent as the upper bound value of the substitution effect and zero as the lower bound value, while cautioning readers that findings based on general equilibrium models, which imply much larger substitution effects, are being ignored. However, because it is so difficult to draw plausible conclusions from existing evidence concerning the size of substitution effects that result from supply-side programmes, it might be better to forego sensitivity analyses for substitution effects in conducting CBAs of such programmes. Instead, readers of the studies can simply be cautioned about the possible existence of substitution effects by briefly reviewing the evidence presented in Tables 2.2 and 2.3 of this report.

- In cost-benefit studies of employment and training employment programmes, base-case estimates should be first derived by assuming that there is no effect on wages. Sensitivity tests of the base-case for large-scale demand-side programmes might then be conducted by assuming that there is a small increase in wages of (say) one or two per cent. Similar sensitivity tests for supply-side programmes do not seem appropriate because the evidence concerning these effects is too weak.

- Although empirical information exists about entry effects, it is insufficient for conducting sensitivity tests. Instead, the likely direction and importance of entry effects should be discussed in reporting cost-benefit findings of employment and training programmes.

3   Subgroup impacts

In this chapter we consider how to account for the fact that the net impact on employment might vary among sub-groups. We further consider a closely related topic: the possibility that the distribution of impacts varies and is not satisfactorily represented by the mean impact. These topics are not currently addressed by the cost-benefit framework.

There are several types of subgroup variables upon which subgroup impacts can be formed:

- firstly, various exogenous variables (that is, variables that are not affected by a programme such as gender and age) can cause impacts to vary;

- secondly, impacts on one endogenous variable (that is, a variable affected by a programme such as hours worked) may affect other endogenous variables (such as whether benefits are received);

- thirdly, impacts of a programme may vary for different outcome subgroups of continuous variables such as earnings (quantiles)[18]. For example, for a programme, the impacts for those in the lowest earnings may differ from the impacts for those at the highest earnings, and the impacts for both groups may be different from the programme's impact at the mean; or for two different programmes which have the same average effect, for one programme the effects may be concentrated at the lower end of the earnings distribution while for the other programme the effects may be concentrated at the higher end.

In this chapter, we provide guidance on how to conduct benefit-cost analysis for each of these three cases.

3.1   Subgroups

In estimating the impacts of training and employment programmes, analysts usually focus on the overall population of programme participants. However, they also estimate programme impacts on subgroups of these persons. Subgroup analysis is often of interest regardless of the finding for the overall programme impact. On the one hand, if the overall impact is positive and statistically significant, then it seems appropriate to focus resources on those most helped. On the other hand, if the overall impact is not statistically significant, then subgroup analysis may be used to see whether there are some persons who are nevertheless helped. Additionally, there may be a 'hard to help' group which is of particular interest – for such a group, a subgroup impact would better assess how the programme meets the needs of this group. When estimated programme effects appear to differ among subgroups, separate cost benefit analyses (CBAs) for each subgroup is a natural next step. This is sometimes done, but not always.

Subgroups may be defined in numerous ways, which will, of course, vary depending on the programme being examined. The list below illustrates some of the more important subgroup possibilities, based on individual characteristics of the claimant:

- gender;

- education;

- family size;

---

[18]   In this context quantiles are subgroupings – points taken at regular intervals from the distribution of an outcome variable, such as the earnings distribution – hence, there can be quantiles of a distribution of earnings impacts.

- ethnicity;

- age;

- health condition;

- benefits status;

- employment history;

- geographic location.

Important informative subgroups can be formed not only from individual characteristics, but from programme delivery aspects or variations in cost. For example, in the New Deal for Disabled People (NDDP) cost benefit analysis (CBA), a key subgroup was formed by job-broker size (Greenberg and Davis, 2007). This example is discussed below in Section 3.2.1.

Subgroups based on individual characteristics, such as employment status and benefit status are typically defined on the basis of status **prior to programme entry** so as not to introduce endogeneity into the analysis. This avoids the problem that occurs with subgroups defined for these variables after programme entry, where something related to the changes in employment/benefit status from the start of programme entry/eligibility and hence the subgroup definition have caused the impact observed in the analysis. Sometimes, however, endogeneity cannot be avoided because the analysis is directly concerned with whether a programme causes people to move from one subgroup to another – for example, from being non-employed and on benefits to being employed and off benefits. As seen in Section 3.2.2, such topics can be examined in a way that does not bias the analysis.

Although analyses of programme impacts are usually most concerned with average effects (that is, with effects on a typical member of the programme group), sometimes there is also an interest in how programme effects are distributed among the programme group – for example, on whether programme impacts are distributed fairly evenly among programme participants or concentrated upon relatively few participants, with effects on the remaining participants small or absent.

> - There may be a group which is of particular interest. If so, a subgroup impact can assess whether the programme meets the needs of this group.
>
> - When estimated programme effects appear to differ among the subgroups, separate CBAs for each subgroup may then be conducted, although this is not always done.
>
> - Subgroup definitions must be carefully defined for analysis to avoid the technical analytical problems of endogeneity.

In the remainder of this section, we first examine the relevance of subgroup impacts for CBA. We then discuss some problems and issues that arise in examining subgroup impacts. We end with a discussion of estimating the distribution of impacts over programme participants.

3.2   Relevance of subgroup impacts for cost benefit analyses

As indicated above, when separate impacts are estimated for different subgroups, separate CBAs are sometimes also conducted for each subgroup. Although this is done fairly frequently, it is obviously practical to do it for only a limited number of subgroups. Moreover, not all of the information needed for the CBA may exist separately by subgroup. For example, programme operating cost information by subgroup needs to be available for a full CBA by subgroup. Knowledge of key subgroups of interest in advance can clearly facilitate the analysis by ensuring that costs and benefits will be available for each subgroup.

3.2.1 Job-broker size

This last point is illustrated by a CBA of the NDDP, a voluntary programme that attempted to increase employment among Incapacity Benefit claimants (Greenberg and Davis, 2007). The NDDP programme was delivered locally by Job Brokers, which were a mixture of public sector, private sector, and charitable organisations. In conducting the cost benefit analysis of the NDDP programme, Greenberg and Davis (2007) provided separate findings for claimants that registered at small Job Brokers and those that registered at large Job Brokers. This seemed important because there was evidence that NDDP was more costly to operate at small Job Brokers than at large Job Brokers and that NDDP reduced the incapacity benefits received by registrants at small Job Brokers by more than it reduced the incapacity benefits received by registrants at larger Job Brokers. Separate CBAs were also conducted for claimants that were already receiving incapacity benefits when NDDP began operating (stock of claimants) and those that began receiving benefits after operations commenced (flow of new/repeat claimants), as NDDP appeared to have larger impacts on the earnings and the receipt of benefits of the first subgroup than on the second. In this case, full CBAs were not possible because while separate estimates of impacts on earnings and benefit receipt were available for the two subgroups of interest, separate estimates of programme operating costs were not. Thus, it was necessary to assume that operating costs were the same for the two subgroups.

3.2.2 Work/benefit outcomes

The CBA of the Pathways to Work programme (Adam *et al.,* 2008), which, like the NDDP, was intended to encourage employment among Incapacity Benefit claimants, used subgroup analysis for a very different purpose. In this case, the programme sample was divided into the following six subgroups, which exhaust all the possibilities:

1   employed 1-15 hours a week, not receiving incapacity benefits;

2   employed 16-29 hours a week, not receiving incapacity benefits;

3   employed over 29 hours a week, not receiving incapacity benefits;

4   employed 1-15 hours a week, receiving incapacity benefits;

5   not employed, receiving incapacity benefits;

6   not employed, receiving incapacity benefits.

An impact analysis was first used to estimate the percentage of the programme sample that would be in each subgroup with and without the Pathways programme, with the difference between the two being a measure of the impact of the programme. Then, **based on the average earnings and the distribution of earnings within each subgroup**, the Institute for Fiscal Studies **micro-simulation model** was used to predict the amount of various benefits (e.g. incapacity benefits, housing benefits, council tax benefits, and tax credits) that would be received and taxes (e.g. income taxes and employee and employer national insurance contributions) that would be paid by the persons in each subgroup in the absence of the programme and with the programme. The difference in each case provided a measure of a programme benefit or cost. This approach may well be unique to the CBA of the Pathways to Work programme. Impacts on benefit and tax payment amounts are more often determined in CBA for a single 'representative' (i.e. average) member of the programme group that is observed to move from the programme into employment. Unlike the Pathways' CBA, however, this approach does not incorporate important non-linearities in the tax and benefit system, and thus may result in biased estimates.

3.3   Key issues in estimating subgroup impacts

3.3.1   Statistical significance, sufficient sample size and intrinsic differences between the groups or they were treated differently under the programme?

Several important issues arise in estimating impacts for separate subgroups and thus in subsequently using these estimates in CBAs. First, although the sample used in an impact analysis may be large enough to detect overall programme impacts, once broken into subgroups the resulting smaller subsamples may produce a finding of non-significance for particular subgroups when a true difference exists.

Second, a statistically significant impact for a particular subgroup may occur because of chance alone. This becomes increasingly likely as the number of subgroups that are examined increase. To minimise this problem, it is important to test whether programme impacts on different subgroups differ significantly from one another, not just whether the individual impacts are statistically significant. Testing whether subgroup differences are statistically significant reduces the possibility of obtaining false positive or false negative findings. It is important to recognise, however, that although significance tests can be performed for programme impacts on earnings and benefits, which are estimated for samples of individuals, this is rarely possible for programme operating costs unless different programme participants are assigned to different service delivery organisations. The CBA of the NDDP programme, which was mentioned above, is an important exception because services were delivered by a number of different Job Brokers – however, the sample size was very small (19 Job Brokers).

Third, it is often unclear as to whether differences in programme impacts on different subgroups are attributable to intrinsic differences between the groups or because they were treated differently under the programme. For example, a difference in impacts on earnings between programme participants with and without previous work experience may result if a welfare-to-work programme provides help in job search for the former and vocational training for the latter. This difference in impacts might or might not have resulted had the two subgroups received the same services. Where differences in subgroup impacts are produced and found, it is important to explore the delivery of services to these subgroups to help understand whether the subgroup impacts are attributable to intrinsic differences between the groups or because they were treated differently under the programme. This is especially important if the programme being examined can be described as a 'black box' or a 'bundle of potential services'.

- It is important to **test whether programme impacts on different subgroups differ significantly from one another using statistical tests**, not just whether the individual impacts are statistically significant.

- When significant differences in subgroup impacts are found, it is important to explore the delivery of services to these subgroups to help understand if the subgroup impacts are attributable to intrinsic differences between the groups or because they were treated differently under the programme.

- Significance tests are rarely possible for programme operating costs. However, there are exceptions, and so this possibility should always be explored.

3.3.2 Confirmatory versus Exploratory subgroup analyses

To minimise problems in subgroup analysis, it is very helpful to define the subgroups that are to be investigated **before** the evaluation begins. Ideally, this should be based on a theory as to why differences in programme impacts on these subgroups are anticipated and what the directions and magnitudes of the differences are likely to be. This helps reduce the number of subgroups that are analysed, thereby minimising the possibility of erroneously finding a statistically significant difference between subgroups when a true difference does not exist. Moreover, it makes it possible to collect data on subsamples that are of sufficient size for each relevant subgroup so that statistically significant impacts can be estimated. With insufficient sized subsamples, the possibility of a false finding of statistically non-significant differences among subgroups is higher. Subgroups to be investigated should be defined at the outset of the evaluation and based on data information for the population defined at programme eligibility point.[19]

> Subgroups to be investigated should be defined at the outset of an evaluation.
>
> To conduct a CBA for subgroup impacts requires that data on costs be collected for these subgroups.

3.3.3 Use all subgroup analyses with caution

Findings from the sort of 'confirmatory' subgroup analysis just described should be treated with considerable caution, even when the differences between the impacts for different subgroups are statistically significant. Findings from an 'exploratory' subgroup analysis in which differences among subgroups are investigated on an ad hoc basis after the overall programme impact is estimated should be treated with much greater caution. This is particularly important when the overall impact of a programme is statistically non-significant and there is great temptation to try to discover a subgroup for which the programme 'works'. Caution is needed because, as set out in Section 3.3.2, the statistical inference is affected by whether a confirmatory or an exploratory subgroup analysis is conducted – the statistical inference for confirmatory subgroup analyses is more reliable and informative, and hence superior to exploratory subgroup analyses.

3.4 Programme effects on the distribution of the outcome

The discussion so far regards how mean impacts might differ across various subgroups of participant types. This section discusses how impacts might differ across subgroups of the outcome variable. This is usually of particular relevance for continuous variables, such as earnings, which have a distribution of values, and the impact at the average may not be the only point of interest.

3.4.1 The current evaluation evidence of average effects informing DWP cost benefit analyses

Distributional effects are particularly important in the context of social policies, as these are often directed at particular targets which reflect only a part of the distribution – such as low earners (not all earners) in order to reduce poverty. An example for such policies is the recent Employment Retention and Advancement (ERA) demonstration, which in particular aimed to improve the situation of low income earners.

---

[19] This can mean that administrative data is the main source for subgroup information in order to provide population information, rather than potentially selected (and hence subject to bias) or endogenous information (such as from the outcome time point or sometime beyond the programme entry point).

Surprisingly, even the ERA programme has so far only been evaluated in terms of the mean effectiveness on earnings (Dorsett *et al.,* 2007), although this mean effect is likely to differ from the effects for low earners or for high earners. In such cases, the estimation of programme effects for particular distribution subgroups (or quantiles[20]) is more informative – continuing with the example about a programme targeting low earners, low earners might be defined as those with earnings in the bottom 25 per cent of the distribution of earnings, and the quantile impact at this point may differ from the mean impact.

3.4.2   Quantile treatment effects

Chernozhukov and Hansen (2005) point out that quantile regression models '*...characterise the heterogeneous impact of variables on different points of an outcome distribution*'. In other words, they can be used to reveal whether the impacts differ across the outcome distribution subgroups.

Distributional effects are in particular important in the context of some welfare-to-work reforms as they may improve the situation of a particular set of workers in the low segment of the wage distribution as they aim at reducing poverty.

Other applications of quantile treatment effects could consider differential effects on particular segments of the distribution of unemployment duration, providing equally important evidence on the policy, e.g. how job seekers at different points in the distribution of unemployment duration are affected by a policy.

When focusing on mean impact, the distributional effects of programmes remain unidentified, and may result in the CBA informing about an average negative outcome, which may in fact consist of positive and negative outcomes for different quantiles along a particular distribution of unemployment or previous benefit duration. Ultimately, disregarding quantile treatment effects may result in factually inadequate advice about the provision of policies for these groups. This is because the evaluation evidence does not inform about the group for whom the programme was targeted.

> Quantile treatment impacts for earnings or duration may be needed if the programme is targeted at (or might be desired to be targeted at) particular subgroups of earnings or duration.

3.4.3   Previous quantile treatment effects empirical evidence

The effects of training on the lower part of the **earnings distribution** can be more interesting for policy makers than the effect at upper quantiles as an explicit objective of many programmes is to reduce poverty for the low-wage groups. There is no existing evidence that we know of on the importance of such heterogeneity of treatment effects for the United Kingdom.

Internationally, Firpo (2007) shows that some programme effects can vary with the quantiles of the earnings distribution – Firpo(2007:39) reanalysed the Dehija and Wahba (1999) Panel Study of Income Dynamics (PSID)[21] experimental data for the United States' National Supported Work Programme and found negative effects of the programme for the extreme upper parts of the

---

[20]   In this context, quantiles are points taken at regular intervals from the distribution of an outcome variable, such as the earnings distribution. Some quantiles have special names and are more commonly used: the 2-quantile is called the median which is also a key measure of central tendency; the 10-quantiles are called deciles, which are often used in income distribution literature; the 100-quantiles are called percentiles.

[21]   The Panel Study of Income Dynamics (PSID) is a longitudinal survey of a representative sample of US individuals and families, which has been ongoing since 1968.

earnings distribution, due to the high earnings of the control sample; however the average impact of the programme on the treated was $1,740, substantially positive and statistically significant. Firpo was also able to use quantile treatment effects to calculate the median impact on the treated, and this was $1,927 – for distributions, the mean and the median are two different measures of central tendency, and for earnings, the median is usually preferred as the measure of the centre of the distribution as the median is not influenced upwards by high value outliers at the upper end of the distribution; in this case, the mean and median impact were relatively close suggesting the centre of the distribution was satisfactorily represented by the average.

Other applications estimating quantile treatment effects focus on **quantiles of particular durations** (e.g. the duration of unemployment or the duration of benefit receipt). Lalive (2007) estimates the impact of a reform of Austrian unemployment benefits for those with extremely long duration in the upper quantiles of the duration of unemployment. Koenker and Bilias (2000) also show how quantile treatment effects can be used to analyse duration data by re-analysing experimental data from the Pennsylvania Reemployment Bonus Experiments.

In estimating duration outcomes, all studies face the difficulty of estimating quantile treatment effects for right-censored durations, as unemployment spells may still continue and so standard estimators are generally not consistent and Censored Quantile Regression algorithms must be used (Powell 1986).

Econometric methods for applying quantile treatment effects are set out in Appendix C.

3.4.4   Recommendation for quantile treatment effects for some programmes

A number of DWP programmes could be assessed for quantile treatment effects as they might affect particular parts of the wage or benefit duration distributions differently. The results of such analyses can inform decisions regarding which groups of the population to target for programme access in the future.

Applications of these methods are arguably most appropriate when a particular part of an outcome distribution (of earnings or benefit duration) was the desired target of a programme.

For example, the DWP programme ERA was designed to improve job retention and career advancement prospects for Britain's low-wage workforce (i.e. some lower part of the distribution). Therefore, for ERA, the quantile treatment effects of earnings are likely to be the main effects of interest for policy makers.

The DWP programme Pathways to Work could be another example, as this could be re-examined with regards to the impact on Incapacity Benefit (IB) durations. The approach taken by Lalive (2007) could be followed as the programme introduction also mimics a natural experiment (with pilot/non-pilot areas, etc.). For Pathways to Work, it would be interesting to investigate the impact of the programme on long-term IB claimants or on particular quantiles of the distribution of benefit durations. However, since Pathways to Work was not randomly assigned and the rollout areas in the programme design were not randomly assigned, controlling for selection effects in an econometric model could be quite complicated.

4   Duration of benefits and employment, wages, and multiple participation

4.1   Duration/sustainability currently used in DWP cost benefit analysis

For some cost benefit analyses (CBAs) of employment and training programmes conducted internally in the Department for Work and Pensions, estimates of the employment generating effect of the programmes originate from studies for a particular cohort (or a stock of participants) at a **particular date** following participation and not for continuous outcomes.

For such analyses, the point estimates of employment impact are multiplied with actual participation figures in order to derive an estimate of the total 'additional jobs' attributable to the intervention. To derive an estimate of the 'total additional job years' created by a programme an estimate of how long a person who obtains one of these additional jobs continues to be employed is also needed (see Greenberg and Knight, 2007). However, when constrained into following this approach, , there is often a lack of direct evidence gained from programme evaluations regarding the duration of post-programme employment.

Deriving estimates of average job duration from the Work and Pensions Longitudinal Study (WPLS) would improve the robustness of the ex post CBAs for a number of DWP employment and training programmes. With the ten years of WPLS data available, some of the longer-term outcomes can be estimated directly for earlier introduced programmes that started near or soon after the beginning of the data, 1999, as these data offer continuous observations of participants and non-participants for a period of up to ten years, allowing the analysis of the sustainability of employment outcomes after the programme and the exit from benefits. For some programmes, two to five years might be available. More recent programmes would have less follow-up years available.

> - We recommend replacing relatively weak assumptions about the duration of a subsequent job by estimating the duration of employment directly using the available information in the WPLS.
>
> - Ideally these should be net impact estimates, where suitable comparison groups are identifiable and where data exists for them

Unfortunately, when using the WPLS for these purposes, substantial data processing is necessary in order to correct for some weaknesses in the data. For example, the employment records' start and end dates have a lot of missing information. Indeed, because of this, some recent evaluations have used WPLS data in addition to survey data requesting detailed information from participants and non-participants in order to estimate the sustainability of employment effects and benefit following participation (see Riccio *et al.*, 2008).

In the following sections, we first review the estimation of programme effects for duration outcomes. We then turn to a number of indicators that we explore in empirical estimates. Finally, we suggest further empirical work which can be done.

4.2   Models to analyse duration outcomes of social programmes

The time spent on benefit, the transition from benefit to employment and the sustainability of employment are variables of interest when evaluating employment and training programmes. It can be in particular relevant to estimate the time spent on unemployment benefit, i.e. the 'survival function' for benefit duration, as some programmes might reduce the time spent on benefit.

There are numerous publications estimating the causal effect of programme participation on duration. Methods available include:

• modelling with a **parametric** baseline hazard rate specification in the context of regression models that allow the prediction of how the survival on benefit is influenced by the participation in programmes;

• models that estimate causal effects on duration **without a parametric** baseline hazard; and

• models evaluating policy effects for duration outcomes, using one of these two approaches, but **within matched samples**.

4.2.1   Hazard models

Hujer *et al.* (1997) evaluate the effects of particular policies combining matching methods that control for observable characteristics with hazard models. A hazard rate is the instantaneous transition (e.g. from unemployment to employment) rate of one individual at **a particular time** per unit of time period.

Hazard models allow one to evaluate the impact of particular policies on the transition from unemployment to employment. One can simulate hazard rates of leaving benefit separately for participants and non-participants. The estimated survivor function can be simulated too in order to show how the probability of remaining unemployed is affected by the policy.

Hazard rate approaches can also simultaneously model the transition into programmes and post-programme outcomes and may more flexibly allow consideration of dynamic selection issues arising from unobserved variation in both programme participation and outcomes (Abbring and Heckman, 2007).

Abbring and van den Berg (2003) elaborate how, under particular assumptions, a non-parametric model of both transition to programme participation and outcomes can be estimated, identifying the causal effect of the treatment in non-experimental data as a function of observed covariates, general dependence and unobserved variation.

The estimated coefficients of these models can be used to evaluate the impact of a particular policy on the duration of unemployment for a particular subgroup or the impact on the average duration of unemployment across all groups. These estimates can then be used to derive estimates of fiscal savings or gains that are due to the policy.

4.2.2   Matching approaches

Hazard models are best suited to estimate policy effects on initial transitions (e.g. into employment). Consideration of the determinants of any later transition (e.g. back into unemployment) is typically

ignored. For some policy questions it may be more informative to estimate outcomes on average days on benefit or employment rates in particular periods following participation directly.

In a number of previous evaluation studies, combined matching approaches with regression functions in matched samples have been used in order to estimate the effect on such employment outcomes directly (see, for example, Bergemann *et al.*, 2009). Such models estimate individual cumulated treatment effects for a particular time after the beginning of a programme in matched samples using propensity score matching. The cumulated programme effect over time can then be estimated for any interval after the beginning of treatment.

4.2.3   Programme effects on post-participation time windows

Hazard models or matching in combination with hazard models are usually related to the particular spell following programme participation and the transition from the programme into employment. In order to analyse the effectiveness of programmes, it may be especially relevant to consider the **sustainability** of such outcomes.

Since WPLS data record, in principle, employment and benefit status on a daily basis[22], we can model sustainability of employment outcomes by analysing the time spent in employment following programme participation. As shown in the evaluation of the longer-term outcomes of Work-Based Learning for Adults (WBLA), this dataset can be used to estimate the gain in days of employment (Speckesser and Bewley, 2006).

WPLS data can also allow one to estimate the average post-participation wage of both participants and non-participants. This would provide an estimate of the net wage effect of programmes and this could be incorporated in CBAs. However, a comparison group must be chosen. If there is uncertainty over the suitability of this comparison group, the report of the estimation results must indicate that they be interpreted with caution.

Speckesser and Bewley (2006) use WPLS data to estimate a variety of long-term outcomes for the evaluation of WBLA. Since WPLS data are organised as spells with start and end dates, outcomes are recorded on a daily basis and allow both the calculation of averages in particular time intervals or cumulative outcomes. Depending on the cohorts chosen for the participation in particular programmes, we may be able to observe outcomes for many months after participation.

As shown in Chapter 5 of the report, in which we report the gross outcomes of the various New Deal programmes on participants, WPLS data can be used to derive estimates to feed into CBAs. In combination with propensity score matching or difference-in-differences methods, the WPLS could be used to gain net impact estimates of the true causal employment, wage, or benefit effects of a programme.

Wage and duration impacts have been estimated for some DWP programmes in the past. The primary obstacle to analysis is that often there is no comparison group and survey data is often limited because of the chosen research design of the survey. In lieu of using survey data one can use administrative data. As some important programme participation data are already included in WPLS data (e.g. the New Deals) or can be merged easily using the encrypted National Insurance numbers

---

[22]   However, note that the WPLS does have recording issues. For example, among other issues, the WPLS records end dates of benefit spells only imprecisely, the self-employed are not included, nor are those whose income is below the taxable threshold.

to other data[23], the production of net programme effects on cohort-specific post-participation time windows would improve the quality of the assumptions that are used when conducting internal CBAs.

> To help inform internal CBAs, steps should be taken to ensure that necessary estimates are collected in order to evaluate programmes. Ideally, net impacts for these items should be estimated. We would suggest to always produce gross estimates for the average:
>
> - number of days employed in the first 12 (and 24) months following programme participation;
>
> - number of days on benefit in the first 12 (and 24) months following participation;
>
> - duration of participants' first jobs following programme participation; and
>
> - wage or earnings of participants' first jobs following the programme participation.

---

[23]   Analyses for impact of the programme on participants (treatment on the treated impact) require programme participation data, which might exist in one separate dataset, as well as employment and benefit spell data which might exist in another data set, and which can act as an outcome variable. These data may need to be merged on an individual level.

5   Using administrative data to inform CBF for duration of benefits and employment, and wages

To value many of the impacts of a programme, we first need an estimate of the total additional time in employment created by the programme (this is the employment that occurs as a result of the programme and would not occur if the programme did not exist). Second, an estimate of the earnings in that additional employment is needed (see Fujiwara, 2010).

In this chapter, empirical estimates of employment and benefit duration and earnings are limited to New Deal programmes. This could be extended in future research where linked programme participation/eligibility data and benefit/employment/earnings outcome data exist. Ideally we would estimate net impacts on duration of benefit receipt, employment, and wages. In practice, we estimate only **gross** outcomes. Net impacts are viable for future research but the identification of a suitable comparison group does need to be possible.

As there is no counterfactual, these figures do not show the programme impact.

5.1   Using administrative data for an analysis of economic outcomes

Much of the information necessary to estimate the costs and benefits of the Department for Work and Pensions' (DWP's) employment and training programmes can be retrieved from various administrative records including the eligible population for the programmes (working-age benefit claimants in the Master Index (MI) data), programme participants (participation databases), benefit outcomes (MI data), and employment outcomes (Her Majesty's Revenue and Customs (HMRC) data).

5.2   Identification of eligible populations

The **National Benefits Database** (NBD) includes those personal characteristics of benefit claimants necessary to **identify the eligible population for particular programmes**. It includes information about benefit payments and beginning and end dates. Since most programmes are offered to the working-age population on Jobseeker's Allowance (JSA) and Incapacity Benefit (IB) (or now Employment and Support Allowance (ESA)), the information from the beginning of the benefit claim can be used to identify those whose benefit claims surpass a given threshold and thus qualify for entry into programmes.

Evaluations of current or previous programmes can use benefit spells from the NBD to identify eligible clients. In some programmes eligibility is determined by a combination of spell information and specific personal conditions. Some examples of eligibility are: those lone parents starting an Income Support (IS) claim conditional on the age of the youngest child qualify for the New Deal for Lone Parents (NDLP); or those starting a new IB/ESA benefit claim due to long-term illness qualify for Pathways to Work (subject to selection based on the nature of the illness, and, due to the rollout since 2003, the individual's location).

In practice, only a fraction of the **eligible** population **participates** in programmes. For example, only those lone parents on benefit who attend their work-focused interviews, feel work-ready, and actively take up the offer of additional services are observed entering the New Deal for Lone Parents. Therefore, the NBD data can sometimes offer the opportunity to identify eligible participants, as well as sometimes offering similar groups of claimants who can be used as a comparison group in a non-experimental evaluation. In particular, programmes that have been piloted in specific areas (e.g. Pathways to Work) can be analysed by comparing eligible benefit claimants in pilot areas who began the programme with those in other areas who remained on the standard benefit.

The NBD has been used in various evaluation studies before, such as the evaluation of Pathways to Work by Bewley *et al.* (2006) and Work-Based Learning for Adults (WBLA) by Speckesser and Bewley (2006).

It should be noted, however, that the NBD has limitations, particularly with respect to available variables to use in analyses. Participants in voluntary programmes like WBLA are usually more work-ready and motivated to intensify their job search than eligible claimants who do not choose to participate. These differences driving participation also influence programme outcomes (i.e. the effect of participants' unobservable characteristics are misattributed to the programme, inflating the programme's success). The NBD data contain only a few variables that would allow one to control for these differences. These include: gender, age, location (postcodes), information on the youngest child and number of children for some benefits (both of which influence work readiness), a part-time work identifier for some of the benefits, and a benefit history (which may offer insights into people's labour market history and work-readiness)[24].

Thomas (2008) and Dolton, Azevedo and Smith (2006) indicate that a strong characterisation of the benefit history can suffice for matching and similar impact methods. Administrative data alone has been found to be sufficient to support robust propensity score matching in evaluation of benefit exit and job entry (see Dolton, Azavedo and Smith, 2006) although some argue that selection issues and unobservables which affect the evaluation validity are not sufficiently controlled for without additional data from surveys (and would require randomisation of programme participation). It seems likely that it might be sufficient to use administrative data, and Thomas (2008) gives further evidence of this. However, essentially the assumption of the Conditional Independence Assumption (CIA) holding for unobservables cannot be proven to be met, as the only evidence must be drawn from observables. As a result, only randomised programme participation can produce data in which the observables can be said to be satisfying the CIA. The administrative data is generally considered to provide the most accurate evidence on benefits, and so evaluation using the administrative data such as the Work and Pensions Longitudinal Survey (WPLS) combined with the MI data can provide accurate estimates of programme impacts on benefit exit. However, it should be recalled that the employment data in the WPLS, which is supplied from HMRC records, is acknowledged to be very imprecisely measured, and certainly has coverage issues – these all combine to result in this data

---

[24]   For the value of these benefit background variables in the New Deal for Lone Parents econometric re-analyses, see Dolton *et al.* (2006) and Knight *et al.* (2006). They have been shown in Dolton *et al.* (2006) to be sufficient to describe and satisfy the conditions for matching using the administrative data records.

producing **different measures** of employment impacts[25] to that found in survey measures. To improve employment estimates, individual survey data can help. This is not only because of the lack of coverage of the HMRC data for all employment, but also because important aspects of the employment such as hours of work are not present. As such, the NBD can be a reasonably powerful resource for evaluating pilot programmes' effects on benefits but would usually benefit from additional data to enable evaluations to produce the full set of impacts on a range of outcomes of interest beyond benefit exit.

Table B.1 summarises the available socio-economic indicators and benefit data which can be retrieved from the NBD. Most of these indicators can also be found in MI data, which are an extract of the NBD for the working-age population on particular benefits. The information is the same as for the NBD, but excludes pensioners' benefits.

5.3   Identification of benefit outcomes

Data on benefit histories is available in the NBD. As illustrated in Table B.2 beginning and end dates of benefit spells allow us to estimate benefit receipt following programme participation. Some assumptions must be made to use this data, for the following reasons:

- benefit spells have imputed end dates. End dates are imputed as a date between benefit register scans (between two to six week intervals depending on the benefit) and since imputation is random, and carried out for all cases, it is unlikely to result in any bias. The start date of the claim is known;

- some information, like the destination on leaving benefits, is not consistently included because it is currently only (partially) recorded for ending a JSA benefit spell.

Despite these two caveats, overall the NBD can be used to give reasonable estimates of programme outcomes including the benefit levels of participants, and net impacts (due to the comparison with non-participants[26]) and for some programmes based on JSA benefits, the destination on leaving benefit. One can use these estimates in cost benefit analyses (CBAs). Furthermore the estimates can be continuously updated using the NBD.

While descriptive evidence on gross outcomes from the NBD, such as those estimated here, can improve cost benefit estimates, net programme impacts would be more beneficial. However, estimating net impacts is more difficult since a comparison group must be identified. A comparison group can be non-participating benefit claimants with similar characteristics, or those not eligible for a programme due to living in a non-pilot area, for example.

---

[25]   The Pathways to Work evaluation projects, Bewley *et al.,* 2009a; Bewley *et al.,* 2009b; and Employment Retention and Advancement evaluation project reports, Dorset *et al.,* 2007, Riccio *et al.,* 2008, had both survey and administrative data-based impacts, and there were reasonably smaller employment estimates from the administrative data which, aside from non-response to surveys, was attributed largely to the fact that the employment in the HMRC records is limited to that collected for National Insurance and tax reasons, and so only covers those eligible to pay these, and also is only recorded intermittently when employers send forms, often with long lags. The issue is not entirely resolved.

[26]   However, the comparison with non-participants needs to be identified as suitable and similar to the programme group.

In estimating and reporting net impacts, it is important to explain the reasons for choosing a comparison group, in order to help establish the validity of this group as suitable and relevant for the evaluation design.

It is recommended that in the establishment of a new policy, careful design for a pilot should take place and include a design for the evaluation which enables identification of a net impact. In the absence of a pilot, included in the design of the policy should be an evaluation design which enables net impacts and consideration of value for money for the policy programme to be established. In discussing a planned policy, an explanation of the considerations made in this design should be made, and such an explanation is also needed in the case where a decision is made not to establish an evaluation design for a policy, together with discussion of the implications this has for identifying the net impacts and value for money for the policy.

5.4   Identification of wage and employment outcomes

The employment data used for this analysis are taken from HMRC tax records for earnings from employment. Data documentation reports that the data were collected January 2004 through June 2009 although some spells **begin** as early as the 1970s. For most programmes, of the individuals observed in various programme databases, approximately 20 per cent of the cases had no employment records (among New Deal for individuals aged 25 plus (ND25+) participants 19 per cent had no employment records, 19 per cent in New Deal for individuals aged 50 plus (ND50+), 21 per cent in NDLP, 20 per cent in New Deal for Disabled People (NDDP), 28 per cent in New Deal for Partners (NDP), and 22 per cent in New Deal for Young People (NDYP)). For cross tabulations included in this report, we assume that those **without** employment records were **not** working[27].

Approximately 0.9 per cent of those individuals (on the programmes analysed in this report over the calendar time period considered) with employment records had an unusually high level of spells with more than ten recorded spells **after recoding** overlapping and duplicated spells. Any spells beyond the tenth spell were ignored.

Earnings data were taken from the P14 file for May 2009. These data included total taxable pay for 2005-2009 tax years. The final year of data, 2009, reports annual earnings, assuming current income streams are maintained, as the records are not necessarily complete for that period yet. Earnings data could not be matched to employment spells, as earnings are reported for taxation purposes in an annual format and are aggregate across jobs. As with employment data, earnings were retained only for programme participants for these analyses. Fewer participants are recorded in the earnings data compared to in the employment data: overall, regardless of employment records, 44 per cent of the ND25+ participants had no earning/income records, 29 per cent of ND50+, 34 per cent of NDLP, 25 per cent of NDDP, 44 per cent NDP, and 25 per cent of NDYP participants.

Table 5.1 illustrates the distribution of missing labour market records for the New Deal programmes. Many cases have employment data without earnings data, but also a few cases have earnings data without employment data. This arises from the separate recording of earnings and employment in the databases, drawn from different administrative tax forms P45 and P14. As illustrated in Table 5.1, in the ND25+ programme among those with earnings records (column) only one per cent do not have employment records, while among those with employment records (row) 55 per cent

---

[27]   The HMRC employment data coverage is not complete and has known deficiencies which mean that this is not precisely true, as those with employment that does need payment of National Insurance or tax is not included, nor is self-employment, for example, while some of the information is also poor quality with regard to start and end date of the employment.

have earnings records. Note that the data used for this analysis include the interpolations and assumptions outlined in the next section – the raw data would have had more missing records.

Table 5.1   Missing labour market data among programme participants

		Earnings records				Earnings records	
ND25+		Exist	Do not exist	ND50+		Exist	Do not exist
Employment	Exists	55%	26%	Employment	Exist	69%	12%
	Do not exist	1%	19%		Do not exist	2%	17%
		Earnings records				Earnings records	
ND25+		Exist	Do not exist	ND50+		Exist	Do not exist
Employment	Exists	66%	13%	Employment	Exist	72%	8%
	Do not exist	0%	21%		Do not exist	3%	17%
		Earnings records				Earnings records	
ND25+		Exist	Do not exist	ND50+		Exist	Do not exist
Employment	Exists	61%	15%	Employment	Exist	61%	15%
	Do not exist	0%	23%		Do not exist	0%	23%

Notes: For example, of individuals observed in ND25+, 54.75 per cent are observed to have at least one HMRC employment record, with at least one earnings record. Appendix B lists variables and databases sourced for these analyses, while Appendix D lists the DWP database information resources. HMRC employment data January 2004–June 2009 (spells can start before 2004).

5.5   Identification of programme participation

A variety of programmes have separate databases with more detailed information on programme participation and some individual characteristics of programme participants[28]. These were defined at the outset of the programme for evaluation purposes, drawing specific variables from the Labour Market System (LMS) and sometimes creating new LMS variables, which were extracted into this database[29].

---

[28]   Individual characteristics are collected differently for different programmes and coverage (the level of missing information) can be very high.

[29]   Such programme databases are a management information tool and can be devised but can also be limited by the time required to plan and setup such databases and the strong limitations on alterations to the LMS data-capture system. Such data are not always well recorded due to the need for caseworkers to enter the data which may not always be prioritised.

While NBD/MI data allows one to identify a programme's eligible population, these programme-specific datasets show only the actual programme participants. In these datasets there are different indicators for when an individual's participation on a programme begins. This is illustrated in Table 5.2. In most programmes the variable **cdpstart** corresponds to the earliest action taken on the programme – usually the date the participant's benefit data details were uploaded to the caseworker's system. Due to the differences in the programme designs of the various New Deals, some programmes do not include **cdpstart** or **cdpstart** indicates some other event, such as the beginning of the Gateway[30] (gwstdt).

Table 5.2 indicates the different start date variables that could be used. While the earliest start does not fully indicate active programme participation, the simple initiation of a case can also generate an employment outcome. For example if the caseworker contacts the benefit claimant with an invitation for a work-focused interview for a New Deal programme, the beneficiary may take a job or end their benefit claim rather than attend the meeting. Most programme effects will arise from actual participation, which can be measured using other start dates such as those indicating that the individual attended an initial interview, started the gateway phase of the programme, or began an option[31]. A priori it cannot be answered whether a broader or narrower definition of programme participation is more appropriate. Therefore we present results using multiple start dates. All empirical analysis using these data should carefully consider and define the start and end dates of programme participation that they use.

Table 5.2 shows the three start dates that were used for each programme. For example, for ND25+ we used cdpstart (the file upload date), gwstdt (when the client began 'gateway'), and nd1intdt (when the individual attended their first interview). The last row in Table 5.2 indicates which of these three variables we would suggest using for each given programme as the 'preferred' start date. Unfortunately, there is little information on the reliability of the recording of these programme data sourced from the LMS, nor on the relative reliability, definition or coverage of each variable.

---

[30] See footnote 30 on the design of the New Deals: 'Gateway' was a period of caseworker services. Hence this is the 'caseload' date, at which this case became part of a caseworker's workload

[31] The New Deal programmes were designed with an initial period (of varying length depending on the specific New Deal programme group), called a Gateway, with caseworker services from Jobcentre Plus, after which those still remaining could enter an 'option' which was a more specific programme service, such as the employment option, which was a wage subsidy, a voluntary option with links to experience in the voluntary sector, training, and the environmental taskforce option. For more information on the design of specific New Deal programmes, consult the DWP Research Report series which has reports from the programme evaluations for each of these programmes.

Table 5.2   Programme start variables

Description	New variable	ND25+ (file: Ltu_0906_ id)	ND50+ (file: N50_0907_ spells_idt)	NDLP (file: Lp_0906_ id)	NDDP (file: Ndstats_ dp_0905_ id)	NDP (file:Ndstats_ npp_0905_id)	NDYP (file: Ndstats_ yp_0905_id)
Loosest definition	start1	cdpstart	cdpstart	cdpstart	contstdt	sfrstart	cpdstart
Medium definition	start2	gwstdt	casestart	firintdt	ibstd	ndpintat	gwstd
Strictest definition	start3	nd1intdt	initint	stcasedt	cdpstart	caselddt	opst1
Preferred start	ndstart	nd1intdt	initint	cdpstart	cdpstart	caselddt	cpdstart

Notes: Appendix B lists variables and databases sourced for these analyses, while Appendix D lists the DWP database information resources. Appendix D lists the database information sources – information about programme databases is available internally to DWP but in practice, information on content of variables is often limited to the labels on the variables/anecdotal/experience.

5.6   Data processing

In impact evaluations published for DWP, data processing decisions tend to vary by evaluation and evaluations vary in the extent to which they report key dataset construction decisions. These decisions can influence the impacts estimated. Given this, in the future, evaluations should carefully report their data processing decisions. Ideally, the best identified methods should be replicated across evaluations. Scioch (2010) explores the importance of this for German evaluations using administrative data.

We recommend that important processing of overlapping spells and inconsistent information in the benefit and employment administrative data used for evaluations should be more standardised by DWP. This could be done by developing basic approved syntax for this, together with a documentary description of the key rules recommended, and supplying this with the data. Variations to this may be needed, and individual evaluations should document these processing decisions. All evaluations should provide information about the key processing rules in appendices. This would mean that replications and critical review are possible.

We recommend that important processing of overlapping spells and inconsistent information in the benefit and employment administrative data used for evaluations should be more standardised by DWP. This should be done by providing a basic approved syntax for this and a documentary description of the key rules recommended. The original data should still be provided together with this, as this then facilitates further understanding or improvements to be made to the processing over time. Data processing can influence the impacts estimated.

All evaluations should provide information about the data processing rules in their reporting.

5.6.1   Standardising socio-economic information and programme options

Most descriptive variables in the New Deal participation data did not need further processing; variables relating to gender, marital status, disability, postcode, rural residence, number of children, and date of birth of youngest child required no recoding. We standardised some of the descriptive

variables available on New Deal participation data to prepare the data for future multivariate analyses across programmes. There was a variety of information: all programmes had data on gender, start date, date of birth, ethnicity, disability status, postcode, end date, and the reason for leaving the programme (this particular variable **exists for all** but has varying degrees of missing data, sometimes very high, and so was considered unusable for these analyses). Some programmes also had data for marital status, the date of birth of youngest child, and programme options (e.g. basic skills training).

For example, to standardise, ethnicity was recoded into five categories: white, black, Asian, mixed, and missing. Most (but not all) of the New Deal programmes used the **srf_eth2** variable which included nine racial categories. We made this information comparable to other programmes reporting only four racial categories, by merging together 'black,' 'black- "Caribbean",' 'black-African,' and 'black-Other' into one category for black clients and by merging Chinese, Bangladeshi, Indian, and Pakistani into one Asian category.

We attempted to include educational level for NDLP and NDYP as the original raw data reported having this information, but in some data sets all observations were missing entries for this variable.

Coding programme options

Coding programme options was the most difficult part of standardising the data. We created five categories: subsidised employment/volunteer work, basic skills training, higher skills training, self-employment assistance, and job readiness. Appendix B lists variables and descriptions and databases sourced for these analyses. See Appendix D for more information on the data items and data sets available at the time of research (supplied November 2008).

Subsidised employment/volunteer work includes:

- ND25+: the Employment Option, Work Placement, Work Experience, and Subsidised Employment (SE);

- NDLP: Worktrial, NDLP job (on other caseload), NDLP in work (on IS), NDLP VS, and NDLP VS Familiarisation;

- NDDP: payptjb, payptse, payftjb, payftse, payptft, pyptftse, pyptsus, pyptsesu, pyftsus, and pyftsesu, all of which indicated a payment combined with some type of employment;

- NDP: no recoding was necessary as F_WEX_ST already indicated subsidised employment/volunteer work;

- NDYP: the employment option, VS, and ETF.

Basic skills training includes:

- ND25+: BET/BS option;

- NDLP: basic skills, other training, WBLA, NDLP B-Skills, NDLP IAP Education and Training, LSC BS Training, and LSC ESOL Start;

- for the NDDP, NDP, and NDYP: no variables indicated a basic skills option.

Higher/occupational skills training (labelled hard skills in the graphs) includes:

- ND25+: FTE, WBLA, ETO, and Training for Work;

- NDLP: Other Education/Training, SJFT, AM Demand Led Training, European Social Fund (ESF) Provision, ESF ILM, Nat/Prog Dev Fund, FTE, FTE training familiarisation, Careers, LSC Occ Training, and LSC Mod App Training;

- NDP and NDYP required no recoding.

Self employment:

- New Deal 25: general self employment option as well as the musicians-option;

- NDLP: Business Advice Agencies, Self Employed Stages 1, 2, and 3, ND Music NDLP MIC, and ND Music NDLP MOLP;

- NDP and NDYP: no self employment option was coded.

Job-readiness option:

- ND25+: IAP training, Soft Skills/Short Motivation, and job search;

- NDLP: (mandatory) job search provision, job club, job interview guarantee (JIG), job search seminar, job fairs, progress to work, Environmental Task Force (ETF) (familiarisation), job search and advice (J&A), Gateway (GW) training (KS, IT, Motivation, or Voc), and Learning and Skills Council (LSC) Information Advice and Guidance (IAG);

- NDP: first start/receiving advice and guidance;

- NDYP: no job readiness option;

- ND50+: There was no available information on options for those in ND50+.

We also standardised the variables indicating **why** individuals left the programme. These codes were simplified to indicate whether an individual left the programme for employment, another benefit, the same benefit, or another reason[32]. For some programmes the original data sometimes specified whether the individual left the programme for full-time or part-time work, self employment, or was no longer eligible.[33]

Overall, the data were rather limited with regards to socio-demographic covariates. The final dataset included: gender, age, ethnicity, rural, disability status, and **for some individuals**, information about the children.

5.6.2   Correction of earnings and employment data from HMRC

Overlapping spells in HMRC data on earnings/employment

The original HMRC employment data included a new spell for each time any information about an individual changed – for example their home address. These spells had to be merged into single spells for these analyses. In addition, since we are generally interested in whether an individual is employed, not whether they maintain the same job, it was assumed that consecutive job spells are only one employment spell.

If one spell lasted exactly the same time period as another or was entirely embedded in the other, the smaller of the two spells was ignored. Two spells were merged together to make one longer spell if the spells overlapped or were within seven days of each other[34]. If there was a gap exceeding seven days between any two employment spells, they were considered independent spells. It is worth noting that the process of merging does not only look at pairs of subsequent spells in

---

[32]   Other reasons usually relate to inactive or jobsearch states but not on the benefit.

[33]   While this variable was not used in the current analysis, we recoded it with the intent of using it in future multivariate analyses.

[34]   An arbitrary number which was selected as reasonable for a change of employment status notification. All data processing rules such as this can affect the results found. This is why we suggest that all evaluations ensure the processing rules are made clear.

the data, but rather compares information from the **recoded** spell in the prior period to the next **unmodified** spell. This means that the following series of spells would be coded as a single spell as shown in Figure 5.1.

Figure 5.1   A series of overlapping spells that would be coded as a single spell
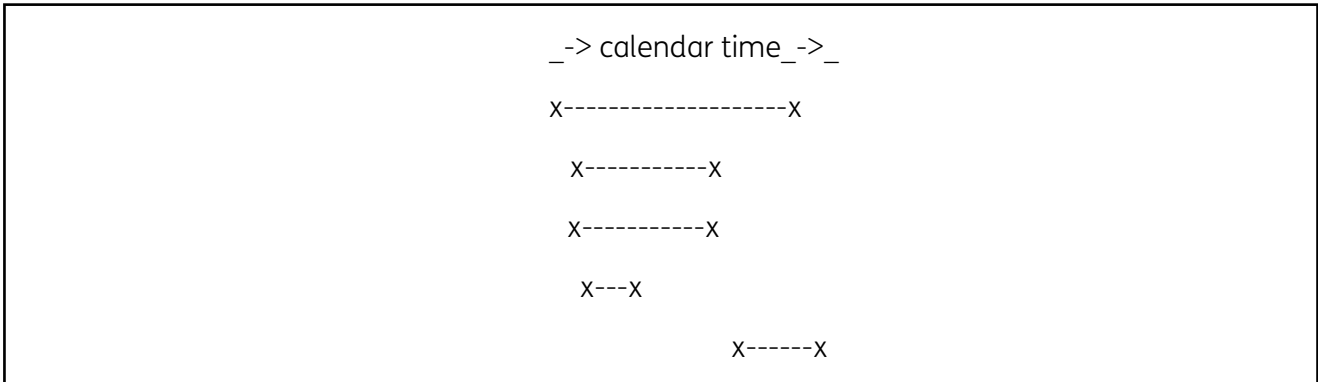
Once the number of spells was reduced, we had to determine how many spells to include in the data to be merged into programme data. To determine the number of spells that were included we first ran the data allowing for 20 spells, finding the distribution of spells among individuals who had at least one employment spell. Figure 5.2 shows the distribution – 99.1 per cent of the people had ten or fewer spells while 99.99 per cent had 18 or fewer spells. We recoded the data in the manner outlined in Figure 5.1 to include a maximum of ten spells before merging employment records to programme data. The small number with such high spell numbers indicates that only a very small outlier group of individuals, potentially highly unusual, are affected by this. This means for these cases, the data reflects employment over the same periods but with fewer changes of employment.

Figure 5.2   Per cent of cases with x employment spells
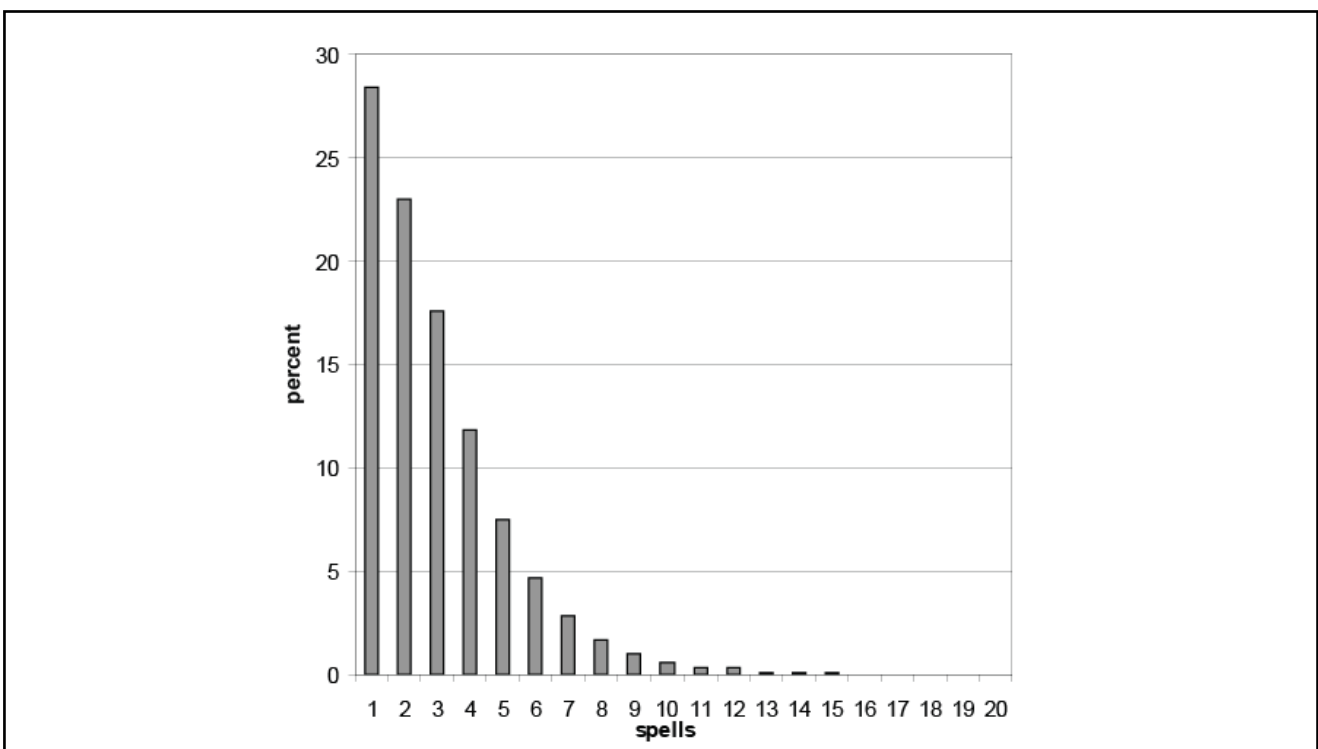
Note: HMRC employment data January 2004 –June 2009 (spells can start before 2004).

5.6.3   Imputation of start/end dates for employment spells

Some employment spells had missing start and end dates, or sometimes start dates were coded as 'Dec 31 9999,' a value usually entered to code for the **end date** for the ongoing spell. Presumably, the person entering the raw information into the system simply mistyped this date into the start date field rather than the end date field. If a spell was missing both the start and end date or if the spell was missing the start date and the end date was Dec 31, 9999, then the missing dates could not be interpolated and the spell had to be dropped[35].

However, **if only a valid start or only a valid end date was available**, then the other date could be interpolated using information about employment spell durations. Often the year of the end/start of the employment spell was available, even if the exact date was not. This meant that when interpolating employment durations, we had to be sure to interpolate using information about average spell durations as well as applying a date falling in the appropriate year.

To interpolate employment spells we use data from the complete employment spells available in the data and fitted the distribution of employment spell duration. We specified a gamma distribution and obtained the best fitting estimates for the parameters of this distribution.

$$pdf = \frac{1}{(\beta^{\alpha}\Gamma\alpha)} t^{\alpha-1} e^{-\frac{t}{\beta}}$$

a = .803

$\frac{1}{\beta}$ *= 1026*

This distribution was then used to pull a **random** start or end date for those spells with missing start or end dates. If the distribution picked a start or end date falling outside of the known year of the missing start or end date, the imputation code pulled another number from the distribution, and continued to do so until the interpolated date fell into the appropriate year.

Figure 5.3 compares the duration of employment spells on the x axis to the per centile in the distribution of employment spells. For example, 800 days is about the 50th per centile spell duration. The dashed line indicates spells with fully observed information, while the solid line indicates spells that we interpolated. The similarity between the two lines shows how closely the observed and the interpolated data now match each other.

---

[35]   Note that due to the format of this data analysis, employment spells were analysed as entirely independent of benefit spells. However, you can use other data to inform imputation: for example, if that case has a benefit spell at the start of the tax year, the employment spell start date could be imputed to fall on a date after the benefit exit spell.

Figure 5.3   Comparing interpolated data with complete data

Note: HMRC employment data January 2004 –June 2009 (spells can start before 2004).
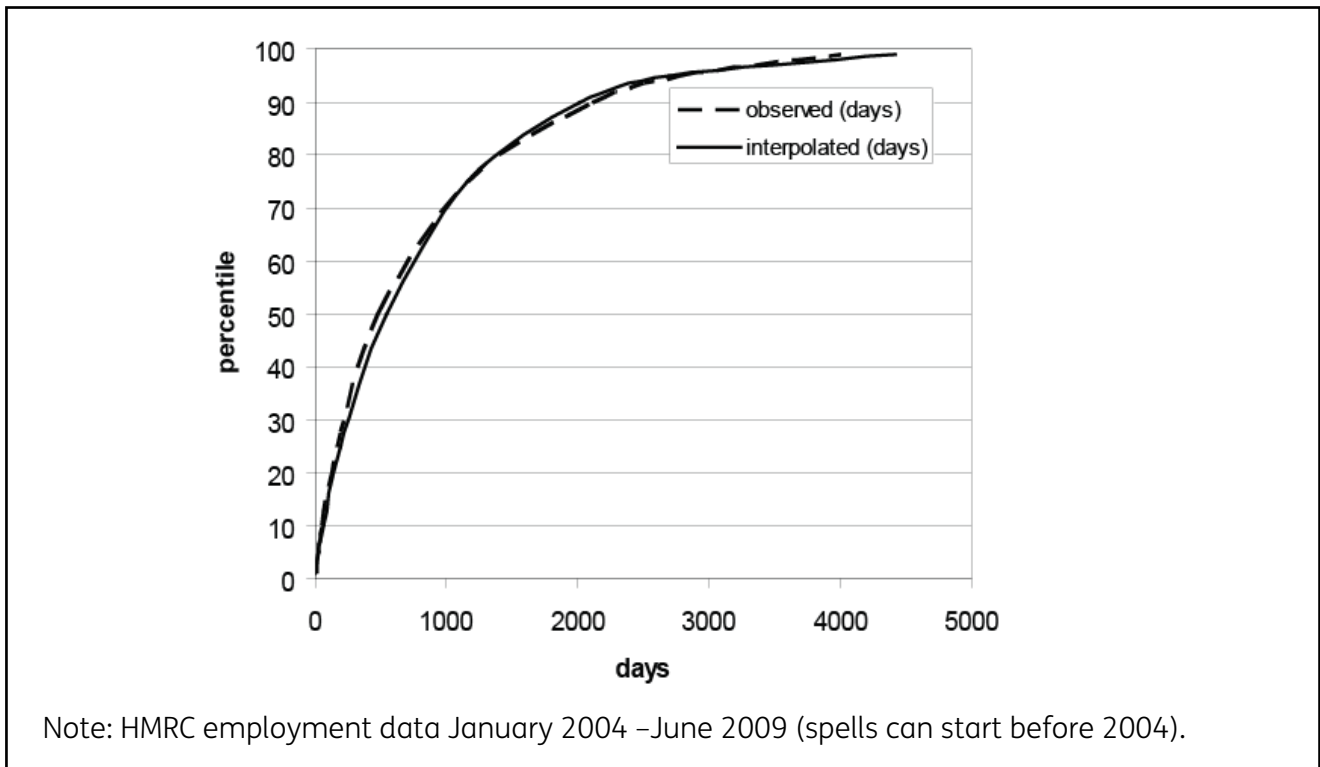
Figure 5.3 also shows the validity of the interpolated spell information by comparing the interpolation to observed spell information.

Finally, employment spell data were merged onto programme data so that we had employment data for participants both pre- and post-programme[36]. The number of spells per participant varied across programmes. Looking at Figure 5.4, a greater share of participants in the 50+ programmes had only one spell of employment.

---

[36]   The current analysis uses only post-programme employment, but the data were compiled with the anticipation that future analyses would wish to control for participants' work histories.

Figure 5.4   Per cent of programme participants with up to 10 employment spells

Note: HMRC employment data January 2004 –June 2009 (spells can start before 2004). Post-programme employment spells.
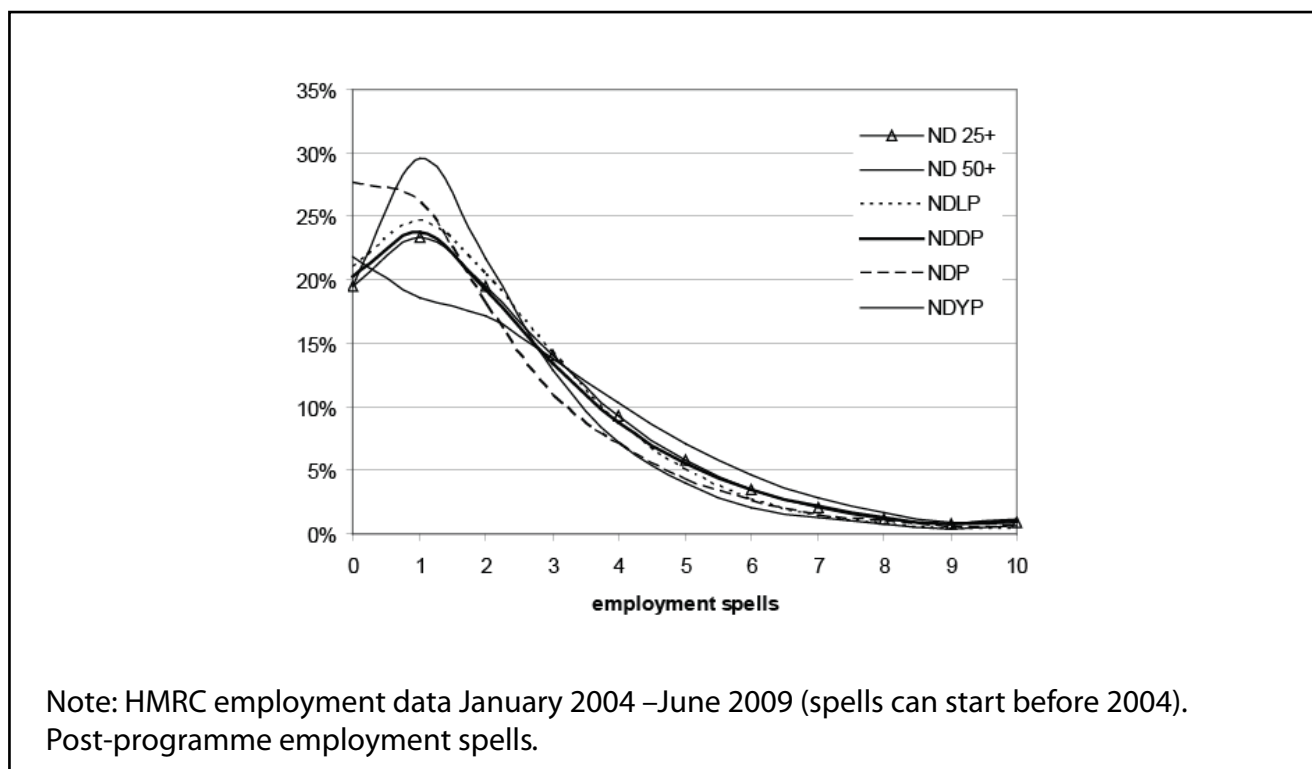
5.6.4   Earnings

The original HMRC P14 file reported taxable earnings with one entry per job per tax year. Since there is unreliable information on the exact start and stop dates for each employment spell and because income is not spell-based but year-based, we cannot match earnings with individual employment spells. Instead, we merge all jobs across each year to create a variable measuring total taxable income across all jobs in a year. We then transposed the data for each individual, so that each record had one variable for taxable income for each fiscal tax year (2005-2009, April-March), for example total taxable earnings recorded for the tax year 2005. Spells of taxable benefit receipt were removed using the flag available in the data.

As described in Section 5.4, earnings were missing for about 58 per cent of the observations with employment spells while employment spell data were rarely missing for those with earnings data. Earnings data were unlikely to be missing at random – those with missing data are likely to have incomes falling below the tax threshold or there was some HMRC reporting issue for the employer. Given the bias of reported information and the preponderance of missing data we decided not to interpolate earnings for those individuals that had employment spells but no earnings. This means that all reported earnings information is limited to those who were employed, for whom the employer has reported tax record information to HMRC and who earned enough to be included in the tax records. As such, the earnings estimates are based on a non-randomly selected sample of the population of people under consideration[37].

---

[37]   This is important when considering analyses of the data and how to construct these validly.

5.7   Benefit status information

5.7.1   Overlapping spells in Master Index data

More than one benefit can be paid simultaneously, meaning that there are often overlapping benefit spells in the data. While the distinction between some benefits is important (e.g. the JSA benefit requires participants to engage in structured job-search activities[38]), in this application we are interested in whether an individual claims **any** benefit in a specified period. In addition, some records in the benefit register may completely overlap with other records or result from changes in circumstances (e.g. address).

Figure 5.5 illustrates how multiple spells can be found for one person in the MI data. This hypothetical case shows that the person has IS and IB claims, which partially overlap. At a later time, the same person has multiple spells of JSA benefits. Benefit data was recoded into daily status variables indicating whether a person was on benefit on each day between 2004 and 2009. For the five years analysed this results in more than 1,400 indicator variables of benefit receipt, retaining information of daily precision, but losing information about which benefit was paid on that particular day. Retaining data for all benefits for the working-age population (IS, incapacity benefits, Severe Disability Allowance, Invalid Care Allowance and JSA) would have required more computational resources. Setting up the data in this manner allows us to estimate the percentage of time spent on benefit or the benefit rate of participants on any day between 2004 and 2009.

Figure 5.5   Recoding benefit data
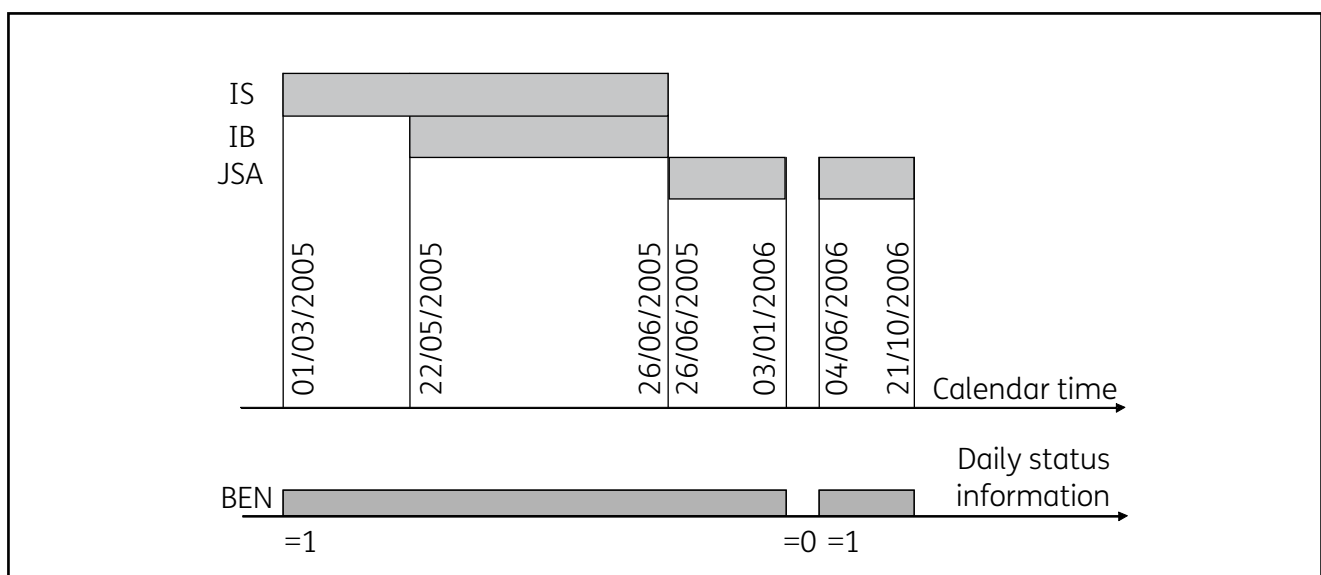
5.8   Participation and outcomes

We analyse the benefit, employment and earnings outcomes for **participants who started on the various New Deals in the financial year 2005-06.** For benefit outcomes, the length of time between the start and the end date of a participant's spell is the duration of interest. Therefore the calendar dates covering a spell depend on when in the financial year the spell began as well as the spell's duration.

---

[38]   The distinction is also important because IB may be paid at a higher rate than JSA, so if a programme causes people to move from JSA to IB this affects the net fiscal benefits.

In this analysis, New Deal participation start dates range from 6 April 2005 to 5 April 2006 and spells are observed for a maximum of 24 months after their start date. Given this, the latest possible observed benefit (or employment) outcome for a start occurs in April 2008. Figure 5.6 illustrates this and shows how benefit outcomes and starting dates of the programme correspond. In the empirical work we consider outcomes at three, six, 12, 18, and 24 months following participation.

Figure 5.6   Moving observation window for programme participation with benefit outcome periods
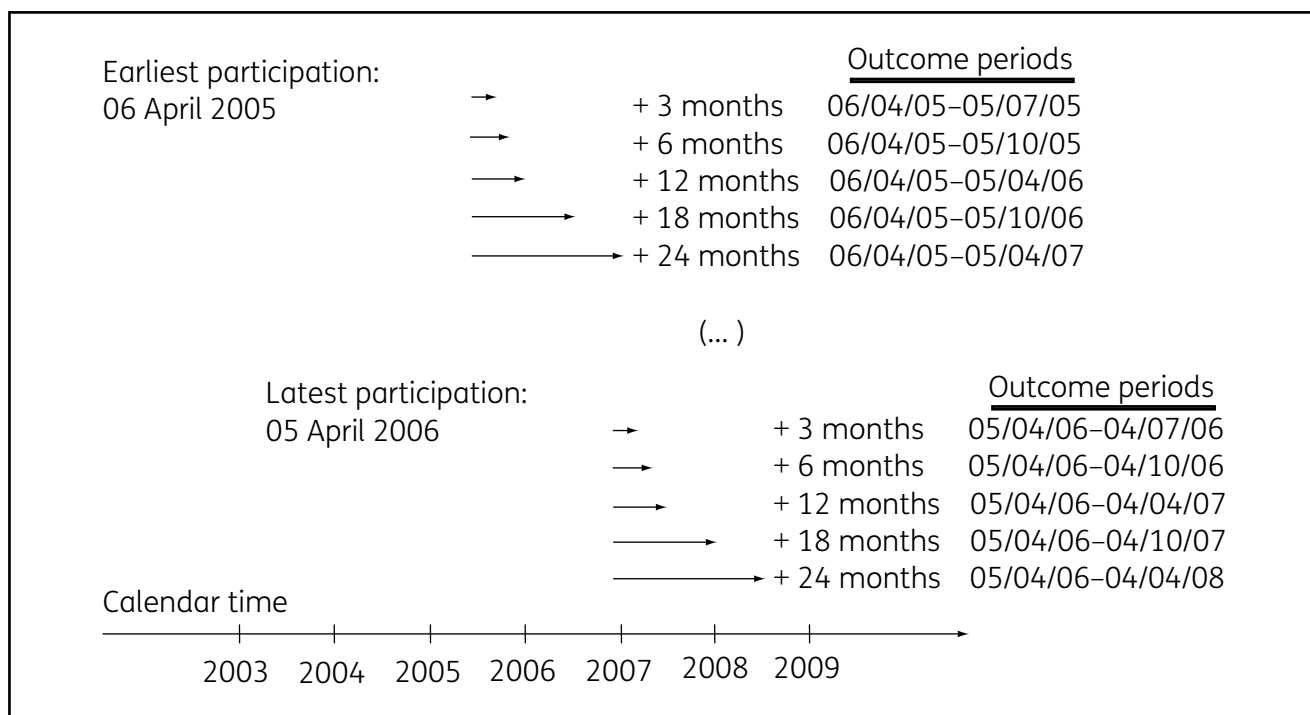
The moving observation window allows an exact alignment of outcome to particular dates of participation. This is analytically preferable compared to outcomes that are related to particular calendar time dates.[39]

> In this report, we analyse the outcomes for participants only. Although this is not causal evidence, it at least informs assumptions on duration of benefit or the length of employment after the programme participation using empirically observed post-programme benefit and employment rates. We recommend using these estimates for CBAs where needed.

5.9   Descriptive analysis: duration of benefits, employment, and earnings

5.9.1   Extraction of participants' cohorts

As discussed in Section 5.5, there are often multiple different types of start dates recorded for participants. This begins with the uploading of participant data, continuing with the first contact by

---

[39]   If a comparison group can be created for net impacts then outcome information can be created for non-participants, so that net impacts on outcomes can be observed for particular post-programme time intervals.

a New Deal adviser, the beginning of the Gateway, or beginning the first option. In practice, with the exception of the NDDP, our results did not find much variation by the start dates we explored.[40]

The measured outcomes (employment, earnings and benefits three, six, nine, 12, 18 and 24 months after participation) depend on the chosen start date definitions. For example, expectations are that the proportion of spells still on benefits will be higher three months after the earliest definition of the starting date compared to three months after the beginning of the gateway of a New Deal programme. This is because there is little active programme participation in the periods immediately after the earliest starting date definitions.

A small number of participants were recorded as having started the same programme twice in 2005-06. We therefore restricted the analysis to outcomes related to the first participation[41].

5.9.2   Gross benefit rates estimates

In this section, we present estimates of benefit claim rates for the programmes under consideration. Estimates are also presented by the programme option categories as set out in Section 5.6.1.

The observed benefit rates for the various programmes are shown in tables contained in Appendix A and are briefly discussed here. Figures 5.7 and 5.8 illustrate the benefit receipt ratios by time since programme start for four selected programmes.

The main findings are:

• About half of the ND25+ participants (who entered in 2005-06) are on benefit 12 months after starting the programme. After two years, the benefit rate is still around 42 per cent (Tables A.1 and A.2). These results are robust to using different definitions of programme start or considering the total per cent of time being spent on benefit for all participants[42].

• For ND50+, one year after programme start, 43 per cent of participants are on benefits. About one third are on benefit after two years[43]. Looking at the time spent on benefit as a percentage of the total time since beginning the programme, more than 50 per cent of the first 12 months was spent on benefit; 45 per cent of time was spent on benefit in the first two years after participation (Tables A.6 and A.7).

• Voluntary programmes, like the NDLP, with entry routes through work-focused interviews or a customer self-referral show more favourable benefit outcomes. This is not surprising as it is expected that more work-ready claimants will begin the programme relative to the compulsory programmes. For NDLP about one third of all the participants are on benefit 12 months after beginning the programme. Over the first year, participants spend up to 50 per cent of the time on benefit. (Tables A.11 and A.12).

---

[40]   For NDDP, using the IB start date variable significantly decreased the observable employment and made the proportion observed in subsidised employment higher than the average for all programmes. It also significantly increased the estimated benefit receipt rates. This possibly relates to the longer benefit (IB) spell lengths for NDDP volunteers relative to programmes based on other benefits.

[41]   This is the simplest best analytical practice where there is multiple participation.

[42]   Intensive Activity Period' (IAP) participation is classified under 'Job-readiness' – see earlier in Section 5.6.1. IAP is variety of assistance (training, work experience etc.) lasting 13 to 52 weeks; it is the second stage of assistance for ND25+, and is voluntary for ND50+.

[43]   The current analysis does not include pension benefits, only working-age benefits.

- In the NDDP (another voluntary programme) benefit rates fall from 47 per cent at three months to 25 per cent after 18 months. About 34 per cent of the first 24 months is spent on benefit (Tables A.16 and A.17). The preferred start date for NDDP is cdpstart (see Table 5.2).

- In the NDP programme, the partners are usually included in the main claimant's benefit claim and have no separate benefit record, so the benefit rates for the 'partners' initially appear low. Over time benefit rates for the partners increase to about 12 per cent after two years, as some participants initiate independent claims (e.g. JSA) (Tables A.21 and A.22).

- Among the mandatory programmes, the NDYP shows the lowest post-participation benefit rates. Twelve months after the start of the programme, about one third of the participants still claim benefits and participants spend about 38 per cent of their first year on benefit. Around 27 per cent of claimants are still on benefit after two years. Due to a few long-term claimants, the average per cent of time spent on benefit for all participants within the first 24 months is rather high at 37 per cent. (Tables A.26 and A.27).

Figure 5.7   Snapshot of benefit receipt for selected New Deal programmes

Note: Participants who started on the various New Deals in the financial year 2005-06. 'hard skills option'=higher/occupational skills training, defined in Section 5.6.1.
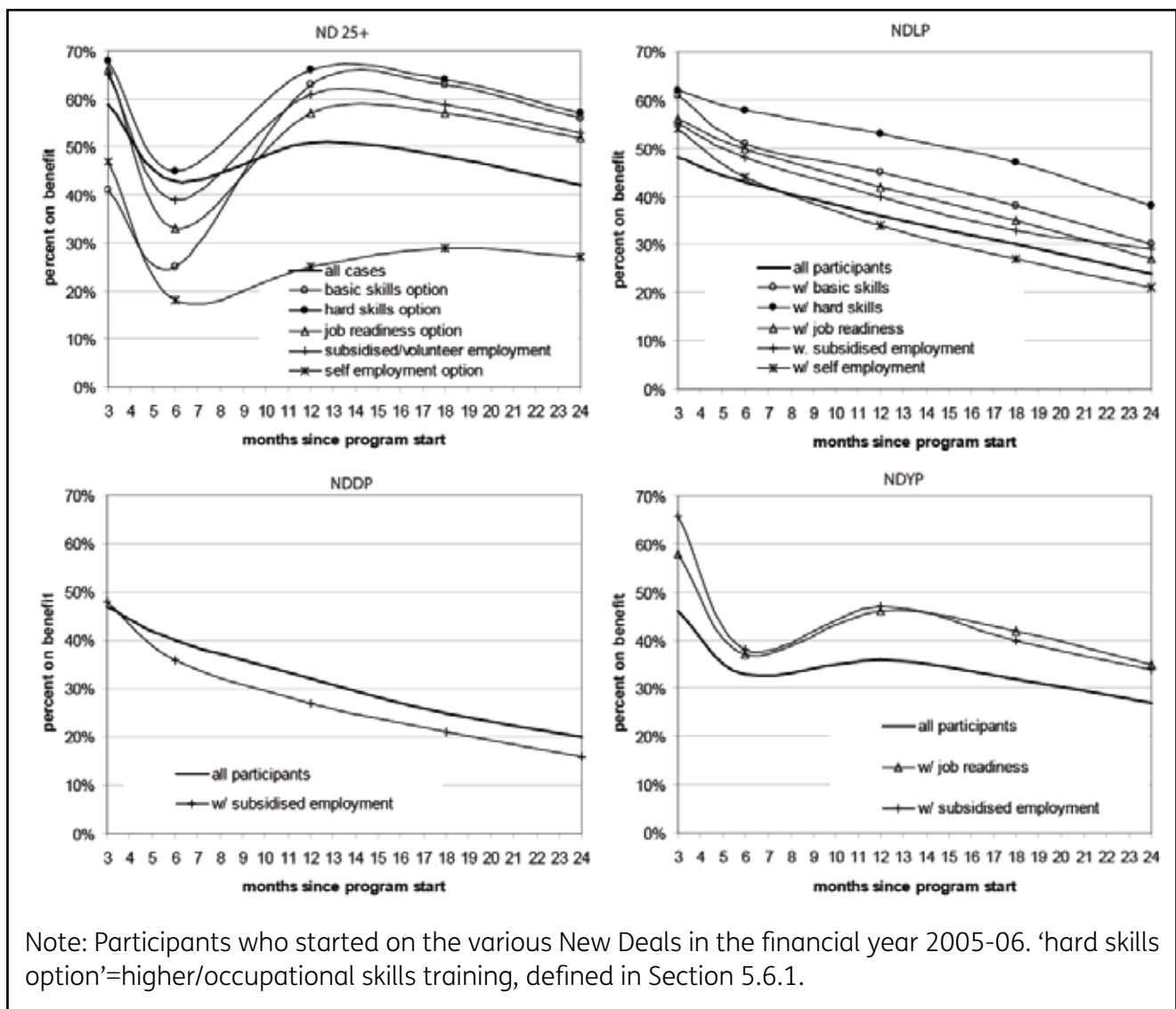
Figure 5.8 Percentage of time spent on benefit for selected New Deal Programmes up to 24 months after participation start
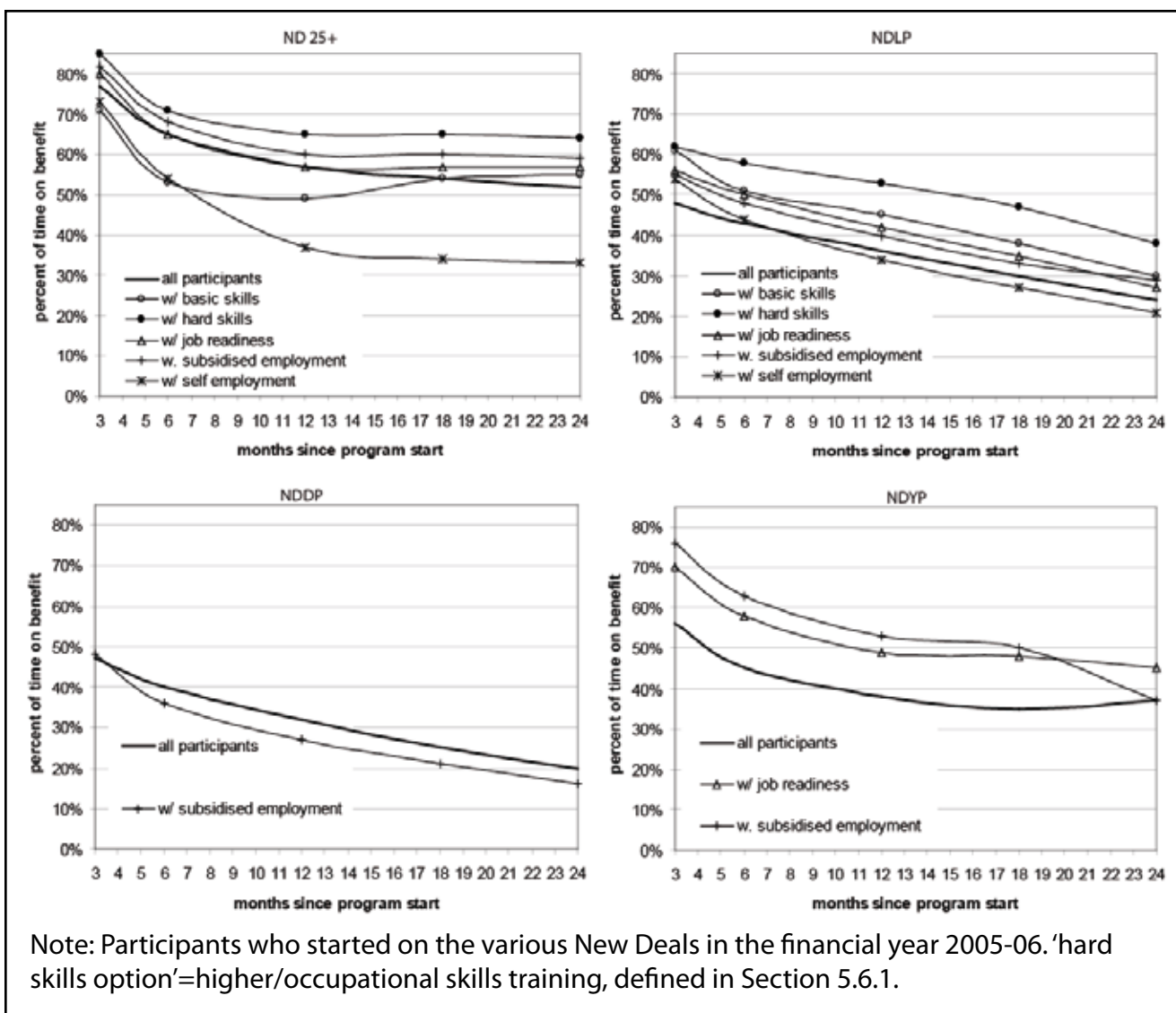
Note: Participants who started on the various New Deals in the financial year 2005-06. 'hard skills option'=higher/occupational skills training, defined in Section 5.6.1.

5.9.3 Gross employment durations

In this section, we present estimates of employment entry for the various programmes under consideration. Estimates are also presented for each programme by the programme option categories as set out in Section 5.6.1.

As with benefits, there are two ways to calculate employment outcomes. The first is to consider the percentage of people who are in employment at given durations after starting the programme. The second is to consider the average percentage of time starts are observed to spend in employment over a given duration window beginning upon programme start date.

For these estimates, we make the assumption that those individuals without employment records were not employed[44]. We also make assumptions relating to the most appropriate start date to specify for a programme (see Section 5.5). We calculated the cross-tabulations using all three start dates, but here we only report our 'preferred' programme start date. For most programmes, with the exception of NDDP, using a different start date did not significantly alter the results.

Figure 5.9, illustrates the per cent of programme participants in employment for a given number of months after their programme start date. For the ND25+ programme, participants have an initial jump in employment followed by a slow steady increase in employment. Those in basic skills training have the worst employment outcomes. This could be (partially at least) due to selection effects, where the least skilled are selected into this option.

The relationship between per cent employed and time since programme start is similar for the NDLP. However, overall employment rates are higher. It is also noticeable that among lone parents those going through 'Higher/Occupational skills training' (labelled 'hard skills') eventually reach employment levels of 40 per cent, which is only slightly surpassed by the 'Subsidised/volunteer employment option'.

Interestingly, NDDP shows the highest employment levels. ND50+ is not illustrated in Figure 5.9. However, the relationship between employment rates and months since programme start looks similar for this programme, with 25 per cent of claimants working three months after the start of the programme and about 30 per cent working two years after the start (Table A.8). The NDP is not presented because the employment results are linked using an ID for the participant's partner (i.e. the main claimant) though these main claimants increase employment from about 36 to 42 per cent employment over the same period (Table A.23).

---

[44]   It is worth noting that for a small number of cases (see Section 5.4) employment data are missing but wage data exist, indicating that this is sometimes a false assumption. However, since employment calculations are made for specific time periods within the year and the wage data only exists for the total year, one is unable to determine whether the individuals with wage data but no employment data were employed in the **specific** period of interest. Given the indeterminacy of the dates as well as how infrequently an individual has wage but not employment information, these individuals are assumed to be not employed.

**Figure 5.9   The per cent of clients employed up to 24 months following
programme start**

Note: HMRC employment data January 2004 –June 2009 (spells can start before 2004).
Post-programme employment spells. Participants who started on the various New Deals in
the financial year 2005-06; 'hard skills option'=higher/occupational skills training, defined
in Section 5.6.1.
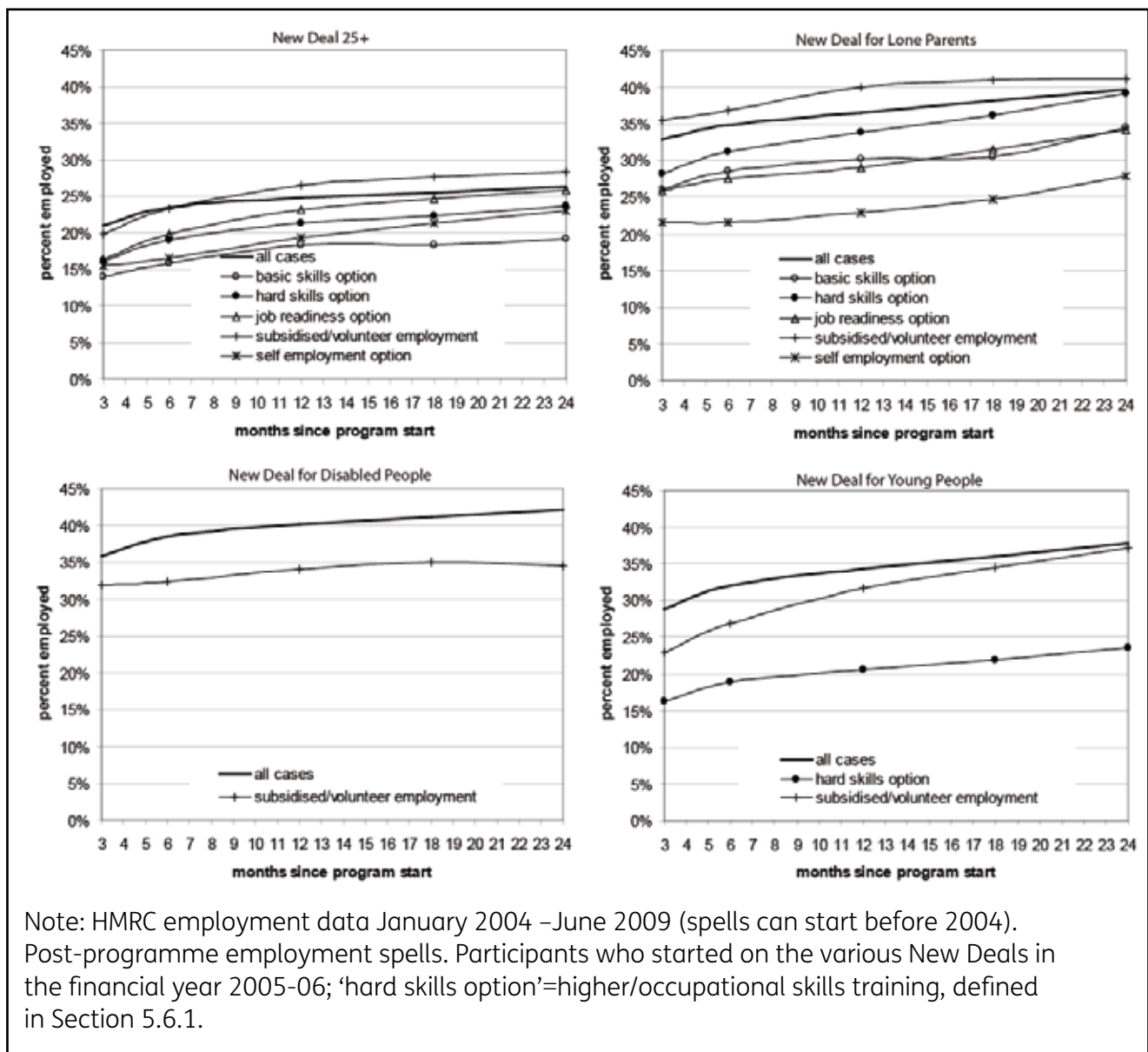
We reach similar conclusions looking at the average per cent of time that individuals are employed.
Looking at Figure 5.10, we see that the same patterns found in snapshots hold true for the window-
based measurement. Most notably, those in subsidised employment or volunteering options
are more likely to be employed and participants in the NDDP people have relatively better gross
employment outcomes.

Figure 5.10 Average per cent of possible time employed that programme recipients were employed 3, 6, 12, 18, and 24 months after programme start[45]

Note: HMRC employment data January 2004 –June 2009 (spells can start before 2004). Post-programme employment spells. Participants who started on the various New Deals in the financial year 2005-06; 'hard skills option'=higher/occupational skills training, defined in Section 5.6.1.
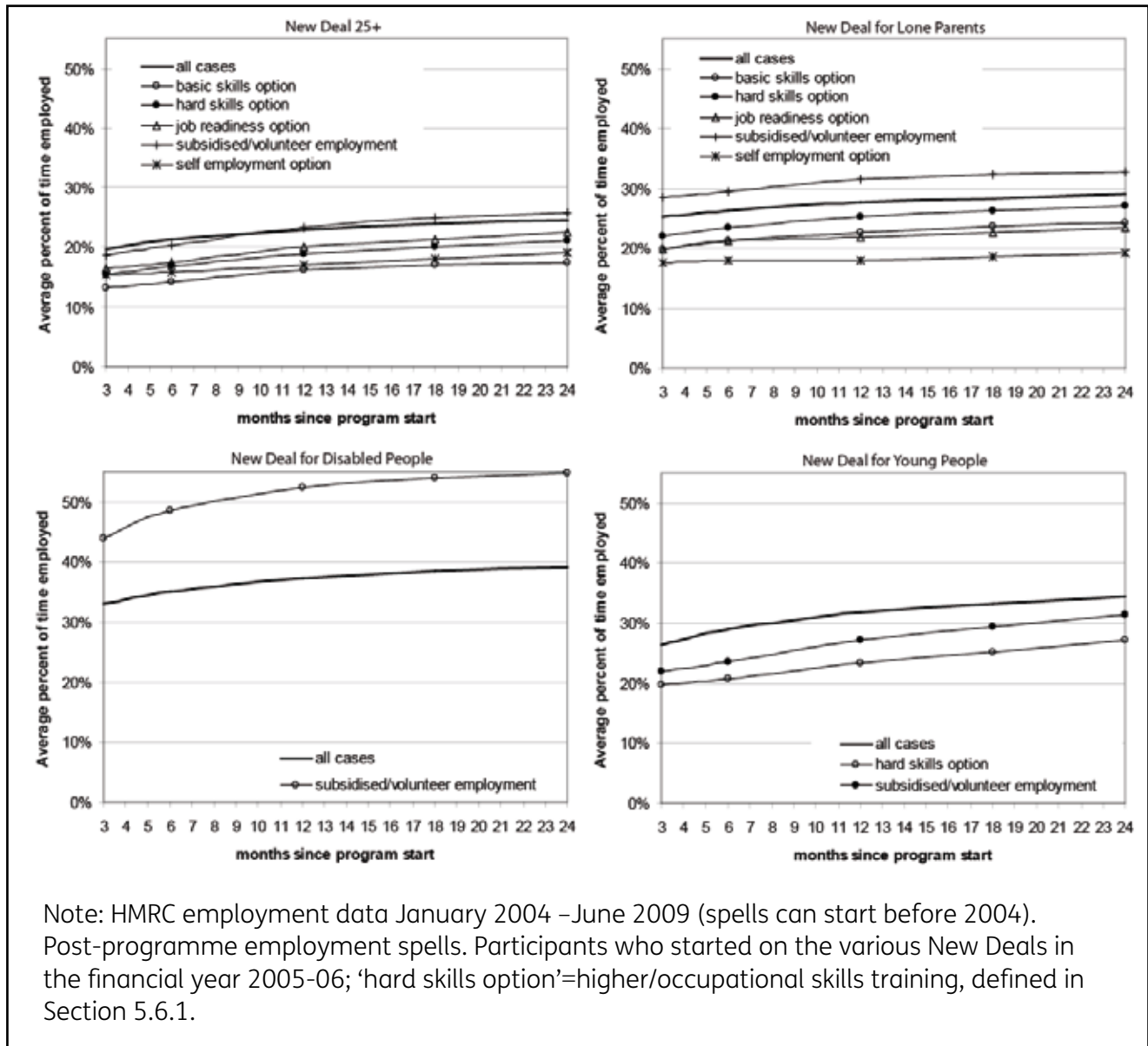
5.9.4   Gross nominal earnings

It is important to note, when looking at earnings, that we selected individuals based on their programme participation in 2005-06. This means that **by definition** the earnings for these individuals would have been lowest in 2005-06 (they were claiming benefits), or else they would not have qualified for the programme. This can be seen in Figure 5.11, where the shaded area is the 2005-06 programme participation period, and which for all indicates 2006 is the lowest point; however, note that the lowest point is not zero.

Earnings data are then reported for programme participants who entered in the 2005-06 tax year. The estimates are for the subset of individuals who were observed with earnings data. Non-taxable earnings (i.e. very low earners) are not observed in the earnings data. Each programme participant

---

[45]   All cases includes those going through options as well as those not going through options.

record from 2005-06 had one variable for taxable income for each fiscal tax year (2005-2009, April-March), for example total taxable earnings recorded for the tax year 2005-06. (This is the total earnings of participants who are observed at some point in 2005-06 with earnings data, of those with earnings data, so there is no one with zero for earnings data in the estimate.)

Looking at mean earnings across all programmes in Figure 5.11, we see that participant earnings declined between 2004 (when they were not on the programme) and 2005 (where they entered the programme) and stayed low through 2006 (when they were likely to still be in the programme). While participants' earnings did recover they never seem to recover to their pre-programme levels. Tables A.5, A.10, A.15, A.20, A.25, A.30 show the estimates underlying Figure 5.11.

Note that these are all nominal earnings and would need to be adjusted for the effects of inflation and also wage growth over time (with an appropriate measure) to reflect the current earnings level at a later date. Earnings can be adjusted for growth over time using the Office for National Statistics annual median earnings measure[46].

Regardless of this, the results seem plausible in interpretation. ND50+ participants are often workers who lose their jobs and then find it difficult to re-enter the workplace given their age. As such, it might be expected that the ND50+ group overall will not recoup their earnings. Similarly, those entering the NDDP programme were not necessarily disabled before entering the programme, but became ill/disabled in 2005-06. Hence they may earn relatively lower earnings upon their return to the workforce. In contrast, because participants who enter NDYP are at the beginning of their careers they recoup their pre-benefit earnings levels upon their return to work. NDLP participants almost entirely recoup their prior earnings levels. The participants of ND25+ never appear to recover to their pre-programme earnings in the follow-up period, as for ND50+.

---

[46]   http://www.statistics.gov.uk/statbase/product.asp?vlnk=13101 The Annual Survey of Hours and Earnings (ASHE) provides information about the levels, distribution and make-up of earnings and hours paid for employees within industries, occupations and regions. In particular, http://www.statistics.gov.uk/cci/article.asp?ID=2370 the Holdsworth series would be useful for overall medians, make-up and distribution of earnings of the 2009 ASHE. The analysis compares the 2009 results with the results from 1997. ASHE focuses on estimates of the median rather than the mean. The median is the value below which 50 per cent of employees fall. The median is preferred to the mean for earnings as it is less affected by extreme values and the skewed distribution of earnings data. However, estimates of the mean are still available in the annually published results.

Figure 5.11 Mean nominal gross earnings for those working with earnings records across programmes[47], for tax years 2004-09, for programme participation defined in 2005-06 tax year

Note: HMRC data January 2004 – June 2009. Participants who started on the various New Deals in the financial tax year 2005-06.
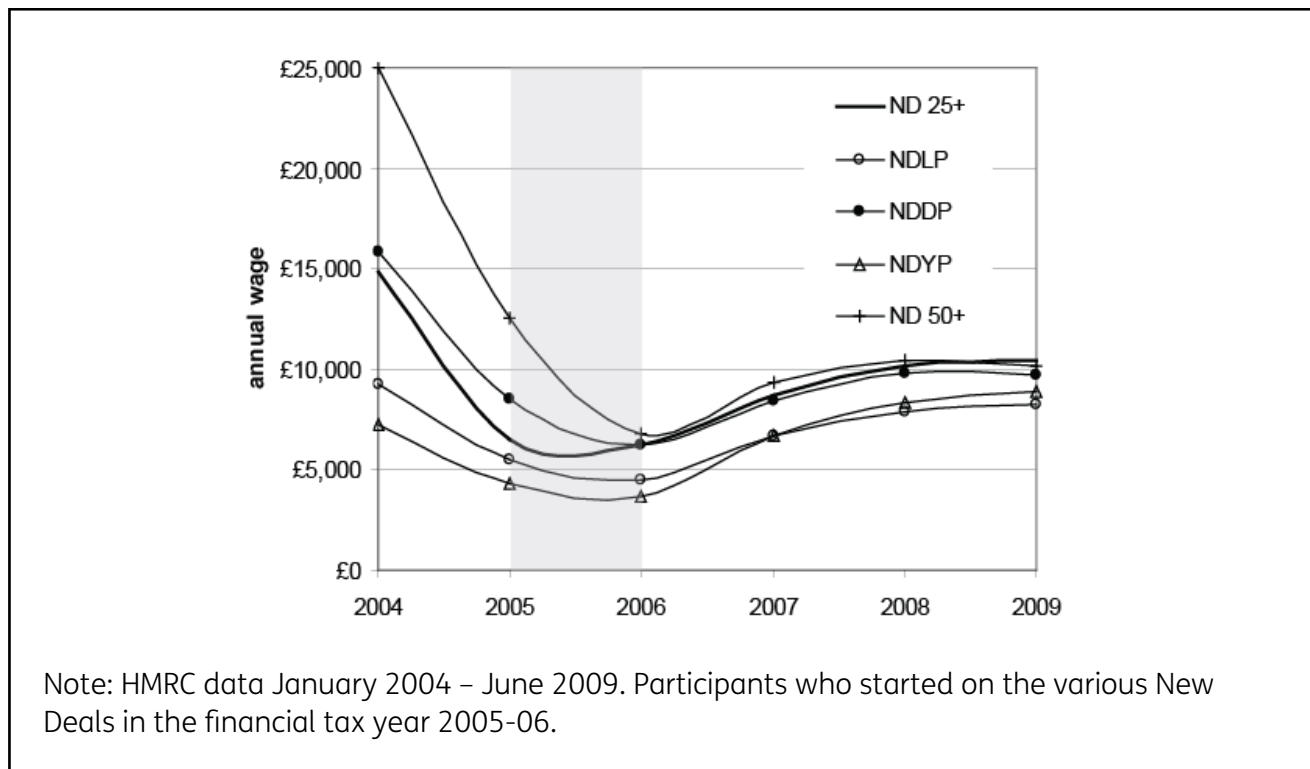
Looking at results by option in Figure 5.12, we can see that those workers with very high pre-programme earnings seem to start the self employment option. However, particularly for the ND25+ participants, individuals going through the self employment option have a permanent post-programme reduction in earnings. This, however, may reflect the fact that HMRC data does not include self-employment income. Those going through 'hard skills' training have a temporary dip in earnings below their basic skills counterparts (who are presumably less well educated), but later catch up. For NDLP, 'job readiness' options seem to yield weak results. For NDDP and NDYP, there is little variation in earnings by option.

---

[47]   Tables A.5 (ND25+), A.10 (ND50+), A.15 (NDLP), A.20 (NDDP), A.25 (NDP) and A.30 (NDYP) show the estimates underlying Figure 5.11.

**Figure 5.12 Mean earnings for those working, for each programme by option, for
tax years 2004–09, for programme participation defined in 2005-06
tax year**

Note: HMRC data January 2004 –June 2009. Participants who started on the various New Deals in the financial year 2005-06; 'hard skills option'=higher/occupational skills training, defined in Section 5.6.1.
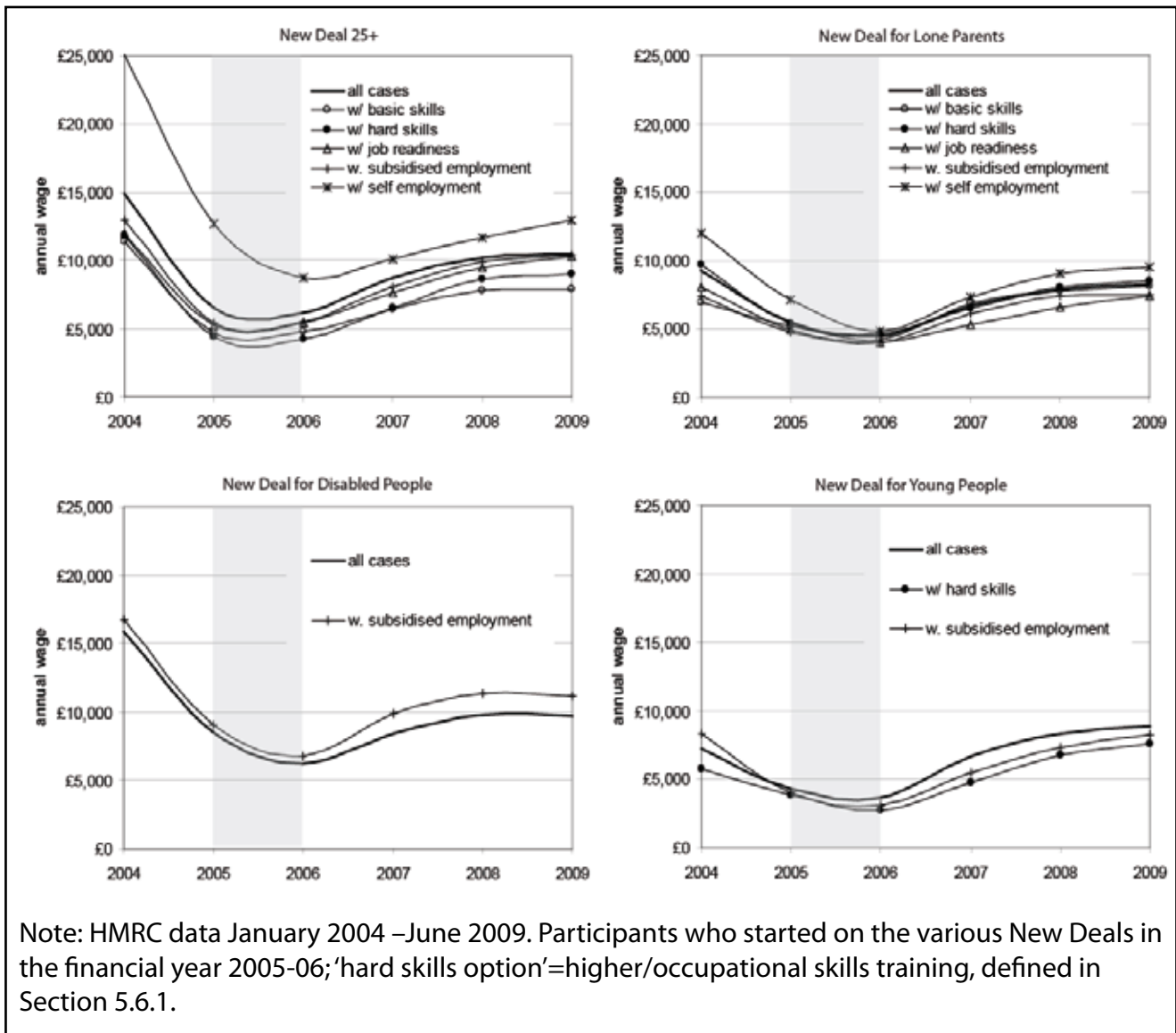
We also examined earnings for some subgroups in each programme, where data was available. Figure 5.13 shows NDYP and NDLP gross earnings by ethnicity. The caution set out in Chapter 3 on subgroup impacts should be applied to interpretation of these results. In addition, these are gross nominal earnings analyses, not net earnings impact analyses and do not represent the programme impact.

Figure 5.13 Mean earnings for those working, for each programme by ethnicity: for tax years 2004-09, for programme participation defined in 2005-06 tax year

Note: HMRC data January 2004 –June 2009. Participants who started on the various New Deals in the financial year 2005-06.
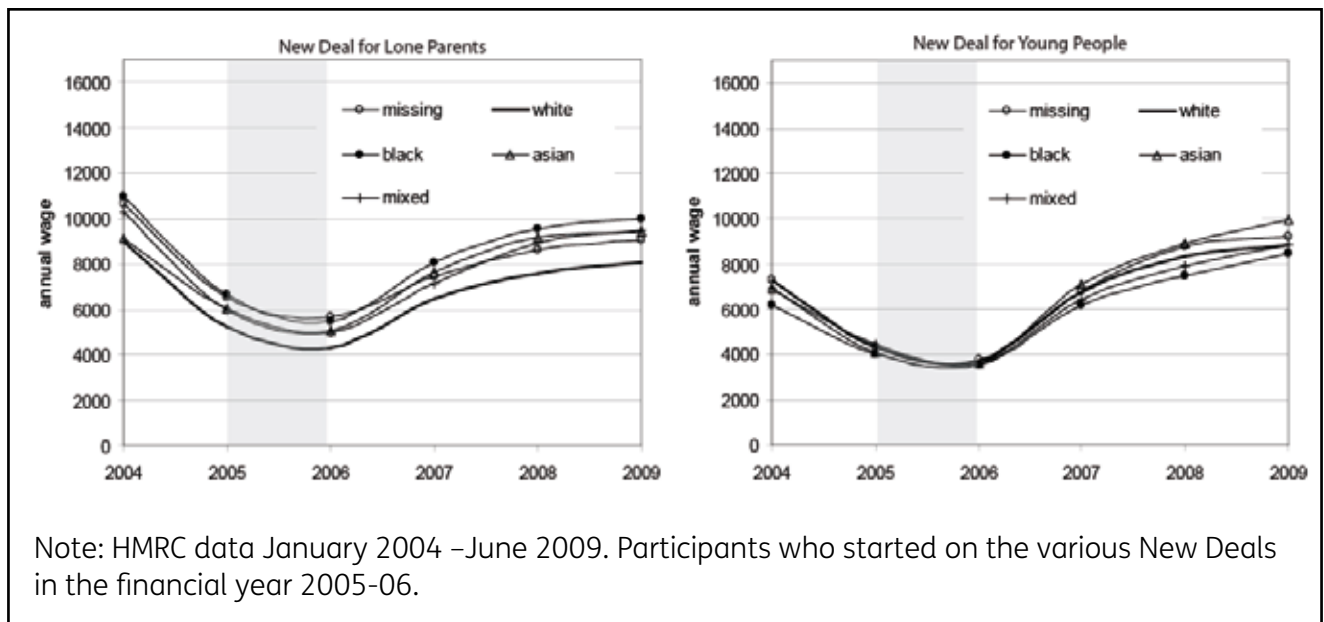
Results are available by marital status for the ND25+ and NDYP programmes. (There are not enough divorced NDYP participants to include a trend line in the graphic.) Results show that those who are married have higher earnings. Note for some, such as the widowed, the trend is unstable due to a low number of observations.

Figure 5.14 Mean nominal gross earnings for those working, for each programme by marital status: for tax years 2004-09, for programme participation defined in 2005-06 tax year

Note: HMRC data January 2004 –June 2009. Participants who started on the various New Deals in the financial year 2005-06.
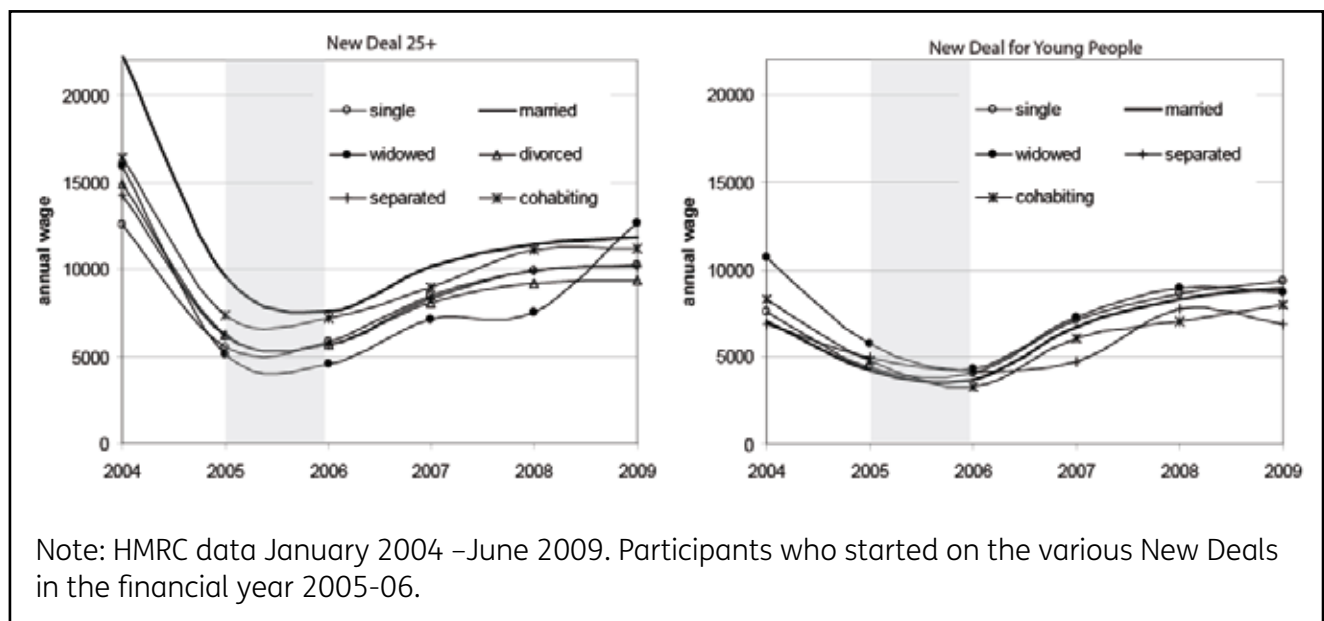
6   Multiple programme participation

Greenberg and Knight (2007) indicate that the minimum that should be done in a cost benefit analysis (CBA) to examine multiple programme participation is an assessment of how many participants are on multiple programmes or repeatedly participate in a programme.

Several types of multiple participation are explored in this chapter:

- a programme can offer a set of different programme options at one time (multiple programmes);

- an individual can be on more than one programme at a single point in time (multiple instantaneous participation – programme combinations);

- an individual can participate in the same programme more than once (repetition – programme combinations);

- an individual can participate in different programmes subsequently over their life (participation dynamics – when people participate and programme combinations).

It is recommended that the best way to reflect multiple participation is through discussion and estimation of the scale of multiple participation observed and exploring the importance of the timing of the gaps between participation. An assessment is also made of what data is available for this and how it can be obtained. This chapter also investigates whether there is a way to measure, for a specific programme, the extent to which the outcomes achieved by the programme are partly due to the intervention regimes that preceded it (repetition).

6.1   Multiple participation in existing cost benefit analysis

The current evaluation evidence on most effects of mandatory and voluntary programmes is based on impact assessments examining the effects of particular cohorts starting a programme.

In CBAs for most programmes, the Department for Work and Pensions (DWP) relies on impact assessments from studies on the effectiveness of programmes. Most of these studies focus on **a particular entry cohort into a programme**. However, mandatory programmes, in particular, may result in cohorts **repeatedly beginning a programme**. This means that some participants of the cohort under consideration may be participating a second or third time. The programme effects on these repeated spells might be quite different than the effects on first-time participants. Thus, an estimated employment outcome for a particular cohort entering a programme potentially covers differential effects for a diversity of participants starting the programme either a first or a repeated time. This may not be entirely informative for CBA in the context of mandatory programmes **enforcing repeated participation**. As a consequence, it is important to outline the programme design for participation and the complexity of the resulting programme combinations over time as the impacts can differ by participation type.

6.2   Identifying the complexity of programme participation

Using Master Index (MI) data[48] allows us to get a clear picture of individual programme participation. As in the secondary evaluation of the New Deal for Lone Parents (NDLP) (Knight *et al.,* 2006), exploiting all longitudinal information offered by these data provides a thorough understanding of the combinations of services delivered to the main target groups, e.g. lone parents or disabled job seekers. MI data can show the full extent of multiple and repeated participation that job seekers experience, reflecting many issues currently neglected in evaluation studies, in particular:

- **Multiple programmes**: There are usually different options for eligible benefit claimants, e.g. the participants on the various New Deals, who can begin training programmes like Basic Skills or programmes allowing them to gain work experience (e.g. an employment option on the 'option stage' of the New Deals).

- **Repeated participation**: Many programmes are mandatory, like the New Deal for Young People (NDYP), and as for now, empirical evidence does not account for the effects of the first (or an initial) programme participation as compared to a later participation. However, the effects of a second participation may be different, and averaging the outcomes of first and second participation may not be the most informative for some types of cost/benefit appraisal questions.

- **Participation dynamics**: Currently, most studies do not account for potential selection problems based on unobservable characteristics or a potential future participation in programmes. However, there are important reasons why evaluation evidence may be biased if such problems are ignored.

6.3   Multiple programmes: programme effects relative to other alternatives

This section explores the multiple programmes available to a particular eligible group. This mostly includes the designed multiplicity of programmes such as NDYP Options – in this case, usually one can only participate in any one Option at a single point in time (multiple instances over time then are covered in the next section). Such multiplicity of programmes for a particular eligible group can arise without design, but is usually a less common case.

As different options exist for a single target group, it is often very important to understand which of the available programmes are most suitable. Even within a particular programme (like Work-Based Learning for Adults (WBLA)), different sub-programmes exist (like Short job-focused training or Basic Employability Training).

Lechner (2001) has pioneered approaches to evaluating multiple programme options. Essentially, an evaluation of relative outcome effects of different programmes can be achieved by comparisons of matched individuals receiving different programme alternatives (and also compared to no participation at all, showing the effect of the programme relative to no participation at all).

Formally, the estimation of relative programme effects in Lechner (2001) identifies effects as the difference in expected values between a particular outcome for a particular treatment compared to the outcome for alternative states. This framework can be generalised for a number of different programme alternatives out of a choice set. As with all non-experimental evaluation studies, the natural measure of comparison ('counterfactual' outcome) of a participant in a programme cannot be observed and the causal effect can only be identified by imposing additional assumptions.

---

[48]   See Appendix D for a description.

The identification of alternative non-treatment outcomes relies on a Conditional Independence Assumption (CIA): conditional on descriptive characteristics variables, the actual observed outcome is independent of the actual participation in either one of the different programmes as long as there are comparable participants in the other options (or non-participants) with similar characteristics.

In practical terms, it is sufficient to compare participants in one particular programme with participants in a particular other alternative as long as both groups show similar characteristics. By means of such comparisons, one can show the relative gain of participating in long-term occupational training as compared to short-term job-focused training. Comparing the effects found for one programme relative to another is informative as the net gain of one programme relative to another can be contrasted to the relative cost differences.

Lechner (2001) also shows that his framework for estimating such relative programme effects can be combined with propensity score matching, which greatly simplifies the practical work of estimating the relative programme effects. The propensity score as a function of the covariates as opposed to a direct match on a large number of different **characteristics** does not result in a lack of similar comparisons, and yet, characteristics would be similar for participants in the different programme alternatives.

Although the econometric methods of estimating such relative programme effects was established some time ago, very few evaluation studies have been published comparing the effects of one programme to another. The reason for the lack of such evidence is primarily a matter of insufficient information available; participants' surveys and administrative data usually do not allow an understanding of **why** individuals were assigned to one particular programme and not to another. Only if such information is known to the researcher (and the CIA is then credible), can one validly compare participants in one programme with participants in another. Therefore, in order to estimate the relative programme effects, one would need to include a number of important implementation variables. Once such information is available, one can estimate the net gain of participating in one programme as relative to the participation in another programme.

Applying the Lechner (2001) framework, one can estimate the effect of one option as compared to other options. There are two informative estimates which can be achieved:

- an estimate of the average programme effect in one particular option compared to non-participation in any of the programmes (which was used for estimating the effects of WBLA (see Speckesser and Bewley (2006)); or

- the comparison of participating in one particular option as compared to another option, 'Pairwise comparison' (which was the design applied in Biewen *et al.* (2008)).

6.4 Repetition, participation dynamics and the effects of programme combinations

If participation in one particular programme is **followed** by another programme, it may be appropriate to combine first and second treatments and think of it as a combined intervention. One would then estimate the effects of the combination of the two programmes, i.e. the effect of the entire sequence.

An estimation of the combined effect of first and further treatments may be justified if the initial selection into such combinations of treatments, i.e. prior to the beginning of the first elements of the treatment combination, can be modelled correctly. This implies that the researchers have a good understanding of the implementation of programmes and know how participants are assigned into particular suites of different programmes.

If, however, the participation in a subsequent element of observed treatment combination **depends** on the outcome of the first treatment, it may be questionable to evaluate combinations of first and second treatments as the real participation process may not justify such an aggregation.

It has been shown in earlier analysis of the combined effect of Lone Parent Work-Focused Interviews (LPWFI) and the NDLP (Knight *et al.*, 2006) that an evaluation of bundles or combinations of programmes may be reasonable for treatments occurring at the same time or not requiring explicit outcomes of a first/initial treatment.

The current method of estimating the relative costs and benefits of employment programmes is structured along the lines of individual programmes rather than particular combinations, like the combination of participation on a specific New Deal programme with some particular option elements like Basic Skills. However, an evaluation of sequences could be informative to find the option bundle for particular target groups in terms of labour market impact and cost effectiveness. This would however require knowledge of the particular assignment mechanism into a suite of programmes, which would only be possible if reasonably high quality data on the process existed.

For better assessment of programme participation, we recommend that more quantitative management information on (caseworker) selection for services is recorded into the Labour Market System (LMS)[49], that the quality of recording is better maintained and checked, and that it is used for evaluation purposes.

6.5   Effects of programme repetition

In the context of some mandatory programmes, some participants may participate in the same programmes more than once. A prominent UK example for repeated participation is the LPWFI and NDLP combinations (see Knight *et al.*, 2006; Dolton *et al.,* 2004).

It is very unlikely that a second participation in the same programme yields the same returns for participating individuals as the first participation. Only a very few evaluation studies have ever considered the issue that the returns of programme participation may differ for a second participation in the same programme. Instead, in most studies, there is a focus on 'programme participation', irrespective of whether individuals have previously participated in the programme (history). This results in an estimate which implicitly averages the programme effects of first and further participations.

In practical research, there can be great difficulties with regard to understanding the allocation of participants to a second or further reiteration of the same programme. This is important to explore in the data and the context to understand what is valid for the treatment effects. For example, it is necessary to understand:

- What are the selection rules, in particular for a second participation in a voluntary programme?

- What are the returns of a second participation if the objective was already set for the first participation?

- Is it possible to evaluate the effect of a particular participation event and assume independence of participation over time (i.e. can we assume independence of the participation spells)?

The recent literature in the context of repeated treatments suggests that different options exist for the evaluation of the effects of a first treatment and of additional incremental effects of later treatments (e.g. Bergemann, Fitzenberger and Speckesser, 2005).

---

[49]   LMS is used for the Jobcentre Plus services supplier. For private sector supplied services, it will be important that sufficient management information data is collected in a similar format, because if the same level of management information from external providers is not available there is a further constraint in the analysis that can be done – for example, it might not be possible to compare impacts or value for money of services from the different supplier(s).

In the theoretical literature, the evaluation of repeated treatments from observational data distinguishes the initial effects of a first participation from incremental effects of a second participation (see Lechner and Miquel, 2002). Such a framework requires separate econometric modelling of the decision to participate at different points in time while a benefit spell is ongoing, i.e. it requires a dynamic framework.

In empirical work (Bergmann *et al.*, 2004), combined effects of a first and second programme participation have been estimated as well as the incremental effect of a second programme participation. This however assumes that selection into a particular combination can be adequately controlled for before the beginning of the first treatment, neglecting a possible sequential assignment into programmes.

A plausible approach for evaluation of the incremental effects of repeated participation has been discussed theoretically by Lechner and Miquel (2002) modelling the decision to participate in the second iteration separately from the decision to participate a first time. The corresponding evaluation framework mirrors such a stepwise decision for participation in a two-stage procedure of the established multiple treatment approach taking into account that participation in the second treatment depends on the outcome of a first treatment. Lechner and Wieler (2007) demonstrate with an application to Austrian data.

A practical example is as follows: if employment depends on the skills an individual offers, then a single participation in a programme like WBLA surely affects this variable. The evaluation therefore has to take account of the outcome of the **first WBLA participation** as an important variable for a potential second participation. Since the outcome of the first participation may lead to successful labour market integration for some participants and may lead into a second participation for others, the comparison group used for the evaluation of the first participation is no longer informative in order to estimate the outcome of non-participation in the first reiteration. A new comparison group needs to be found for the group participating another time – potentially consisting of the participants in the first programme. Out of the group of programme participants, the group showing a second participation should then ideally be compared with **individuals participating in the programme only once who show on average similar characteristics**. As long as the observed characteristics are similar between both groups, then CIA is again assumed to hold, i.e. that the comparison group represents the outcome that would have materialised if there had been no second participation. Note that historically many United Kingdom programmes are targeted at specific groups, for example a lone parent claiming Income Support can go on NDLP) but is unlikely to go on NDDP (New Deal for Disabled People) (available to IB claimants), due to the eligibility requirements. These methods are only relevant where similar participants might participate in different programmes.

The implication is that we **need to understand the process** leading to a second participation in order to validate the treatment estimation. A repeated participation could be because of the mandatory nature of the programme (like the New Deals) or because of individuals opting in to participate in some activity twice. As with the evaluation of multiple programme alternatives, a research design for the evaluation of incremental programme effects exist (see Lechner and Miquel, 2002), but has rarely been applied to programme evaluations.

The outcome of a first participation can be very important for modelling the selection into the second participation, but **only if the outcome of the first programme is considered to be exogenous from the participation in the second programme** (Lechner and Miquel, 2002: 13). An example for such an outcome could be an insufficient result, e.g. a failed certificate of achievement for some particular training, which could then be obtained in a repetition. However, this may not be a credible assumption on how programmes can influence adult learning. If the outcome of the first

participation in a programme changes because of the second participation and the success of the first programme participation only materialises during the second participation, then those who participate only once will not be a realistic comparison group. Such a possibility is very likely if the programme content differs, e.g. for training programmes. The learning on a second programme may then actually help achieving the outcome of the first programme. Under such circumstances, it is no longer clear whether we can attribute the impact of a programme to a first or a second participation of the programme. If we attributed everything to the second participation, we would incorrectly overstate the impact. If we attributed everything to the first participation, we would understate it. Either way, the group of participants who did the programme but who did not retake it does no longer allow the estimation of the non-participation in the second round: this group has had the programme impact of the first programme and – with no second programme participation – an effect of the first programme induced by the participation in the second programme will not materialise.

The evidence of employment outcomes in the context of multiple treatments is weak due to the statistical difficulties of modelling a second or further participation. Hence, evaluation should be constrained to be for the first treatment. In the case of the UK programmes, there has been no exploration of how we can evaluate first and further increments of particular programmes. However, the descriptive evidence (Knight *et al.,* 2006) shows how important this potentially is (in particular, in the context of mandatory programmes) – see Section 6.7.

Programme repetition needs to be measured – what share of participants have one, two, three or more participations?

In the event of substantive second or further participation, compare the characteristics of those with 'two participations' against those with 'one participation', if statistically similar (t tests should be used to determine whether they differ) then the comparison group for one participation can be the same as that for two participations, i.e. the overall impact of all treatments is reasonable but may contain some bias. Reporting should clearly indicate that the impact could represent an average impact of different participations and there may be some bias.

The process of participation needs to be explored in the reporting, in the context of whether it might be reasonable to consider the second participation to be unaffected (exogenous) by the outcome of the initial participation.

Where there is substantial second or further treatment, and the characteristics of the groups statistically differ, we recommend that valid impact analysis should **restrict the treatment to the first participation** in a programme. This means for the cost-benefit that the total additional time is only for the members of a cohort who are on their first participation. The reporting should clearly state this restriction.

There can be important information for programme design purposes in estimating the incremental effects of second and further participations, and then conducting the cost-benefit by these subgroups. This does have information requirements by subgroup, such as costs by subgroup. For example, the incremental impact of a second treatment may be very small relative to the costs and it can inform that restricting to one participation is cost-effective.

6.6   Differences in effects due to different durations of benefit before participation

The duration of unemployment or benefit receipt before participation in any of the programmes is an important explanation for the assignment to treatment. Indeed, for some programmes, individuals are mandated to join after given lengths of time spent on benefits. However, in most evaluation studies, the time spent on job search before participation in a programme is widely ignored.

Fredriksson and Johansson (2003) show that ignoring the participation dynamics of **when** people participate (or decide not to) may result in biased estimates of the employment effect. First, individuals who are not treated up to the end of the time period might be participants in a programme after this time. Bearing this in mind, a comparison group partially includes later participants and may therefore have less favourable employment prospects than true non-participants. If this is taken into account, the comparison group may not be an appropriate estimate of the non-treatment outcomes, even when conditioning on observable characteristics, and would result in biased estimates. Second, evaluation studies usually consider participation in a particular time period, e.g. the first six months of unemployment or so. However, the comparison group may not represent an unbiased estimate for the treatment group as some of the non-participants may have actively refused programme participation because of an anticipated job entry. Thus, even when conditioning on observable characteristics, the comparison group would indeed have more favourable employment prospects and the treatment effects would be biased.

One can never rule out that individuals who do not receive treatment today may not be participants in the future. This could lead to bias in estimates of employment effects. To avoid this, it could be assumed that the timing of treatment matters. For example, if treatment in one year differs from the treatment in the next year, then a comparison group whose members do not start treatment in the specific period could provide a valid non-treatment outcome.

Hazard model approaches are another important method of evaluating programme effectiveness in the presence of dynamic selection issues as they allow a simultaneous modelling of the transition into programmes and outcomes.

These approaches are increasingly important because they allow dynamic selection issues to depend on unobserved heterogeneity[50] in both participation and outcomes (Abbring and Heckman, 2007). Abbring and van den Berg (2003) explain how, under particular assumptions, a non-parametric model of both transition to programme participation and outcomes can be used to identify the causal effect of the treatment in non-experimental data conditional on observed covariates, general dependence and unobserved heterogeneity.

As far as we know, there are no studies for the UK considering the dynamics of the comparison group.

6.7   Empirical evidence

A number of studies have published a comprehensive evaluation of multiple and mutually exclusive treatments. These include Biewen *et al.* (2008), Fitzenberger *et al.* (2009), Lechner *et al.* (2005), Hotz *et al.* (2006) and Knight *et al.* (2006).

---

[50]   Both true state dependence in outcome variables (e.g. employment outcomes for individuals are conditional on previous employment outcomes) or unobserved heterogeneity can explain why causal effects should be modelled so as to be aligned to individual employment (or programme participation) dynamics.

Some of these recent studies also consider the selection into programmes as a dynamic problem. They compare the participation with non-participation or later participation based on Sianesi's (2004) paper outlining participation dynamics as an important concept. In the following, we review some of the most important findings of these studies and discuss the applicability of the approach in the context of the UK.

Biewen *et al.* (2008), Fitzenberger *et al.* (2009) and Lechner *et al.* (2005) use similar data from the German Social Insurance Records and compare the relative effectiveness of different training programmes available for particular target groups among the unemployed. The design of these evaluation studies always focuses on particular entry cohorts into unemployment and the participation in a particular programme alternative after some elapsed duration of unemployment. These studies find, in most cases, positive effects of the different alternatives compared to non-participation, but also considerable differences in the programme effectiveness when comparing two alternatives with each other. In addition, the programme effectiveness seems to be higher for job seekers having been unemployed six months or more before the beginning of the programme.

Hotz *et al.* (2006) evaluate the differential effects of labour force attachment (LFA) versus human capital development (HCD) training components from California's Greater Avenues to Independence (GAIN) programme. They find stronger positive effects of LFA in the short term and stronger positive effects of HCD in the long term.

Knight *et al.* (2006) evaluate the impact of different combinations of LPWFIs and NDLP on the movements off benefit by eligible lone parents, with NDLP, LPWFI, or the combination of both, as particular 'routes' through the benefit system. The paper focuses on the differential effects, applying a static multiple treatment framework for the first programme participation of the eligible population in the light of very substantial repeated participation. The main findings support a strong positive effect of NDLP, but do not clearly show additional benefits of LPWFI.

While all these studies applied a framework evaluating mutually exclusive alternative treatments for particular entry cohorts, there are few papers using Lechner and Miquel's (2002) identification strategy of effects of reiterated treatments. While there was strong evidence for repeated participation in the secondary evaluation of the NDLP (Knight *et al.*, 2006), so far there has not been a consistent evaluation of a second or further participation in the same programme except for the Austrian analysis by Lechner and Wieler (2007) and Bergeman *et al.* (2000). If the full variety of services available for a particular eligible group was considered, we would possibly see many more persons participating in particular elements of welfare-to-work programmes more than once. However, there has not been a consistent evaluation of repeated participation for the UK programmes.

Generally, the methodology for the evaluation of sequential treatments is established. However, it is uncertain as to whether data are sufficiently informative to apply these methods.

> Improved recording quality of LMS management information would likely allow one to estimate incremental or combined or relative effects. This could greatly improve cost-benefit estimates for programmes.

6.8   Exploring the magnitude and impact of multiple participation in the UK

The estimates of employment effects used in cost benefit analysis (CBA) in DWP are usually based on the results of particular evaluation studies. These estimates are based on specific entry cohorts.

> We recommend re-analysing the employment effect of the cohorts currently underlying the employment effects included in DWP's CBA, exploring participation repetition and characteristics and identifying the extent of multiple reiterated treatments, focusing on the most important programmes (e.g. the programmes for job seekers, like the New deals).

In particular, based on merged data of MI and Work and Pensions Longitudinal Survey (WPLS), explore these items:

* How many of the participants participate more than once in the same programme?

* How much time, on average, is spent between the first and a further participation in the programme and what the participants did between these iterations? This should give some indication about the mutual dependency of both treatments, and whether we can assume some weak independence between a first and a second participation or not.

* What are the employment outcomes for participants who only participated once, e.g. on the average employment in a particular period following participation? In the light of multiple repeated treatments, what is the initial effect of a first participation – what is the incremental effect of a second participation separately? This should give some indication on how the average effect currently used for DWP internal CBA compares to an effect which explicitly considers the multiplicity of the programme participation.

The available data from administrative records can be used for an empirical exploration of the magnitude of multiple programme participation and the estimation of employment and benefit effects of first and later participations. WPLS data have a comprehensive history of all benefit and Her Majesty's Revenue and Customs employment spells as well as any participation in one of the following programmes:

* New Deal 50+;

* NDYP;

* New Deal 25+ (Long-term unemployed);

* NDLP;

* NDDP;

* New Deal for Partners of the Unemployed;

* Basic skills;

* WBLA;

* Employment Zones.

In principle, the extent of multiple participation can be shown for all programmes implemented by the DWP in recent years as data can usually be linked by the encrypted National Insurance numbers. Consequently, many more programme combinations and repetition issues could be explored (for programmes such as Action Teams Adult Learning Option Pilot, Workstep, Progress2Work, Minority Ethnic Outreach, IB Reforms, Access to Work, Referrals). For larger programmes, like the New Deals and Basic Skills in particular, repetition is more likely for they are mandatory programmes for particular customers.

We recommend estimating the effect of a first participation, and then, when relevant, also estimating the effect of a second participation in the largest recent mandatory programmes (e.g. the New Deals). First, the extent of initial and later programme participation should be identified. Based on available data (WPLS/MI) employment and benefit outcomes of first and later participation in these programmes should be estimated separately.

Estimates of programme effects which have disregarded the particular number of repetitions of the programme should be replaced by estimated effects which reflect the true characteristics of the participants under consideration, i.e. whether this is a first or a later participation.

With better impact estimation of multiple participation impacts, it might be necessary to have more detailed cost information as costs may differ across participations. With such further information, the sensitivity of cost-benefit estimates to assumptions regarding cost differences can be explored.

Appendix A
Benefit duration, employment duration and wages results by programme

Results begin with the New Deal for individuals aged 25 plus (ND25+) programme, for which all subgroups were available. For some programmes, rows are missing when there is no data available. For example, the New Deal for individuals aged 50 plus (ND50+) programme does not have programme options. Options can start at a date after the programme entry.

Subsidised employment/volunteer work includes:

- ND25+: the Employment Option, Work Placement, Work Experience, and Subsidised Employment (SE).

- New Deal for Lone Parents (NDLP): Worktrial, NDLP job (on other caseload), NDLP in work (on IS), NDLP VS, and NDLP VS Familiarisation.

- New Deal for Disabled People (NDDP): payptjb, payptse, payftjb, payftse, payptft, pyptftse, pyptsus, pyptsesu, pyftsus, and pyftsesu, all of which indicated a payment combined with some type of employment.

- New Deal for Partners (NDP): no recoding was necessary as F_WEX_ST already indicated subsidised employment/volunteer work.

- New Deal for Young People (NDYP): the employment option, VS, and ETF.

Basic skills training includes:

- ND25+: BET/BS option.

- NDLP: basic skills, other training, WBLA, NDLP B-Skills, NDLP IAP Education and Training, LSC BS Training, and LSC ESOL Start.

- For the NDDP, NDP, and NDYP: no variables indicated a basic skills option.

Higher/occupational skills training (labelled hard skills option in figures) includes:

- ND25+: FTE, WBLA, ETO, and Training for Work.

- NDLP: Other Education/Training, SJFT, AM Demand Led Training, European Social Fund (ESF) Provision, ESF ILM, Nat/Prog Dev Fund, FTE, FTE training familiarisation, Careers, LSC Occ Training, and LSC Mod App Training.

- NDP and NDYP required no recoding.

Self employment:

- ND25+: general self employment option as well as the musicians-option.

- NDLP: Business Advice Agencies, Self Employed Stages 1, 2, and 3, ND Music NDLP MIC, and ND Music NDLP MOLP.

- NDP and NDYP: no self employment option was coded.

Job-readiness option:

- ND25+: IAP training, Soft Skills/Short Motivation, and job search.

- NDLP: (mandatory) job search provision, job club, job interview guarantee (JIG), job search seminar, job fairs, progress to work, Environmental Task Force (ETF) (familiarisation), job search and advice (J&A), Gateway (GW) training (KS, IT, Motivation, or Voc), and Learning and Skills Council (LSC) Information Advice and Guidance (IAG).

- NDP: first start/receiving advice and guidance.

- NDYP: no job readiness option.

ND50+: There was no available information on options for those in ND50+.

ND25+

Table A.1   New Deal 25+ benefit receipt rates for selected start dates of participation, for periods between three and 24 months after participation

		Percentage share on benefit (at selected points in time following start)				
		Months since programme start				
Start	**Type of option**	**3**	**6**	**12**	**18**	**24**
Cdpstart	All participants	63%	44%	51%	49%	43%
	With basic skills	48%	26%	62%	63%	57%
	With higher skills	73%	45%	66%	64%	58%
	With job readiness	69%	34%	58%	58%	51%
	With subsidised employment	71%	39%	60%	59%	54%
	With self employment	53%	21%	25%	29%	27%
Gwstdt	All participants	59%	43%	51%	48%	42%
	With basic skills	41%	25%	63%	63%	56%
	With higher skills	68%	45%	66%	64%	57%
	With job readiness	66%	33%	57%	57%	52%
	With subsidised employment	65%	39%	61%	59%	53%
	With self employment	47%	18%	25%	28%	28%
nd1intdt	All participants	59%	43%	51%	48%	42%
	With basic skills	41%	25%	63%	63%	56%
	With higher skills	68%	45%	66%	64%	57%
	With job readiness	66%	33%	57%	57%	52%
	With subsidised employment	65%	39%	61%	59%	53%
	With self employment	47%	18%	25%	29%	27%

Note: For example, row 1, cdpstart: of all individuals starting ND25+, 63 per cent are observed on benefit three months after the programme start (defined by Cdpstart), 44 per cent at six months, etc.

Table A.2   **New Deal 25+ average per cent of time on benefit for selected start dates of participation, for periods between three and 24 months after participation**

		Percentage of time spent on benefit in periods following participation				
		Months since programme start				
Start	**Type of option**	**3**	**6**	**12**	**18**	**24**
cdpstart	All participants	77%	65%	57%	54%	52%
	With basic skills	71%	53%	49%	54%	55%
	With higher skills	85%	71%	65%	65%	64%
	With job readiness	80%	65%	57%	57%	57%
	With subsidised employment	82%	68%	60%	60%	59%
	With self employment	73%	54%	37%	34%	33%
gwstdt	All participants	73%	61%	55%	53%	51%
	With basic skills	64%	48%	48%	53%	55%
	With higher skills	82%	67%	63%	64%	63%
	With job readiness	77%	61%	56%	57%	56%
	With subsidised employment	79%	63%	59%	59%	59%
	With self employment	70%	50%	35%	33%	31%
nd1intdt	All participants	73%	61%	55%	53%	51%
	With basic skills	64%	48%	48%	53%	55%
	With higher skills	82%	67%	63%	64%	63%
	With job readiness	77%	61%	56%	57%	56%
	With subsidised employment	79%	63%	59%	59%	59%
	With self employment	69%	49%	35%	33%	31%

Table A.3   New Deal 25+ per cent employed for selected start dates of participation, for periods between three and 24 months after participation

Start variable		Type of option	Months since programme start				
			3	6	12	18	24
start1	cdpstart	All cases	21%	23%	25%	25%	26%
		With basic skills	14%	16%	18%	18%	19%
		With hard skills	16%	19%	21%	22%	24%
		With job readiness	16%	19%	23%	24%	26%
		With subsidised employment	19%	23%	27%	27%	28%
		With self employment	15%	16%	19%	21%	23%
start2	gwstdt	All cases	21%	23%	25%	26%	26%
		With basic skills	14%	16%	18%	18%	19%
		With hard skills	16%	19%	21%	22%	24%
		With job readiness	16%	20%	23%	25%	26%
		With subsidised employment	20%	23%	27%	28%	28%
		With self employment	16%	16%	19%	21%	23%
start3	nd1intdt	All cases	21%	23%	25%	26%	26%
		With basic skills	14%	16%	18%	18%	19%
		With hard skills	16%	19%	21%	22%	24%
		With job readiness	16%	20%	23%	25%	26%
		With subsidised employment	20%	23%	27%	28%	28%
		With self employment	16%	16%	19%	21%	23%

Note: For example, row 1, cdpstart: of all individuals starting ND25+, 21 per cent are observed in HMRC employment three months after the programme start, 23 per cent at six months, etc.

Table A.4   New Deal 25+ average per cent of time employed for selected start dates of participation, for periods between three and 24 months after participation

Start variable		Type of option	Months since programme start				
			3	6	12	18	24
start1	cdpstart	All cases	19%	21%	23%	24%	24%
		With basic skills	13%	14%	16%	17%	17%
		With hard skills	15%	16%	19%	20%	21%
		With job readiness	15%	17%	19%	21%	22%
		With subsidised employment	19%	20%	23%	24%	25%
		With self employment	15%	16%	17%	18%	19%
start2	gwstdt	All cases	20%	21%	23%	24%	25%
		With basic skills	13%	14%	16%	17%	18%
		With hard skills	15%	17%	19%	20%	21%
		With job readiness	15%	17%	20%	21%	23%
		With subsidised employment	19%	20%	23%	25%	26%
		With self employment	26%	16%	17%	18%	19%
start3	nd1intdt	All cases	20%	21%	23%	24%	25%
		With basic skills	13%	14%	16%	17%	18%
		With hard skills	15%	17%	19%	20%	21%
		With job readiness	16%	17%	20%	21%	23%
		With subsidised employment	19%	20%	23%	25%	26%
		With self employment	15%	16%	17%	18%	19%

Table A.5   New Deal 25+ mean annual nominal taxable wages observed for employed (£)

Subgroup	Type of option	Year					
		2004	2005	2006	2007	2008	2009 (May)
By skill	All cases	14,864	6,495	6,190	8,720	10,158	10,475
	With basic skills	11,433	4,650	4,822	6,404	7,820	7,901
	With hard skills	11,874	4,452	4,217	6,495	8,619	8,981
	With job readiness	11,974	5,225	5,413	7,609	9,490	10,276
	With subsidised employment	12,941	5,429	5,479	8,073	9,894	10,378
	With self employment	25,054	12,713	8,699	10,118	11,695	12,998
By ethnicity	Missing	15,055	6,742	6,615	9,247	10,811	11,035
	White	15,138	6,675	6,233	8,693	10,135	10,403
	Black	12,991	5,361	6,115	9,084	10,337	11,017
	Asian	13,587	4,906	4,986	8,104	9,313	10,273
	Mixed	12,440	5,099	5,577	8,174	9,812	10,175
By marital status	Unknown	17,907	9,619	6,292	8,626	10,610	9,688
	Single	12,556	5,486	5,830	8,457	9,902	10,203
	Married	22,317	9,621	7,561	10,143	11,414	11,859
	Widowed	15,926	51,131	4,532	7,133	7,530	12,642
	Divorced	14,900	6,260	5,708	8,070	9,191	9,358
	Separated	14,275	6,126	5,715	8,337	9,886	10,153
	Cohabiting	16,420	7,388	7,180	8,925	11,102	11,216

Note: 2004 is pre-programme, 2005 is programme entry year.

ND50+

Table A.6   New Deal 50+ benefit receipt rates for selected start dates of participation, for periods between three and 24 months after participation

Start variable	Months since programme start				
	3	6	12	18	24
cdpstart	61%	53%	43%	35%	28%
casestart	61%	53%	43%	35%	28%
initint	63%	55%	45%	37%	30%

Table A.7   New Deal 50+ average per cent of time on benefit for selected start dates of participation, for periods between three and 24 months after participation

| Start variable | Months since programme start | | | | |
	3	6	12	18	24
cdpstart	70%	63%	55%	50%	45%
casestart	70%	64%	56%	50%	45%
initint	72%	65%	57%	52%	47%

Table A.8   New Deal 50+ per cent employed for selected start dates of participation, for periods between three and 24 months after participation

| Start variable | Months since programme start | | | | |
	3	6	12	18	24
cdpstart	25%	28%	30%	31%	31%
casestart	19%	20%	22%	23%	23%
initint	21%	23%	24%	25%	26%

Table A.9   New Deal 50+ average per cent of time employed for selected start dates of participation, for periods between three and 24 months after participation

| Start variable | Months since programme start | | | | |
	3	6	12	18	24
cdpstart	23%	25%	27%	28%	29%
gwstdt	18%	19%	20%	21%	22%
nd1intdt	19%	21%	22%	23%	24%

Table A.10   New Deal 50+ mean annual taxable wages for employed (£)

| Subgroup | Type of option | Year | | | | | |
		2004	2005	2006	2007	2008	2009 (May)
	All cases	24,981	12,535	6,798	9,385	10,471	10,174
By ethnicity	Missing	23,145	11,076	6,366	8,766	9,200	9,660
	White	25,349	12,789	6,911	9,511	10,616	10,289
	Black	18,795	9,622	5,663	7,724	9,374	8,464
	Asian	23,272	10,721	5,554	8,094	9,687	9,250
	Mixed	22,493	10,910	5,820	9,449	10,745	9,667
dis	If disabled	19,041	9,233	5,561	7,568	8,442	8,378
	Not disabled	27,568	13,972	7,361	10,192	11,362	10,958

NDLP

Table A.11 New Deal Lone Parents benefit receipt rates for selected start dates of participation, for periods between three and 24 months after participation

		Months since programme start				
		3	6	12	18	24
cdpstart	All participants	48%	43%	36%	30%	24%
	With basic skills	61%	51%	45%	38%	30%
	With higher skills	62%	58%	53%	47%	38%
	With job readiness	56%	50%	42%	35%	27%
	With subsidised employment	55%	48%	40%	33%	29%
	With self employment	54%	44%	34%	27%	21%
firintdt	All participants	46%	41%	34%	28%	23%
	With basic skills	60%	50%	44%	37%	29%
	With higher skills	60%	56%	52%	45%	36%
	With job readiness	55%	49%	41%	33%	25%
	With subsidised employment	50%	42%	36%	30%	26%
	With self employment	51%	42%	34%	26%	20%
stcasedt	All participants	44%	39%	33%	28%	23%
	With basic skills	59%	48%	42%	37%	30%
	With higher skills	57%	55%	52%	45%	37%
	With job readiness	52%	46%	38%	31%	24%
	With subsidised employment	49%	43%	36%	30%	27%
	With self employment	51%	39%	32%	25%	19%

Table A.12 New Deal Lone Parents average per cent of time on benefit for selected start dates of participation, for periods between three and 24 months after participation

		Months since programme start				
		3	6	12	18	24
Cdpstart	All participants	55%	50%	45%	41%	37%
	With basic skills	65%	61%	54%	50%	46%
	With higher skills	66%	63%	59%	56%	53%
	With job readiness	61%	57%	51%	47%	43%
	With subsidised employment	63%	57%	50%	46%	42%
	With self employment	59%	54%	47%	41%	37%
firintdt	All participants	53%	48%	43%	39%	36%
	With basic skills	65%	60%	53%	49%	45%
	With higher skills	65%	61%	58%	55%	51%
	With job readiness	61%	56%	50%	46%	41%
	With subsidised employment	58%	53%	45%	41%	38%
	With self employment	59%	53%	45%	40%	35%
stcasedt	All participants	51%	46%	41%	37%	34%
	With basic skills	64%	59%	52%	48%	44%
	With higher skills	62%	59%	56%	54%	50%
	With job readiness	58%	53%	47%	43%	39%
	With subsidised employment	58%	52%	45%	41%	38%
	With self employment	56%	50%	44%	38%	34%

Table A.13 New Deal Lone Parents per cent employed for selected start dates of participation, for periods between three and 24 months after participation

Start variable		Type of option	Months since programme start				
			3	6	12	18	24
start1	cdpstart	All cases	33%	35%	37%	38%	40%
		With basic skills	26%	29%	30%	31%	35%
		With hard skills	28%	31%	34%	36%	39%
		With job readiness	26%	28%	29%	32%	34%
		With subsidised employment	36%	37%	40%	41%	41%
		With self employment	22%	22%	23%	25%	28%
start2	firintdt	All cases	27%	29%	30%	31%	32%
		With basic skills	23%	24%	26%	27%	30%
		With hard skills	25%	27%	30%	32%	34%
		With job readiness	22%	24%	25%	27%	29%
		With subsidised employment	32%	34%	35%	36%	36%
		With self employment	18%	18%	19%	20%	23%
start3	stcasedt	All cases	27%	29%	30%	31%	32%
		With basic skills	22%	23%	24%	24%	28%
		With hard skills	24%	25%	28%	29%	31%
		With job readiness	22%	22%	23%	25%	27%
		With subsidised employment	32%	32%	34%	35%	35%
		With self employment	18%	17%	18%	19%	22%

Table A.14   New Deal Lone Parents average per cent of time employed for selected start dates of participation, for periods between three and 24 months after participation

Start variable	Type of option	Months since programme start				
		3	6	12	18	24
start1 cdpstart	All cases	31%	32%	34%	35%	36%
	With basic skills	24%	26%	27%	29%	30%
	With hard skills	37%	28%	31%	32%	34%
	With job readiness	25%	26%	27%	28%	29%
	With subsidised employment	31%	33%	36%	37%	38%
	With self employment	22%	22%	22%	23%	24%
start2 gwstdt	All cases	25%	27%	28%	29%	30%
	With basic skills	21%	22%	24%	25%	26%
	With hard skills	23%	25%	27%	28%	29%
	With job readiness	21%	22%	23%	24%	25%
	With subsidised employment	29%	31%	33%	34%	34%
	With self employment	33%	18%	19%	19%	20%
start3 nd1intdt	All cases	25%	26%	28%	28%	29%
	With basic skills	20%	21%	23%	24%	24%
	With hard skills	22%	23%	25%	26%	27%
	With job readiness	20%	21%	22%	23%	23%
	With subsidised employment	29%	30%	32%	32%	33%
	With self employment	18%	18%	18%	19%	19%

Table A.15   New Deal Lone Parents mean annual taxable wages for employed (£)

Subgroup	Type of option	Year					
		2004	2005	2006	2007	2008	2009 (May)
By skill	All cases	9,238	5,475	4,523	6,659	7,840	8,285
	With basic skills	6,938	5,211	4,232	6,840	7,822	8,137
	With hard skills	9,729	5,364	4,633	6,505	8,014	8,543
	With job readiness	8,102	4,937	3,996	5,345	6,638	7,373
	With subsidised employment	7,431	4,730	4,069	6,107	7,405	7,525
	With self employment	11,963	7,131	4,814	7,335	9,105	9,530
By ethnicity	Missing	10,629	6,526	5,655	7,426	8,637	9,083
	White	8,969	5,263	4,332	6,466	7,596	8,052
	Black	10,960	6,665	5,464	8,056	9,570	9,953
	Asian	9,120	6,048	5,027	7,651	9,199	9,402
	Mixed	10,309	5,960	4,986	7,115	8,930	9,487

NDDP

Table A.16   New Deal Disabled People benefit receipt rates for selected start dates of participation, for periods between three and 24 months after participation

Start variable	Type of participation	Months since programme start				
		3	6	12	18	24
ibstd	All cases	74%	57%	37%	30%	24%
	With subsidised employment	79%	53%	29%	23%	18%
cpdstart	All cases	47%	40%	32%	25%	20%
	With subsidised employment	48%	36%	27%	21%	16%

Table A.17   New Deal Disabled People average per cent of time on benefit for selected start dates of participation, for periods between three and 24 months after participation

Start variable	Type of participation	Months since programme start				
		3	6	12	18	24
ibstd	All cases	79%	72%	59%	50%	44%
	With subsidised employment	87%	76%	57%	46%	40%
cpdstart	All cases	55%	49%	42%	38%	34%
	With subsidised employment	61%	51%	41%	35%	31%

Table A.18   New Deal Disabled People per cent employed for selected start dates of participation, for periods between three and 24 months after participation

Start variable	Type of participation	Months since programme start				
		3	6	12	18	24
contstdt	All cases	42%	41%	39%	37%	36%
	With subsidised employment	36%	37%	40%	41%	41%
ibstd	All cases	27%	27%	28%	29%	30%
	With subsidised employment	32%	34%	35%	36%	36%
cpdstart	All cases	36%	38%	40%	41%	42%
	With subsidised employment	32%	32%	34%	35%	35%

Table A.19 New Deal Disabled People average per cent of time employed for selected start dates of participation, for periods between three and 24 months after participation

Start variable	Type of participation	Months since programme start				
		3	6	12	18	24
cdpstart	All cases	42%	42%	41%	40%	39%
	With subsidised employment	46%	46%	45%	45%	44%
gwstdt	All cases	27%	27%	27%	28%	28%
	With subsidised employment	31%	31%	33%	34%	35%
nd1intdt	All cases	33%	35%	37%	38%	39%
	With subsidised employment	44%	49%	52%	54%	55%

Table A.20 New Deal Disabled People mean taxable wages for employed (£)

Subgroup	Type of participation	Year					
		2004	2005	2006	2007	2008	2009 (May)
	All cases	15,884	8,499	6,266	8,458	9,793	9,731
	With subsidised employment	16,727	9,045	6,815	9,897	11,338	11,200
by ethnicity	Missing	14,256	8,551	6,207	8,732	10,489	10,296
	White	15,977	8,530	6,311	8,507	9,820	9,746
	Black	16,380	7,755	5,356	7,281	8,726	8,862
	Asian	14,154	8,117	5,769	7,747	9,250	9,522

NDP

In the New Deal for Partners programme, the partners are usually included in the main claimant's benefit claim, so the benefit rates initially appear lowih Over time benefit rates increase, as some participants initiate independent claims.

Table A.21 New Deal Partner benefit receipt rates for selected start dates of participation, for periods between three and 24 months after participation

	Months since programme start				
Start variable	3	6	12	18	24
sfrstart	9%	11%	13%	12%	12%
ndpintat	NA	NA	NA	NA	NA
caselddt	9%	11%	13%	12%	12%

Table A.22 New Deal Partner average per cent of time on benefit for selected start dates of participation, for periods between three and 24 months after participation

	Months since programme start				
Start variable	3	6	12	18	24
sfrstart	7%	9%	10%	11%	12%
ndpintat	NA	NA	NA	NA	NA
caselddt	8%	9%	10%	11%	12%

Table A.23 New Deal Partner per cent employed for selected start dates of participation, for periods between three and 24 months after participation

	Months since programme start				
Start variable	3	6	12	18	24
sfrstart	42%	41%	39%	37%	36%
ndpintat	27%	27%	28%	29%	30%
caselddt	36%	38%	40%	41%	42%

Table A.24 New Deal Partner average per cent of time employed for selected start dates of participation, for periods between three and 24 months after participation

	Months since programme start				
Start variable	3	6	12	18	24
sfrstart	42%	42%	41%	40%	39%
ndpintat	27%	27%	27%	28%	28%
caselddt	33%	35%	37%	38%	39%

Table A.25 New Deal Partner mean annual taxable wages (£)

Subgroup	Year					
	2004	**2005**	**2006**	**2007**	**2008**	**2009 (May)**
All cases	15,884	8,499	6,266	8,458	9,793	9,731
White	15,977	8,530	6,311	8,507	9,820	9,746
Black	16,380	7,755	5,356	7,281	8,726	8,862
Asian	14,154	8,117	5,769	7,747	9,250	9,522

NDYP

Table A.26 New Deal Young People benefit receipt rates for selected start dates of participation, for periods between three and 24 months after participation

Start variable	Type of participation	Months since programme start				
		3	**6**	**12**	**18**	**24**
cpdstart	All cases	46%	33%	36%	32%	27%
	With hard skills	58%	37%	46%	42%	35%
	With subsidised employment	66%	38%	47%	40%	34%
gwstd	All cases	45%	33%	37%	32%	28%
	With hard skills	53%	35%	47%	41%	35%
	With subsidised employment	61%	36%	47%	39%	34%
opst1	All cases	45%	33%	37%	32%	28%
	With hard skills	53%	35%	47%	41%	35%
	With subsidised employment	61%	36%	47%	39%	34%

Table A.27 New Deal Young People average per cent of time on benefit for selected start dates of participation, for periods between three and 24 months after participation

Start variable	Type of participation	Months since programme start				
		3	**6**	**12**	**18**	**24**
cpdstart	All cases	56%	45%	38%	35%	37%
	With hard skills	70%	58%	49%	48%	45%
	With subsidised employment	76%	63%	53%	50%	37%
gwstd	All cases	59%	48%	42%	40%	37%
	With hard skills	66%	54%	47%	47%	44%
	With subsidised employment	72%	59%	51%	49%	46%
opst1	All cases	59%	48%	42%	40%	37%
	With hard skills	66%	54%	47%	47%	44%
	With subsidised employment	72%	59%	51%	49%	46%

Table A.28 New Deal Young People per cent employed for selected start dates of participation, for periods between three and 24 months after participation

Start variable	Type of participation	Months since programme start				
		3	6	12	18	24
cpdstart	All cases	29%	32%	34%	36%	38%
	With hard skills	16%	19%	21%	22%	24%
	With subsidised employment	23%	27%	32%	35%	37%
gwstd	All cases	26%	28%	30%	32%	33%
	With hard skills	16%	19%	21%	22%	24%
	With subsidised employment	24%	28%	32%	35%	37%
opst1	All cases	26%	28%	30%	32%	33%
	With hard skills	16%	19%	21%	22%	24%
	With subsidised employment	23%	28%	32%	35%	37%

Table A.29 New Deal Young People average per cent of time employed for selected start dates of participation, for periods between three and 24 months after participation

Start variable	Type of participation	Months since programme start				
		3	6	12	18	24
cpdstart	All cases	26%	29%	32%	33%	35%
	With hard skills	20%	21%	23%	25%	27%
	With subsidised employment	22%	23%	27%	29%	31%
gwstd	All cases	23%	26%	28%	29%	30%
	With hard skills	20%	21%	24%	26%	28%
	With subsidised employment	22%	24%	28%	30%	32%
opst1	All cases	23%	26%	28%	29%	30%
	With hard skills	20%	21%	24%	26%	28%
	With subsidised employment	22%	24%	28%	30%	32%

Table A.30 New Deal Young People mean annual taxable wages (£)

Subgroup	Type of option	2004	2005	2006	2007	2008	2009 (May)
By skill	All cases	7,191	4,287	3,698	6,700	8,328	8,928
	With hard skills	5,790	3,839	2,750	4,733	6,780	7,575
	With subsidised employment	8,359	3,996	3,080	5,449	7,362	8,256
By ethnicity	Missing	7,314	4,324	3,775	6,809	8,791	9,225
	White	7,266	4,305	3,717	6,711	8,323	8,865
	Black	6,207	4,033	3,527	6,195	7,506	8,447
	Asian	6,939	4,061	3,556	7,087	8,894	9,951
	Mixed	6,948	4,452	3,612	6,428	7,938	8,814
By marital status	Unknown	9,250	6,649	7,224	8,893	10,537	11,269
	Single	7,563	4,370	4,036	7,105	8,634	9,289
	Married	7,049	4,215	3,659	6,669	8,307	8,927
	Widowed	10,663	5,777	4,271	7,259	8,951	8,701
	Divorced	Not enough observations					
	Separated	6,829	4,908	4,170	4,675	7,771	6,866
	Cohabiting	8,313	4,794	3,256	6,021	6,980	7,989

Appendix B
Source data tables

Table B.1   Source Data: Benefits

Path	Dataset/programme	SAS name	Variable contents	Relevant population
Benefit duration (to be constructed based on the following variables)				
/secure/DSU2b/project/nbd/published7	nbd	CDSTART	Date benefit claim started – corrected as part of the overlapping claims processing.	All benefit claimants (for the benefits covered in NDB
/secure/DSU2b/project/nbd/published7	nbd	CDEND	Date benefit claim ended – corrected as part of the overlapping claims processing.	All benefit claimants
/secure/DSU2b/project/nbd/published7	nbd	CDESTND	Date benefit claim ended (randomly allocated between two scan periods for non-JSA benefits)	All benefit claimants
/secure/DSU5/project/longitudinal/P45/database2/enc	hmrc P45	BENFLAG	Identifies records which are for benefit claims (JSA or IB)	All sampled in HMRC (presumably all persons who have ever been on benefit as NBD seems to be the basic selection)
Type of benefit				
/secure/DSU2b/project/nbd/published7	nbd	CDSTART	Date benefit claim started – corrected as part of the overlapping claims processing.	All benefit claimants
/secure/DSU2b/project/nbd/published7	nbd	CNBDCTYP	Type of JSA benefit claim	JSA claimants
/secure/DSU2b/project/nbd/published7	nbd	CXBENTYP	Type of benefit	All benefit claimants
Personal characteristics				
/secure/DSU2b/project/nbd/published7	nbd	CCCUSTBD	Record type	All
/secure/DSU2b/project/nbd/published7	nbd	CCSEX	Sex of claimant	All
/secure/DSU2b/project/nbd/published7	nbd	CDDOB	Claimants date of birth	All
/secure/DSU2b/project/nbd/published7	nbd	CLIENTEN	Client group that the claimant belonged to at the end of their claim	All
/secure/DSU2b/project/nbd/published8	nbd	CLIENTST	Client group that the claimant belonged to at the start of their claim	All
/secure/DSU2b/project/nbd/published7	nbd	CNAGEST	Age at the start of current claim	All
/secure/DSU2b/project/nbd/published8	nbd	CNEMPIND	Flag to indicate if claimant is in employment for current claim	ICA
/secure/DSU2b/project/nbd/published9	nbd	CNBDPART	Number of part-time hours worked	IS, PC, JSA
/secure/DSU2b/project/nbd/published10	nbd	CNBDPINC	Most recent amount of part-time income	IS, PC, JSA

Continued

Table B.1   Continued

Path	Dataset/ programme	SAS name	Variable contents	Relevant population
/secure/DSU2b/project/nbd/published11	nbd	ISDIS	Indicates whether a Disability Premium has ever been claimed during an IS benefit claim	IS
/secure/DSU2b/project/nbd/published7	nbd	CNBDPART	Number of part-time hours worked	IS, PC, JSA
/secure/DSU2b/project/nbd/published7	nbd	CNBDPINC	Most recent amount of part-time income	IS, PC, JSA
Family/dependants				
/secure/DSU2b/project/nbd/published7	nbd	ISLP	Indicator of a lone parent for IS benefit claims	IS
/secure/DSU2b/project/nbd/published7	nbd	JCPAREA	Pre 2004/05 JCP district codes (derived from 2005 LA codes)	All
/secure/DSU2b/project/nbd/published7	nbd	CXBDWORK	Flag to indicate whether claimant or partner works	IS, PC
/secure/DSU2b/project/nbd/published7	nbd	DDYCH_EN	Date of birth of youngest child at end of benefit claim	IS, JSA
/secure/DSU2b/project/nbd/published7	nbd	DDYCH_ST	Date of birth of youngest child at start of benefit claim	IS, JSA
/secure/DSU2b/project/nbd/published7	nbd	DNCHI_EN	Number of children at end of claim	IS, PC, JSA, ICA, IB, SDA, PIB, WB, BB, RP
/secure/DSU2b/project/nbd/published7	nbd	DNCHI_ST	Number of children at start of claim	IS, PC, JSA, ICA, IB, SDA, PIB, WBn, BB, RP
/secure/DSU2b/project/nbd/published8	nbd	DNPARTNREN	Flag that indicates whether claimant had a partner at the end of the benefit claim	IS, PC, JSA
/secure/DSU2b/project/nbd/published9	nbd	DNPARTNRST	Indicates whether the claimant had a partner at the start of the benefit claim	PC, IS, JSA

Continued

Table B.1   Continued

Geographic indicators

Path	Dataset/ programme	SAS name	Variable contents	Relevant population
/secure/DSU2b/project/nbd/published7	nbd	CCBRANEN	Branch office number at the end of the claim	All
/secure/DSU2b/project/nbd/published7	nbd	CCBRANST	Branch office number at start of benefit claim	All
/secure/DSU2b/project/nbd/published7	nbd	CNBDPONO	Post office number	IS, PC, JSA
/secure/DSU2b/project/nbd/published7	nbd	GMS_ PCODEEN	Post code at the end of the claim from GMS extracts	All
/secure/DSU2b/project/nbd/published7	nbd	GMS_PCODEST	Post code at the start of the claim from GMS extracts	All
/secure/DSU2b/project/nbd/published7	nbd	JCPAREA0405	2004/05 JCP district codes (derived from 2005 LA codes)	All
/secure/DSU2b/project/nbd/published7	nbd	JCPAREA0506	2005/06 JCP district codes (derived from 2005 LA codes)	All
/secure/DSU2b/project/nbd/published7	nbd	JCPAREA0607	2006/07 JCP district codes (derived from 2005 LA codes)	All
/secure/DSU2b/project/nbd/published7	nbd	JCPAREA0809	2008/09 JCP district codes (derived from 2005 LA codes)	All
/secure/DSU2b/project/nbd/published7	nbd	JCPEN	JCP cluster at end of claim allocated using the postcode at the end of the claim	All
/secure/DSU2b/project/nbd/published7	nbd	JCPST	JCP cluster at start of claim allocated using the postcode at the start of the claim	All
/secure/DSU2b/project/nbd/published7	nbd	LA_CODE	2005 LA Code – allocated based on the postcode at the end of the claim	All
/secure/DSU2b/project/nbd/published7	nbd	WFIEN	WFI cluster at end of claim	All
/secure/DSU2b/project/nbd/published7	nbd	WFIST	WFI cluster at start of claim	All
/secure/DSU2b/project/nbd/published7	nbd	ONE_EN	One marker at end of claim. The value of branch office at the end of the claim is used to derive the one area which the branch belongs to	All

Table B.1   Continued

Path	Dataset/ programme	SAS name	Variable contents	Relevant population
/secure/DSU2b/project/nbd/published7	nbd	ONE_ST	One marker at start of claim. The value of branch office at the start of the claim is used to derive the one area which the branch belongs to	All
Benefit payments				
secure/DSU2b/project/nbd/published9	nbd	AMOUNTEN	Weekly amount of benefit in payment at the end of the benefit claim or at the time of the most recent update if the claim is still live	All
secure/DSU2b/project/nbd/published10	nbd	AMOUNTST	Weekly amount of benefit in payment at the start of the benefit claim	All
secure/DSU2b/project/nbd/published12	nbd	CNBDAMT12	Amount of one parent benefit in payment	IS, PC, JSA
secure/DSU2b/project/nbd/published13	nbd	CNBDAMT40	Amount of occupational pension	IS, PC, JSA
secure/DSU2b/project/nbd/published14	nbd	CNBDAMT61	Amount if IB in payment	IS, PC, JSA
secure/DSU2b/project/nbd/published17	nbd	CXBDHB	Flag to indicate if claimant is in receipt of Housing Benefit	IS, PC, JSA
secure/DSU2b/project/nbd/published18	nbd	CXBDLOAN1	Housing loan type 1	IS, PC, JSA
secure/DSU2b/project/nbd/published19	nbd	CXBDLOAN2	Housing loan type 2	IS, PC, JSA

Table B.2   Source Data: Benefit duration and type

Path	Dataset/ programme	SAS name	Variable contents
Benefit duration (to be constructed based on the following variables)			
/secure/DSU2b/project/nbd/published7	nbd	CDSTART	Date benefit claim started – corrected as part of the overlapping claims processing
/secure/DSU2b/project/nbd/published7	nbd	CDEND	Date benefit claim ended – corrected as part of the overlapping claims processing
/secure/DSU2b/project/nbd/published7	nbd	CDESTND	Date benefit claim ended (randomly allocated between two scan periods for non-JSA benefits)
/secure/DSU5/project/longitudinal/P45/database2/enc	hmrc P45	BENFLAG	Identifies records which are for benefit claims (JSA or IB)
Type of benefit			
/secure/DSU2b/project/nbd/published7	nbd	CDSTART	Date benefit claim started – corrected as part of the overlapping claims processing
/secure/DSU2b/project/nbd/published7	nbd	CNBDCTYP	Type of JSA benefit claim
/secure/DSU2b/project/nbd/published7	nbd	CXBENTYP	Type of benefit
Destination when leaving benefit			
/secure/DSU2b/project/nbd/published7	nbd	DESTCODE	Destination code – reason the person left benefit

Table B.3   Source Data: HMRC employment

Path	Dataset/ programme	SAS name	Variable contents
Wages			
/secure/DSU5/project/longitudinal/P14/database2/enc	hmrc P14	CCORCID	Unique identifier for DWP customer (database 2 only)
/secure/DSU5/project/longitudinal/P14/database2/enc	hmrc P14	NINO	National Insurance number. Only on DB2
/secure/DSU5/project/longitudinal/P14/database2/enc	hmrc P14	EMPLOYMENT_ INSTANCE_ID	Employment instance identifier assigned by HMRC
/secure/DSU5/project/longitudinal/P14/database2/enc	hmrc P14	TOTAL_PAY	Total pay for the tax year in pounds as recorded on the P14.
/secure/DSU5/project/longitudinal/P14/database2/enc	hmrc P14	TOTAL_TAX	Total tax for the tax year in pounds as recorded on the P14. NB: this value is received as pence but converted to pounds during processing
/secure/DSU5/project/longitudinal/P14/database2/enc	hmrc P14	PREVIOUS_TAX	Pay received in previous employments in pounds for the tax year as recorded on the P14. NB: this value is received as pence but converted to pounds during processing
/secure/DSU5/project/longitudinal/P14/database2/enc	hmrc P14	PAY_THIS_ EMPLOYMENT	Calculation, Total_Pay – Previous_Pay. This is calculated by HMRC and not during our processing
/secure/DSU5/project/longitudinal/P14/database2/enc	hmrc P14	TAX_THIS_ EMPLOYMENT	Calculation, Total_Tax – Previous_Tax. This is calculated by HMRC and not during our processing
/secure/DSU5/project/longitudinal/P14/database2/enc	hmrc P14	DATE_OF_ STARTING	The start date of this employment as recorded on the P14 or P60. Note: the start date is only entered on the P14/P60 by the employer if it falls within the year to which the return relates
/secure/DSU5/project/longitudinal/P14/database2/enc	hmrc P14	DATE_OF_ LEAVING	The end date of this employment as recorded on the P14 or P60. Note: the end date is only entered on the P14/P60 by the employer if it falls within the year to which the return relates
/secure/DSU5/project/longitudinal/P14/database2/enc	hmrc P14	TAX_CREDITS	Tax credits paid in this employment in pounds as recorded on the P14. NB: this value is received as pence but converted to pounds during processing

Continued

Table B.3 Continued

Employment outcomes

Path	Dataset/ programme	SAS name	Variable contents
/secure/DSU5/project/longitudinal/P45/database2/enc	hmrc P45	CCORCID	Unique identifier for DWP customer (database 2 only)
/secure/DSU5/project/longitudinal/P45/database2/enc	hmrc P45	EMPLOYMENT_ START_DATE	Start of an employment record. Dates of 6th April are largely artificial, missing dates are blank
/secure/DSU5/project/longitudinal/P45/database2/enc	hmrc P45	EMPLOYMENT_ END_DATE	End of an employment record. Dates of 5th April are largely artificial, missing dates are blank, live employments are coded as '31stDec9999'
/secure/DSU5/project/longitudinal/P45/database2/enc	hmrc P45	BENFLAG	Identifies records which are for benefit claims (JSA or IB).
/secure/DSU5/project/longitudinal/P45/database2/enc	hmrc P45	DATE_OF_ EXTRACT	Numeric version of extract variable (above). Indicates when updates have been made to the databases. Any records with date_of_extract=190001 were received in the cut-over data, i.e. one-off download of historical records received in April 2004
/secure/DSU5/project/longitudinal/P45/database2/enc	hmrc P45	DATE_OF_ DOWNLOAD	Month of the P45/P60 download that the specific job was first received (i.e. if this record is an amendment to a job that was first received in Jun2005, then date_of_download is 200506, i.e. June05 regardless of when the amendment was received)
/secure/DSU5/project/longitudinal/P45/database2/enc	hmrc P45	FIRST_ ENDDATE	Month of the P45/6P0 download that the specific job first received an end date (i.e. employment end date not equal to missing or 31Dec9999)

Appendix C
Quantile treatment effects

This section discusses how the quantile treatment effects can be produced econometrically.

Empirical strategies for quantile treatment effects

A quantile treatment effect can be regarded as the difference across the treatment status in two outcome quantiles (Bitler *et al.,* 2003: 16). In the absence of problems of selection bias, such as in an ideal experimental allocation of treatment and control group, the 50 per cent quantile effect (the median or quantile treatment effect of the .5 quantile) would be defined as the sample median of the treatment group subtracted from the sample median of the control group. Other quantile effects can be estimated accordingly (ibid.). It is important to note that a causal model can only estimate the quantile treatment effects, but it will not be able to estimate quantiles of the treatment effect distribution (ibid.).

In the absence of **ideal** experimental data[51], we have the problem that we can only observe the outcome for programme participants and non-participants. Since the programme participants differ in terms of observable and unobservable characteristics from the non-participants, a simple comparison of particular distributional outcomes between both groups does not represent an unbiased estimate of the programme effect on the particular part of the distribution.

Just as non-experimental evaluation studies estimate the mean effect of a programme on a particular group of participants, the evaluation of outcomes for particular parts of the distribution has to estimate a non-programme outcome for the participants. The distribution of the counterfactual control outcome can never be observed for the treatment group.

Conditional Independence Assumption

The first strand of methods applied in order to estimate the distribution of the non-treatment outcome is based on a Conditional Independence Assumption (CIA). This implies that one can estimate the outcome without participation irrespective of whether persons were actually participants or non-participants as long as non-participants are similar to the participants with respect to observable characteristics.

This assumption is only plausible if data are informative and allow the observation of characteristics that affect both the decision to participate and the outcomes. If satisfied, a group of non-participants similar in characteristics allows estimation of the distribution of the non-participation outcome for the participants.

A huge disadvantage for the CIA is the 'curse-of-dimensionality', as it might be difficult to find similar non-participants with respect to a large set of observable characteristics. Therefore, most evaluation studies use the result by Rosenbaum and Rubin (1983) that the CIA also holds with respect to the probability of treatment ('propensity score') as a function of all observable characteristics.

---

[51]   In practice, most experimental data will require the non-experimental estimation methods.

In empirical applications for welfare-to-work programmes in the USA, Bitler *et al.,* 2003 propose to implement an inverse propensity score weighting of quantile regressions. Similarly, Firpo (2007) proposes a framework of estimating quantile treatment effects based on a Conditional Independence Assumption. He implements a weight by the inverse of a nonparametric estimate of the propensity score. This was first implemented by Hirano, Imbens and Ridder (2003) for Average Treatment Effects, who show that the procedure can leads to an efficient estimate of the average treatment effect.

Implementing propensity score weighting would be a feasible strategy in order to estimate the effect of the British welfare-to-work policies on particular parts of the distribution, e.g. on the very short and very long durations of unemployment as there are often informative survey data available that have been used for earlier evaluation studies implementing propensity score matching (e.g. the econometric evaluation of the New Deal for Lone Parents by De Azevedo, Dolton and Smith 2006). These can provide the very wide set of characteristics that are used in estimation to satisfy the credibility of the CIA.

Instruments

A second means of estimating quantile treatment effects by Chernozhukov and Hansen (2005) extends the non-experimental framework established in Heckman and Robb (1986). As for the models implementing an inverse propensity score weighting, this paper first makes an identifying assumption how a particular policy can affect features of the distribution of potential outcomes, e.g. the quantiles of potential outcomes under various treatments, conditional on observed characteristics. This particular distributional outcome is denoted as the 'quantile treatment response function', and the quantile treatment effect summarises the differences in the impact of treatments on the quantiles of potential outcomes. By integrating over the various quantile effects, the average effects can be defined.

The problem of estimating either one of these is again the difference in observable characteristics between participants and non-participants inducing selection bias, so that conventional quantile regressions of observed **outcomes** are inappropriate for measuring the impact and estimating the quantile treatment effect.

To solve this, Chernozhukov and Hansen (2005) propose estimating quantile treatment effects using **instruments**, which are variables that affect the participation in the particular programme, but which are independent of potential outcomes.

One may think of instruments as representing some unobserved characteristics about participants and non-participants, e.g. ability, and can control for that unobserved information which can affect outcomes and be correlated with participation. Such an instrument is interesting for describing and learning the structure of differing treatment effects and controlling for unobserved information about participants. They present an example of an instrumented analysis of the returns to schooling, where they treat schooling as endogenous to the wage outcome, and the instrument (weak) is the variation in the geographic measure of proximity to the school which proxies for years of education in a quantile regression of wages on education, which they used to uncover the variation in the returns to schooling.

The estimation of quantile treatment effects assumes that, given informative characteristics as covariates, an instrument exists which is related to participation in the programme, but not to outcomes. In other words, this crucially depends on the availability of instruments, which can reflect exogenous variation affecting programme participation (e.g. the distance to the job centre or being part of a programme pilot) or ultimately the random assignment status in a controlled experiment. Finding measured data items which can act as instruments is not straightforward.

Similarly, Froelich and Melly (2008) assume the availability of instruments as a potential strategy for estimating quantile treatment effects. However, in this case, the instrumental variable may only really affect a narrow group of all participants in their potential outcomes and potential treatment status, and not the entire distribution. As Chernozhukov and Hansen (2005), they assume independence given available instruments, but only close to a threshold affecting the participation status. The resulting quantile treatment effect would only be an effect for participants (and non-participants) close to such a threshold (a 'local' treatment effect), extending the concept of Imbens and Agrist (1994) to quantile treatment effects. Note that this means that the impact is then limited in interpretation.

Quantiles of the treatment effect?

While **quantile treatment effects can be identified, quantiles of the treatment effects** are another item and these cannot be identified without further knowledge or assumptions about the joint distribution. The following two assumptions exist in real-world studies on the identification of the quantiles of the treatment effect:

1   A distribution fully identified by the mean impact: Bitler *et al.* (2003) refer to an example how the treatment effects are recovered. An assumption could be that the treatment effects is equal for all observations and fully identified by the mean impact, i.e. the cumulative distribution of treatment effects is fully characterised by a mean and supports only that one value. This is very unlikely for a real-world distribution of unemployment duration, for example.

2   Rank preservation: Any person whose outcome in the counterfactual control distribution is the qth quantile will also have an outcome that is the qth quantile in the counterfactual treatment distribution. Then, the quantile treatment effect is the quantile of the treatment effect (see Bitler *et al.,* 2003: 17).

Since the quantile treatment effect is always identified by the difference of marginals at quantile q, it follows that cumulative distribution of treatment effects is also identified by sorting the set of estimated differences. However this sorting may allow the rank preservation to hold for a large proportion of the distribution, but there are parts like the bottom of the distribution where the rank preservation fails. Quantile treatment effects can be seen as differences in the treated and comparison distribution, not the treatment effects for identifiable persons in either distribution.

In the case where rank preservation fails, there are still important features of the joint distribution which can be identified, like the mean treatment effect, see Bitler *et al.* (2003) for details. If this is the objective of interest, the marginal distributions are just as informative as the joint distribution. Even with heterogeneous effects some important features of the joint distribution can be identified, depending on the estimated quantile treatment effects.

They find there are a number of results that hold:

1   For any particular quantile, the minimum treatment effect is no larger than the smallest quantile treatment effect. If we find a negative quantile treatment effect, at least one treatment effect must also be negative.

2   The same holds for negative treatment effects: If any quantile treatment effect is positive, at least one treatment effect must also be positive.

3   The variance of the distribution of treatment effects is at least as great as the variance of the quantile treatment effects.

If subgroups are defined with respect to characteristics that are either permanent or fixed over the period of study, the above results hold within subgroups, which may yield further information (for example, the maximum quantile treatment effect may be greater in a subgroup than in the pooled sample).

Appendix D
Information regarding DWP datasets

(Supplied by the Department for Work and Pensions (DWP) November 2008, sourced from the DWP intranet.)

Programme databases

New Deal 50 plus

New Deal 50 plus (ND50+) is a voluntary programme where customers aged 50 and over can receive specialist advice in assisting them back into work.

New Deal for Disabled People

New Deal for Disabled People (NDDP) offers support to help people with health conditions and disabilities to find and keep jobs. NDDP is voluntary and offers eligible people access to a network of Job Brokers from private, voluntary and public sector organisations. A number of datasets are derived from the NDDP data:

- NDDP Registrants – This dataset includes all clients who volunteer for NDDP.

- NDDP Referrals (to Job Brokers) – Includes the details of customers Jobcentre Plus have referred to Job Brokers in order to match with registrants to evaluate take-up rates. The data is used in the evaluation of NDDP and in the production of National Statistics.

- NDDP LABOUR MARKET SYSTEM (LMS) Jobs – Includes details of NDDP customers who have a job outcome recorded on the Jobcentre Plus benefit system.

- NDDP Eligible Population. This extract is used to identify all customers who are eligible for a NDDP Gateway interview, who are then contacted.

New Deal for Lone Parents

New Deal for Lone Parents (NDLP) aims to help lone parents to overcome barriers into work and improve their job readiness. NDLP was introduced as a voluntary programme nationally in October 1998 for those lone parents claiming Income Support (IS). It was extended in November 2001 for all lone parents who are not working or who are working less than 16 hours per week.

New Deal for Long Term Unemployed

New Deal for Long Term Unemployed (NDLTU) was introduced nationally on 29 June 1998 for those who had been claiming Jobseeker's Allowance (JSA) for at least two years. From April 2001, NDLTU was extended and enhanced to provide clients with access to a greater and more tailored range of support and provision. Eligibility has been extended to include those who had been claiming JSA for 18 months or more.

New Deal for Partners

New Deal for Partners (NDP) was re-launched in April 2004. The programme offers non-working partners of working-age persons claiming JSA (excluding Joint Claims couples), IS, Incapacity Benefit (IB), Severe Disablement Allowance (SDA) or Carer's Allowance (CA) a similar level of help and support currently available to lone partners through the NDLP programme. This dataset contains all partners that have gone through the NDP programme.

New Deal for Young People

New Deal for Young People (NDYP) was introduced nationally on 6 April 1998 and is aimed at those aged 18-24 who have been claiming JSA for at least six months. Eligible clients will progress through a Gateway period, lasting up to four months, where advisers will work to improve employability and to find unsubsidised jobs for as many as possible. Those who do not find a job will move into one of four options: Subsidised employment, Full-time education/training, Environmental Task Force, or Voluntary Sector jobs. Clients' reaching the end of their option without obtaining a job enter a follow-through period, where they receive intensive help to find a job.

The Master Index

The Master Index is an individual-level spells database that brings together the historical programme evaluation databases from the Employment Service and the 100 per cent National Benefit Databases (NBD), previously known as the Working Age Statistical Database (WASD). It aims to provide, for the first time, easy cross benefit and programme analysis. The Master Index is the primary DWP feed for the Work and Pensions Longitudinal Study (WPLS) and the Job Outcome Target (JOT) data.

The Master Index covers the main New Deal programmes and further programmes, pilots and pathfinders. The dataset also contains ten benefits taken from NBD and also JOT spells data. The main variables are a unique identifier (ORCID), start and end dates of spells and the benefit/programme type for that period.

Work and Pensions Longitudinal Study

From January 2004, the WPLS links benefit and programme information held by DWP on its customers, with employment records from Her Majesty's Revenue and Customs (HMRC). New data-sharing provisions introduced in the Employment Act 2002 have opened the way for DWP to receive more data on employment from HMRC and use the information for more purposes. DWP and HMRC have been working together to enable this data sharing to take place and to develop safeguards for the initiative.

The P14 and P45 data dictionaries contain details of the variables for each dataset, plus frequency counts of the decode values for each variable.

The 100 per cent National Benefit Databases

Formerly known as the WASD. These databases were created to evaluate the impact the introduction of Jobcentre Plus offices had on the Labour Market. The series of databases contain information about clients' claims and spells on the main DWP benefits from June 1999 and is updated every month. The data comes from the different benefits systems and covers the following benefits – JSA, IS, IB, SDA, ICA (CA), PIB, Widows Benefit (WB)/Bereavement Benefit (BB), Disability Living Allowance, Pension Credit, Retirement Pension, and Attendance Allowance.

- The 'National Database' datasets holds claim-level information on what benefits a person has claimed (Great Britain only).

- The **'Spells Database'** holds a record for each spell someone has had on benefit (i.e. one or more consecutive benefit claims).

- The **'Client Group History Table'** contains a complete history of which client groups a person has belonged to during their time on benefit. The datasets also records what their previous and next client group was.

References

Abbring, J., and Heckman, J. (2007). Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation, in Heckman, J. and Leamer, E. (ed.), *Handbook of Econometrics*, edition 1, volume 6, chapter 72 Elsevier.

Abbring, J., and van den Berg, G. (2003). The non–parametric identification of treatment effects in duration models, *Econometrica*, 71, 1491–1517.

Adam, S., Bozio, A., Emmerson, C., Greenberg, D., and Knight, G. (2008). *A cost-benefit analysis of Pathways to Work for new and repeat incapacity benefit claimants*. Department for Work and Pensions (DWP) Research Report No. 498. Leeds: Corporate Document Services.

Bergemann A., Fitzenberger, B., and Speckesser, S. (2009). Evaluating the Dynamic Employment Effects of Training Programs in East Germany Using Conditional Difference-in-Differences, *Journal of Applied Econometrics*, vol. 24, issue 5, pages 797-823.

Bergemann A., Fitzenberger, B., and Speckesser, S. (2004). Evaluating the Dynamic Employment Effects of Training Programs in East Germany Using Conditional Difference-in-Differences, *IZA Discussion Paper* 1648, Bonn: Institute for the Study of Labor (IZA).

Bergemann A., Fitzenberger, B., Schultz, B., and Speckesser, S. (2000). Multiple Active Labor Market Policy Participation in East Germany: An Assessment of Outcomes, *Applied Economics Quarterly*, 51, 195-244.

Bewley, H., Dorsett, R., and Haile, G. (2007). *The impact of Pathways to Work*, DWP Research Report No. 435. DWP, Norwich.

Bewley, H., Dorsett, R., and Salis, S. (2009a). *The impact of Pathways to Work on work, earnings and self-reported health in the April 2006 expansion areas*. Research Report No. 601. Norwich: DWP.

Bewley, H., Dorsett, R., and Salis, S. (2009b). *The impact of Pathways on benefit receipt in the expansion areas*. DWP Research Report No. 552, Norwich: DWP.

Biewen, M., Fitzenberger, B., Osikominu, A., and Waller, M. (2007). *Which Program for Whom? Evidence on the Comparative Effectiveness of Public Sponsored Training Programs in Germany*, ZEW Discussion Paper 07-042, Center for European Economic Research.

Bitler, M., Gelbach, J., and Hoynes, H. (2003). *What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments*, NBER Working Papers 10121, Cambridge: National Bureau of Economic Research.

Chernozhukov, V., and Hansen, C. (2005). An IV Model of Quantile Treatment Effects, *Econometrica*, 73, 245-261.

Cox, D. (1972). Regression Models and Life Tables (with Discussion), *Journal of the Royal Statistical Society*, Series B 34:187-220.

Dolton, P., de Azevedo, J. P., and Smith, J. (2006). *The econometric evaluation of the New Deal for Lone Parents*. DWP Research Report No. 356. DWP, Norwich.

Dorsett, R., Campbell-Barr, V., Hamilton, G., Hoggart, L., Marsh, A., Miller, C., Phillips, J., Ray, K., Riccio, J., Rich, S., and Vegeris, S. (2007). *Implementation and first-year impacts of the UK Employment Retention and Advancement (ERA) demonstration*, DWP Research Report No. 412.

Firpo, S. (2007). Efficient Semiparametric Estimation of Quantile Treatment Effects, *Econometrica*, 75, 259-276.

Fitzenberger B., and Speckesser, S. (2007). Employment Effects of the Provision of Specific Professional Skills and Techniques in Germany, E*mpirical Economics*, 32, 530-573.

Fitzenberger, B., Osikominu, A., and Völter, R. (2006, 2009). Get Training or Wait? Long-Run Employment Effects of Training Programs for the Unemployed in West Germany, IZA Discussion Paper No. 2121, ZEW Discussion Paper No. 06-039 forthcoming in *Annales d'Economie et Statistique*.

Fredriksson, P., and Johansson, P. (2003). *Program Evaluation and Random Program Starts*, Working Paper 2003:1 (revised), Department of Economics, Uppsala: Uppsala University.

Froelich, M., and Melly, B. (2008). *Quantile treatment effects in the regression discontinuity design*, IZA Discussion paper 3638, Bonn: IZA.

Greenberg, D., and Davis, A. (2007). *Evaluation of the New Deal for Disabled People: The cost and cost-benefit analyses.* DWP Research Report No. 431. Leeds: CDS. http://www.dwp.gov.uk/asd/asd5/rports2007-2008/rrep431.pdf

Greenberg, G., and Knight, G. (2007). *Review of the DWP Cost Benefit Framework and how it has been applied*, DWP Working Paper No. 40, DWP.

Heckman, J., and Robb, R. (1986). Alternative Methods For Solving The Problem of Selection Bias in Evaluating The Impact of Treatments on Outcomes, in Wainer. H. (ed.) *Drawing Inference From Self Selected Samples*, Heidelberg: Springer (reprinted 2001).

Hirano, K., Imbens, G. W., and Ridder, G. (2003). Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score, *Econometrica*, 71, 1161-1189.

Horowitz, J., and Mammen, E. (2004). Nonparametric Estimation of an Additive Model with a Link Function, *Annals of Statistics*, 32, 2412-2443.

Hotz, J., Imbens, G. W., and Klerman, J. (2006). Evaluating the Differential Effects of Alternative Welfare-to-Work Training Components: A Re-Analysis of the California GAIN Program, *Journal of Labor Economics*, 24(3), July 2006, 521-566.

Hujer, R., Maurer, K., and Wellner, M. (1997a). *Estimating the Effect of Training on Unemployment Duration in West Germany, A Discrete Hazard–Rate Model with Instrumental Variables*, Frankfurter Volkswirtschaftliche Diskussionsbeiträge, 73, Frankfurt: Johann Wolfgang Goethe University.

Imbens, G., and Angrist, J. (1994). Identification and Estimation of Local Average Treatment Effects, *Econometrica*, 62(2), 467-75.

Knight, G., Speckesser, S., Smith, J., Dolton, P., and de Azevedo, J. P. (2006). *Lone parents Work Focused Interviews/New Deal for Lone Parents: combined evaluation and further net impacts*, DWP Research Report No. 368. http://www.dwp.gov.uk/asd/asd5/rports2005-2006/rrep368.pdf

Koenker, R., and Bilias, Y. (2001). Quantile Regression for Duration Data: A Reappraisal of the Pennsylvania Reemployment Bonus Experiments, *Empirical Economics* 26, 199–220.

Lalive, R. (2007). How Do Extended Benefits Affect Unemployment Duration? A Regression Discontinuity Approach, *Journal of Econometrics*, 2007, 142 (2), 785-806.

Lechner, M. (2001). Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption, in: Lechner, M., F. Pfeiffer (eds), *Econometric Evaluation of Labour Market Policies*, Heidelberg: Physica, 43-58.

Lechner, M., and Miquel, R. (2002). *A Potential Outcome Approach to Dynamic Programme Evaluation – Part I: Identification*, Discussion Paper 2001-07, Department of Economics, University of St. Gallen (Version April 2002).

Lechner, M., Miquel, R., and Wunsch, C. (2005a). *Long-Run Effects of Public Sector Sponsored Training in West Germany*. IZA Discussion Paper No. 1443.

Lechner, M., and Wiehler, S. (2007). *Does the Order and Timing of Active Labour Market Programmes Matter?*, CEPR Discussion Papers 6521, C.E.P.R. Discussion Papers http://ideas.repec.org/p/cpr/ceprdp/6521.html

Maier, M. (2008). *Three Essays on Semiparametric Econometric Evaluation Methods, Inaugural dissertation*, University of Constance, downloadable: URL: http://www.ub.uni-konstanz.de/kops/volltexte/2008/5685/

Powell, J. (1986). Censored regression quantiles, *Journal of Econometrics*, 32(1), 143-155.

Riccio, J., Bewley, H., Campbell-Barr, V., Dorsett, R., Hamilton, G., Hoggart, L., Marsh, A., Miller, C., Ray, K., and Vegeris, S. (2008). *Implementation and second-year impacts for lone parents in the UK Employment Retention and Advancement (ERA) demonstration*, DWP Research Report No. 489.

Scioch,P. (2010). *The Impact of Cleansing Procedures for Overlaps on Estimation Results – Evidence for German Administrative Data*, FDZ method report 03/2010 http:doku.iab.de/fdz/reporte/2010/MR_04-10-EN.pdf

Sianesi, B. (2004). An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s, *Review of Economics and Statistics* 86, 133-155

Speckesser, S., and Bewley, H. (2006). The longer term outcomes of Work-Based Learning for Adults: Evidence from administrative data, DWP Research Report No. 390.

Thomas, A. (2008). *Evaluating the econometric evaluations of active labour market programmes using administrative data: evidence from the Jobseeker's Allowance pilots*. DWP Working Paper No. 50.

General equilibrium review references

Adam, S., Bozio, A., Emmerson, C. Greenberg, D. and Knight, G. (2008). *A cost-benefit analysis of Pathways to Work for new and repeat incapacity benefits claimants*, DWP Research Report No. 498, Norwich: DWP.

Bartik, T. J. (2002). Spillover Effects of Welfare Reforms in State Labor Markets. *Journal of Regional Science*, 42(4), 667-701.

Belleville, A. (2001). L'utilisation des Aides a L'emplei par les Entreprises: Permanence ou Logique Conjoncturelle? *Premieres Syntheses*, 25(1), 1-11.

Blundell, R., Dias, M. C., Meghir, C., and Van Reenen, J. (2004). Evaluating the Employment Impact of a Mandatory Job Search Program. *Journal of the European Economic Association*, 2(4), 569-606.

Breen, R., and Halpin, B. (1989). *Subsidising Jobs: An Evaluation of the Employment Incentive Scheme*. The Economic and Social Research Institute, Dublin.

Calmfors, L., Forslund, A., and Hemstrom, M. (2002). *The effects of active labour market policies in Sweden: What is the evidence?*

Calmfors, L., and Skedinger, P. (1995). Does active labour-market policy increase employment? Theoretical considerations and some empirical evidence from Sweden. *Oxford review of economic policy*, 11(1), 91-109.

Card, D., and Robins, P. K. (2005). How important are 'entry effects' in financial incentive programs for welfare recipients? Experimental evidence from the Self-Sufficiency Project. *Journal of Econometrics*, 125(1-2), 113-139.

Chang, F. (1996). *Evaluating the Impact of Mandatory Work Programs on Two-Parent Welfare Caseloads*, unpublished doctoral dissertation, Baltimore: University of Maryland Baltimore County.

Cohen, M. S. (1969). The Direct Effects of Federal Manpower Programs in Reducing Unemployment, *The Journal of Human Resources*, 4 (1), pp. 491-507.

Dahlberg, M., and Forslund, A. (1999). *Direct Displacement Effects of Labour Market Programmes: The Case of Sweden*, Uppsala University Working Paper 1999: 22, Uppsalal.

Dahlberg, M., and Forslund, A. (2005). Direct Displacement Effects of Labour Market Programmes. *Scandinavian Journal of Economics, 107*(3), 475-494.

Davidson, C., and Woodbury, S. A. (1993). The Displacement Effect of Reemployment Bonus Programs. *Journal of Labour Economics*, 11(4), 575-605.

DeKoning, J., Gravesteijn-Ligthelm, T. N., and Verkaik, A (1992). *Werkt de Kra?*, Arbeidsvoorziening, Rijswijk, The Netherlands.

DeMunnik, R. (1992). *Recent Evaluation Studies: Part Two*, Rijswijk, The Netherlands.

Department of Employment, Education and Training (1989). *Jobstart Evaluation*, Canberra, Australia.

Dixon, P. B., and Rimmer, M. T. (2006). The Displacement Effect of Labour-Market Programs: MONASH Analysis. *Economic Record*, 82, S26-S40.

Eriksson, T., and Pehkonen, J. (1995). *Unemployment Flows in Finland: A Time-Series Analysis*, Working Paper No. 156, School of Business and Economics, Juniversity of Jyvaskyla.

Farkas, G., Alton Smith, D., and Stromsdorfer, E. W. (1983). The Youth Entitlement Demonstration: Subsidized Employment with a Schooling Requirement, *Journal of Human Resources*, 18, pp. 557-573.

Fertig, M., Schmidt, C., and Schneider, H. (2002). *Active Labor Market Policy in Germany – Is there a Successful Policy Strategy?*, IZA Discussion Paper 576 Bonn RWI Schriften 74, Berlin, Duncker, Humblot.

Fujiwara, D. (2010). *The Department for Work and Pensions Social Cost-Benefit Analysis Framework: Methodologies for estimating and incorporating the wider social and economic impacts of work in Cost-Benefit Analysis of employment programmes*, DWP Working Paper No. 86.

Garasky, S. (1990). Analyzing the Effect of the Massachusetts ET Program on the States' AFDC-Basic Caseload, *Evaluation Review*, 14, pp. 701-710.

Hagen, T., and Steiner, V. (2001). *Von der Finanzierung der Arbeitslosigkeit zur Forderung von Arbeit*, ZEW Wirtschaftsanalysen Bd. 51, Baden-Baden.

Hamermesh, D. S. (1972). The Secondary Effects of Manpower Programs, *Economic and Business Bulletin*, 24 (3), pp. 18-26.

Hall, R. E. (1979). Comment on `The Labor Market Displacement Effect in the Analysis of the Net Impact of Manpower Training Programs by George Johnson, in *Evaluating Manpower Training Programs: Research in Labor Economics, Supplement 1*, JAI Press, Inc., pp. 255-257.

Johnson, G. E., *Does Public Service Employment Pay Off?*, Unpublished paper, Ann Arbor, MI.: University of Michigan, August 1972.

Johnson, G. (1979). The labor Market Displacement Effect in the Analysis of the Net Impact of Manpower Training Programs, in *Research in Labor Economics*, Supplement 1, pp. 227-254.

Johnson, G. E., and Tomola, J. D. (1977). The fiscal substitution effect of alternative approaches to public service employment policy. *Journal of Human Resources*, 12(1), 3-26.

Johnson, T. R., Klepinger, D. H., and Dong, F. B. (1990). *Preliminary Evidence from the Oregon Welfare Reform Demonstration*, unpublished paper.

Jongen, E., Van Gameren, E., Graafland, J. (2003). Exploring the Macroeconomic Impact of Subsidized Employment, *De Economist* 151(1), pp.81-118.

Kangasharju, A., and Venetoklis, T. (2003). *Do Wage Subsidies Increase Employment in Firms?*, VATT Discussion Paper 304, Helsinki.

Katz, L. F. (1994). Active Labor Market Policies to Expand Employment and Opportunity, in *Reducing Unemployment: Current Issues and Policy Options*, Proceedings from a Symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, pp. 239-290.

Lise, J., Seitz, S., and Smith, J. (2004). *Equilibrium Policy Experiments and the Evaluation of Social Programs*, National Bureau of Economic Research Working Paper 10283, Cambridge MA.

Kluve, J. (2005). *Study on the Effectiveness of ALMPs*, RWI, Essen, Germany.

Knabe, A. (2008). A Note on the Displacement Effects of Marginal Wage Subsidies. *Academia Economic Papers*, 36(1), 1-23.

Kopits, G. F. (1978). Wage Subsidies and Employment – Analysis of the French Experience. *International Monetary Fund Staff Papers*, 25(3), 494-527.

Manski, C. F. and Garrfinkel, I. (1992). Introduction in Charles F. Manski and Irwin Garfinkel, eds., *Evaluating Welfare and Training Programs*, Cambridge MA: Harvard University Press, pp. 1-22.

Marcotte, J. (1990). E*ffect of Washington State FIP on the Caseload Size*, unpublished paper. Washington, D.C.: The Urban Institute.

Moffitt, R. A. (1992). 'valuation Methods for Program Entry Effects, in Manski, C. F. and Garfinkel, I. eds. *Evaluating Welfare and Training Programs*, Cambridge: Harvard University Press, pp. 231-252.

Moffitt, R. A. (1996). The effect of employment and training programs on entry and exit from the welfare caseload. *Journal of Policy Analysis & Management*, 15(1), 32-50.

Nathan, R. P. (1978). *Job Creation through Public Service Employment*, Washington, DC: National Commission on Manpower Policy.

National Planning Association (1974). *An Evaluation of the Economic Impact Project of the Public Employment Program*, Final Report, Washington, DC.

Phillips, E. H. (1993). *The Effect of Mandatory Work and Training Programs on Welfare Entry: The Case of GAIN in California*, unpublished doctoral dissertation, Madison: University of Wisconsin.

O'Neill, J. (1990). *Work and Welfare in Massachusetts: An Evaluation of the ET Program*, Pioneer Paper No. 3, Boston: Pioneer Institute for Public Policy Research.

Pehkonen, J. (1997). Displacement Effects of Active Labour Market Policy: The Youth Labour Market in Finland. *Empirica*, 24, 195-208.

Perloff, J. M., and Wachter, M. L. (1979). The New Jobs Tax Credit: An Evaluation of the 1977-78 Wage Subsidy Program. *American Economic Review*, 69(2), 173-179.

Puhani, P. A. (1999). *Public training and outflows from unemployment: an augmented matching function approach on Polish regional data*: Centre for Economic Policy Research.

Puhani, P. A. (2003). Active labour market policy and employment flows – Evidence from Polish regional data. *International Journal of Manpower*, 24(8), 897-915.

Rantala, J. (1995). *The Effect of Active Labor Market Programs on Employment,* VATT, Research Reports.

Rotger, G. P., and Arendt, J. N. (2010). *The Effects of a Wage Subsidy on Subsidised Firm's Ordinary Employment*, AKF Working Paper, Copenhagen, Danish Institue of Governmental Research (AKF).

Schiller, B. R. (1973). Empirical Studies of Welfare Dependency: A Survey, J*ournal of Human Resources*, 8 (Supplement), pp. 19-32.

Schiller, B. R., and Brasher, C. N. (1993). Effects of Workfare Saturation on AFDC Caseloads, *Contemporary Policy Issues*, 11 (1), pp. 39-49.

Wiseman, M. (1976). Public Employment as Fiscal Policy, *Brookings Papers on Economic Activity*.

Wissoker, D. A., and Watts, H. W. (1994). *The Impact of FIP on AFDC Caseloads*, The Urban Institute, Washington, DC.

Wolf, D. A. (1990). *Caseload Growth in the Evaluation Sites: Is There a FIP Effect?*, unpublished paper, Washington, D.C.: The Urban Institute.

This report has been produced with the primary aim of informing cost benefit estimates within the Cost-Benefit Framework (CBF). The Department for Work and Pensions CBF is a guidance document for the production of cost-benefit information. The research consisted of two components: a literature review and new empirical estimates from DWP administrative data.

The following areas were covered in the literature review:
• general equilibrium effects;
• subgroup impacts and distribution of impacts;
• impacts on duration of benefits and employment, and wages;
• multiple participation in programmes or other interventions.

This report reviews what is known about these topics and discusses when they are likely to be important, with recommended actions in the context of the CBF net impact analyses and cost-benefit analyses. For general equilibrium effects, estimates from the literature are presented and recommendations are made for to account for these effects in cost-benefit analyses. These estimates can be used to guide sensitivity tests. For duration of benefits and employment, the analysis of the DWP administrative data provides empirical estimates of gross duration of benefits and employment, and annual nominal taxable HMRC earnings. These estimates have been produced for a range of New Deal programmes.

If you would like to know more about DWP research, please contact:
Kate Callow, Commercial Support and Knowledge Management Team,
Upper Ground Floor, Steel City House, West Street, Sheffield, S1 2GQ.
http://research.dwp.gov.uk/asd/asd5/rrs-index.asp

DWP Department for Work and Pensions