

Working paper

Lone Parent Obligations and New Services for Lone Parents: a feasibility study of an impact assessment

by Mike Brewer, James Browne and Tom Crossley

Department for Work and Pensions

Working Paper No 92

Lone Parent Obligations and New Services for Lone Parents: a feasibility study of an impact assessment

Mike Brewer, James Browne and Tom Crossley

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

© Crown copyright 2010.

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

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

This document/publication is also available on our website at:
<http://research.dwp.gov.uk/asd/asd5/rrs-index.asp>

Any enquiries regarding this document/publication should be sent to us at:
Department for Work and Pensions, Commercial Support and Knowledge Management Team,
Work and Welfare Central Analysis Division, 3rd Floor, Caxton House, Tothill Street, London SW1H 9NA

First published 2010.

ISBN 978 1 84712 891 1

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

Contents

Acknowledgements.....	vii
The Authors.....	viii
Abbreviations.....	ix
Glossary of terms.....	xi
Summary.....	1
1 Introduction.....	5
2 What is the policy change and what groups are of interest?.....	7
2.1 What is the policy change?.....	7
2.1.1 <i>Withdrawal of eligibility to IS for lone parents with older children</i>	7
2.1.2 <i>Preparing for the transition from IS to JSA/ESA</i>	8
2.1.3 <i>Options and Choices events</i>	9
2.1.4 <i>Final Year Quarterly WFIs</i>	9
2.1.5 <i>New in-work support</i>	9
2.2 When does the lone parent policy change begin for a particular lone parent?.....	9
2.3 Who is the population of interest?.....	10
2.4 Summary.....	11
3 The empirical approach.....	13
3.1 Evaluation techniques and the role of the comparison group.....	13
3.2 Some alternative evaluation strategies.....	16
3.2.1 <i>Simple-difference, regression, and matching</i>	16
3.2.2 <i>Regression discontinuity</i>	17
3.2.3 <i>'The Norway Approach'</i>	18
3.3 The difference-in-difference design.....	19
3.3.1 <i>Extending the DiD design: more than two groups or time periods</i>	21
3.3.2 <i>Threats to the internal validity of the DiD design</i>	22
3.4 Choosing a comparison group.....	26
3.5 Assessing whether common trends holds between lone parents affected by the LPPC and lone parents with younger children.....	27

3.5.1	<i>Assessing the suitability of DiD with lone parents with younger children as a comparison group: analysis using simple averages</i>	28
3.5.2	<i>Correct inference in DiD models</i>	37
3.5.3	<i>Assessing the suitability of DiD with lone parents with younger children as a comparison group: analysis using linear regression</i>	39
3.5.4	<i>Overall assessment of the suitability of the DiD method, using lone parents with younger children as a comparison group</i>	47
3.6	Summary.....	48
4	The empirical approach: further refinements and detailed considerations.....	51
4.1	Dynamic selection and duration models	51
4.2	Estimating the impact of the roll-out of the lone parent policy change as well as the steady-state	53
4.3	Separating the impact of the different elements of the lone parent policy change	55
4.4	Summary.....	56
5	Outcomes of interest and likely sources of data.....	57
5.1	What outcomes could be addressed by an impact assessment?	57
5.2	Estimating the impact of the lone parent policy change on new and repeat claimants.....	65
5.3	Substitution, displacement and other issues relevant to a cost-benefit analysis.....	66
5.4	How precisely can any impacts be estimated?.....	68
5.5	Identifying which lone parents are affected by the lone parent policy change, and when.....	69
5.6	Summary.....	70
6	Potential confounding factors.....	71
6.1	Anticipation effects	71
6.2	Other policies.....	72
6.2.1	<i>Policies which have already taken effect</i>	72
6.2.2	<i>Policies which have yet to take effect</i>	74
6.3	Summary.....	75

7	Conclusions and recommendations: what is and is not feasible in an impact assessment of the lone parent policy change.....	77
7.1	Empirical methods	77
7.2	Choice of datasets	78
7.3	Detailed implementation issues for an impact assessment	78
Appendix A	Constructing outcome measures from the Work and Pensions Longitudinal Study.....	81
Appendix B	Comparison groups for each of the phases of the roll-out.....	85
Appendix C	Likely sample sizes for different phases of the roll-out by date Income Support eligibility is withdrawn	91
Appendix D	A formalisation of the regression discontinuity design.....	93
References	95

List of tables

Table 3.1	A simple difference-in-difference design	21
Table 3.2	Outcome: off benefit three years after initially observed on IS, Phase 1	42
Table 3.3	Outcome: off benefit three years after initially observed on IS, Phase 2	42
Table 3.4	Outcome: off benefit three years after initially observed on IS, Phase 3	43
Table 3.5	Outcome: proportion of three years spent off benefit, Phase 1	43
Table 3.6	Outcome: proportion of three years spent off benefit, Phase 2	44
Table 3.7	Outcome: proportion of three years spent off benefit, Phase 3	44
Table 3.8	Outcome: in work three years after initially observed on IS, Phase 1.....	45
Table 3.9	Outcome: in work three years after initially observed on IS, Phase 2.....	45
Table 3.10	Outcome: in work three years after initially observed on IS, Phase 3.....	46
Table 3.11	Outcome: proportion of three years spent in work, Phase 1.....	46
Table 3.12	Outcome: proportion of three years spent in work, Phase 2.....	47
Table 3.13	Outcome: proportion of three years spent in work, Phase 3.....	47
Table 5.1	Summary of datasets	59
Table 5.2	Outcomes	61
Table 5.3	Feasibility of estimating various outcomes needed for a cost-benefit analysis	67
Table B.1	Comparison groups for each of the phases of the roll-out.....	86
Table C.1	Likely samples sizes for different phases of the roll-out, by date Income Support eligibility is withdrawn	92

List of figures

Figure 3.1	Comparison of the treatment and comparison groups after treatment is introduced	20
Figure 3.2	Results of failure of the common trend assumption.....	23
Figure 3.3	Failure of the common trend assumption if programme eligibility depends on mean-reverting lagged outcomes	24
Figure 3.4	Result if the outcome is limited and the comparison group is near it.....	25
Figure 3.5	Comparing trends of benefit outcomes for lone parents with differently aged youngest children (12 to 15 compared to three to four)	29
Figure 3.6	Comparing trends of benefit outcomes for lone parents with differently aged youngest children (10 to 11 compared to three to four)	30
Figure 3.7	Comparing trends of benefit outcomes for lone parents with differently aged youngest children (seven to nine compared to three to four)	30
Figure 3.8	Comparing differences in benefit outcomes for lone parents with differently aged youngest children (12 to 15 compared to three to four)	31
Figure 3.9	Comparing differences in benefit outcomes for lone parents with differently aged youngest children (10 to 11 compared to three to four)	31
Figure 3.10	Comparing differences in benefit outcomes for lone parents with differently aged youngest children (seven to nine compared to three to four)	32
Figure 3.11	Comparing trends of work outcomes for lone parents with differently aged youngest children (12 to 15 compared to three to four)	34
Figure 3.12	Comparing trends of work outcomes for lone parents with differently aged youngest children (10 to 11 compared to three to four)	34
Figure 3.13	Comparing trends of work outcomes for lone parents with differently aged youngest children (seven to nine compared to three to four)	35
Figure 3.14	Comparing differences in benefit outcomes for lone parents with differently aged youngest children (12 to 15 compared to three to four)	35
Figure 3.15	Comparing differences in benefit outcomes for lone parents with differently aged youngest children (10 to 11 compared to three to four)	36
Figure 3.16	Comparing differences in benefit outcomes for lone parents with differently aged youngest children (seven to nine compared to three to four)	36
Figure D.1	An example of the regression discontinuity design	94

Acknowledgements

This feasibility study was commissioned by the Department for Work and Pensions (DWP) and the project managers have been Kris Chapman, Alison Herrington, Jane Hall and Karen Elsmore. The authors are grateful to these and other DWP officials who have helped in their understanding of the policy environment and the Work and Pensions Longitudinal Study (WPLS). They also thank Monica Costa Dias and Jonathan Shaw for helpful discussions, and have benefited from feedback from an Advisory Group convened by the DWP.

This research was done in two parts: Stage 1 considered the feasibility of undertaking an impact assessment of the policy change known as ‘Lone Parent Obligations and New Services for Lone Parents’ (together called ‘lone parent policy change’, or LPPC). Stage 2, which involved secondary analysis of DWP administrative data, provided evidence on the suitability of the difference-in-difference methodology, and the use of lone parents with younger children as a comparison group, both of which were recommended by the research team after concluding Stage 1. This report, however, presents the findings as a unified whole. The study was undertaken between 2009 and 2010, during which time there was a change of Government. Consequently, not all policies and initiatives referred to are necessarily the policy of the current Government.

The Authors

Mike Brewer is director of the Direct Tax and Welfare research sector at the Institute for Fiscal Studies. His main interests are in welfare reforms, and the way the tax and benefit system affects families with children. Within this broad area, he has analysed the generosity of the tax and benefit system for families with children, and the impact of the same on work incentives. He took part in a major evaluation of Working Families' Tax Credit, examining its impact on labour supply and take-up, and more recently has evaluated the labour market impact of In-Work Credit. Other research has looked at fertility and public policy, how women's labour market behaviour changes after they have children, and trends in childcare use. He has an interest in measuring poverty and living standards, and has closely followed the current Government's ambition to eradicate child poverty.

James Browne is a Senior Research Economist in the Direct Tax and Welfare sector at the Institute for Fiscal Studies. His general research interest is in the economic impacts of the personal tax and benefit system. In particular, he has recently examined the policies that the Government could introduce to meet its child poverty targets, how tax and benefit reforms over the last 30 years have affected the distribution of income and incentives to work, and the effect of In-Work Credit on the employment and benefit outcomes of lone parents.

Dr. Thomas F. Crossley is a Research Fellow at the Institute for Fiscal Studies and a Reader in Economics at the University of Cambridge. He is also director of the consumption and saving research sector at the Institute for Fiscal Studies.

Abbreviations

BHPS	British Household Panel Survey
CTC	Child Tax Credit
DID	Difference-in-difference
DLA	Disability Living Allowance
DWP	Department for Work and Pensions
ESA	Employment and Support Allowance
EZ	Employment zone
FND	Flexible New Deal
FRS	Family Resources Survey
FYQWFI	Final Year Quarterly Work Focused Interviews
GHS	General Household Survey
HMRC	Her Majesty's Revenue & Customs
IA	Impact assessment
IB	Incapacity Benefit
IFS	Institute for Fiscal Studies
IS	Income Support
ITT	Intention to treat
IWC	In-Work Credit
IWRP	In Work Retention Pilot
JSA	Jobseeker's Allowance
LEPs	Local Employment Partnerships
LFS	Labour Force Survey
LPO	Lone Parent Obligations
LPPC	Lone parent policy change
NBD	National Benefits Database
NDLP	New Deal for Lone Parents
RD	Regression discontinuity
WPLS	Work and Pensions Longitudinal Study

Glossary of terms

Anticipation effect	A response to a policy change by an individual that occurs before they are actually affected by it.
Attenuation bias	A bias that reduces the magnitude of estimated policy effects, usually caused by measurement error.
Common trends	A necessary assumption for the difference-in-difference (DiD) estimator to be valid – requires that outcomes in the treatment and comparison groups would change by the same amount in the absence of any policy change.
Comparison group	In DID analysis, a set of individuals who are not directly affected by the policy change under consideration, and whose outcomes are used to estimate the change in outcomes that would have occurred for the treatment group had the policy change not taken place.
Dynamic selection bias	Bias that arises when examining the effects of a policy on an outcome that is only measured for those who make a transition and the likelihood of making the transition is affected by the policy.
External validity	External validity refers to the extent to which that estimate can be generalised to other populations or time periods.
General equilibrium effects	Effects of a policy on those who are not directly affected, for example through changes to wages and prices in the whole economy.
Intention to treat	A variable which describes the average effect of a policy on all those who are meant to be affected by it rather than all those who actually are, acknowledging that some individuals will choose not to take advantage of the policy or will be overlooked (because of administrative mistakes).
Internal validity	Internal validity refers to the ability of a research design to deliver a good estimate of the effect in question.
Longitudinal data	Panel data, where individuals or households are followed over a period of time.
Roll-out	The period between November 2008 and January 2011 during which the oldest age of youngest child consistent with being eligible to Income Support (IS) as a lone parent will fall from 16 to seven.
P45/P46	Forms sent to Her Majesty's Revenue & Customs (HMRC) when an individual leaves or starts a job.

Steady-state	The period after January 2011, when IS eligibility for lone parents will have reduced to those with a youngest child aged seven. The Government has announced this age will reduce to five from 2012.
Treated	Used to describe individuals who are directly affected by the policy intervention being considered.
Treatment group	Set of individuals directly affected by the policy change under consideration.
Untreated	Used to describe individuals who are not directly affected by the policy intervention being considered.

Summary

This report investigates the feasibility of undertaking an impact assessment of the policy change known as ‘Lone Parent Obligations and New Services for Lone Parents’ (together called ‘lone parent policy change’, or LPPC).

Any impact assessment of the LPPC will have to overcome many difficulties. Some of these are due to its design: the policy was rolled out quickly in a uniform way across the country. Some of these reflect its timing: it was rolled out during a recession, and there has not been a comparable recession for over 15 years. Some of these reflect the wider policy environment: other policy changes have happened or are planned which have affected or will affect lone parents receiving Income Support (IS). However, the researchers’ assessment is that an impact assessment for some groups of affected clients and for the main benefit and employment outcomes is, in principle, feasible using the difference-in-differences (DiD) approach, although any DiD estimator is always subject to a number of untestable assumptions. It also seems unlikely that the DiD approach could reliably detect small impacts (less than two percentage points).

Key implications of the design of the lone parent policy change for an impact assessment (Chapter 2)

The date at which a lone parent will cease to be eligible to IS will depend on whether they are already claiming, the date of birth of their youngest child, and the date of their Work Focused Interviews (WFIs), but in a predictable manner with virtually no personal adviser discretion. This predictability is of great help when designing an impact assessment.

Lone parents will be offered additional support to move closer to the labour market during the 12 months before they lose entitlement to IS as a lone parent. Those entering work will also receive additional support. These changes are part of the LPPC. For this reason, and to allow for anticipation effects in response to the impending loss of entitlement to IS, the date on which the policy first affects a lone parent should be considered to be at least 12 months before the loss of their entitlement to IS. A variant in which the policy is assumed to affect lone parents 24 months before they lose IS entitlement should be considered; this will be particularly important in the steady-state, when lone parents will have known about the LPPC for a long time. The drawback is that this increases the differences in ages of the children in the treatment and comparison groups.

Empirical methods and comparison groups (Chapters 3 and 4)

The recommended empirical strategy is to use a DiD approach. This will compare the outcomes of lone parents directly affected by the LPPC with those of lone parents whose children are sufficiently young so that they are not affected; this comparison will then be itself compared to the difference in outcomes of similar groups of lone parents observed before the LPPCs. The idea is that differences in outcomes of lone parents with differently-aged youngest children before the LPPC begins are informative about what the differences in outcomes of lone parents with differently-aged youngest children would have been in the absence of the LPPC.

The key assumption needed for a DiD estimate to be valid is known as the ‘common trends’ assumption. In the case of the LPPC, common trends would fail if the labour market behaviour of lone parents with younger children changed in a way which was different from the change in the labour market behaviour of lone parents with older children. There is some evidence that the

‘common trends’ assumption did not hold during the period from 2001 to 2007. But the divergences are usually small, and may not be statistically significant, and they could be accounted for by using a trend-adjusted DiD model. However, even if this assumption holds in the pre-programme period, there is no guarantee that it will still hold after the LPPC is introduced. A particular concern for this impact assessment (IA) is that the LPPC is being implemented during a recession, and there is no pre-programme data to test whether common trends held during previous recessions. Over and above this, though, the main threats to the DiD approach are anticipation effects, and the impact of other contemporaneous policy changes.

Some outcomes, such as the duration on benefit and the number of lone parents starting work, could be examined using either a conventional linear DiD regression or a duration model. However, a DiD estimator cannot be used to estimate the effect of the LPPC on work-contingent outcomes such as job duration or earnings (were suitable data available) due to the dynamic selection bias problem discussed in Section 3.4. A duration model is required to examine these outcomes.

Naïve analysis using the DiD method often produces estimated standard errors which are too small, and the conventional ways of addressing these (with clustered standard errors) may not be appropriate in the particular case of the LPPC. An eventual impact assessment should examine the ways in which standard errors can be estimated correctly in this situation, but it should be noted that these methods are at the cutting-edge of modern econometrics and the literature has not yet come to a consensus. The overall assessment of the research team is that, if the actual impact of LPPC is expected to be smaller than two percentage points, then it seems unlikely that such impacts could be detected reliably. But if an eventual IA is intended to test whether the LPPC had impacts as large as five to ten percentage points (compared with no effect), then a DiD or a trend-adjusted DiD model should provide robust answers.

There is a three-way trade-off between examining outcomes for a longer period of time, allowing for longer anticipation effects, and reducing the difference between the treatment and comparison groups in terms of the age of the youngest child. For some of the early phases of the roll-out, the difference in ages of the treatment and comparison groups will be large. A small difference in ages is desirable to make the common trends assumption more plausible, but a larger difference is needed in order to ensure that the comparison group is completely unaffected by the programme.

There is a 27 month ‘roll-out’ period between November 2008 and January 2011 before the policy reaches a steady-state¹. There are different problems to contend with when examining the impact of the roll-out and the steady-state policy. On balance, the internal validity of estimates based on the roll-out is probably higher than those based on the steady-state, but estimates of the roll-out of the LPPC may be different from (and therefore, a poor guide to) estimates based on the steady-state. Whether these represent significant disadvantages depends on whether there is interest in understanding the impact of the roll-out in itself, rather than merely using the estimates from the roll-out as a guide to the likely impact of the steady-state policy. For example, the roll-out provides the only opportunity to estimate the impact of the LPPCs on lone parents whose youngest child is aged ten or more.

No quantitative method could robustly separate the overall impact of the LPPC into that due to ‘anticipation effects’, ‘support’ effects and ‘after effects’, or separate the different elements of the support package, even were additional data available in the Work and Pensions Longitudinal Study (WPLS). Any IA of the LPPC would give an estimate of the effect of the whole of the LPPC.

¹ Since undertaking the feasibility study, the Government has announced that lone parents with a youngest child aged five and over will lose eligibility to IS from 2012.

Data and outcomes (Chapter 5)

Two groups of lone parents will be directly affected by the LPPC: those receiving IS, who will lose entitlement to it sooner than they would have done otherwise, and those who will no longer be entitled to claim IS. Furthermore, all may be affected by general equilibrium effects, such as substitution, displacement or other spillover effects.

For the first group, administrative data (the WPLS) can be used to measure outcomes relating to benefits and, if tax credit data in the WPLS is of sufficient quality, work outcomes. If information on earnings and tax credit payments were made available within the WPLS, then administrative data could be used to partially estimate the effect of the LPPC on family income. However, this would not allow researchers to estimate the impact of the LPPC on child poverty, as currently defined by the Government.

It may be possible to use the WPLS to estimate the impact of the LPPC – including so-called deterrent effects – on repeat claimants, but cannot, through its nature, be used to estimate the impact of deterrent effects on new customers.

The impact of the LPPC on benefit, work, income and poverty outcomes could also be measured using household surveys, such as the Labour Force Survey (LFS) and Family Resources Survey (FRS), if administrative data on these was of low quality or unavailable. However, small sample sizes would mean that impacts would be estimated imprecisely, and it would be unlikely that one could be confident that the LPPC had had any impact at all.

It will not be possible for an IA to separate out the effects of the different elements of the LPPC, or examine displacement or substitution effects on the population as a whole. This is because of methodological barriers rather than inadequacies of the available data.

Potential confounding factors (Chapter 6)

There are several potential confounding factors that could affect various parts of an eventual IA. Most of these can be dealt with within a DiD or duration model, although there are drawbacks, and it is possible that future developments in policy affecting lone parents could derail the proposed strategy totally.

Anticipation effects can be dealt with by changing the effective start date of the programme in the model. The drawback is that this increases the differences in ages of the children in the treatment and comparison groups, making the ‘common trends’ assumption underlying the DiD approach less plausible.

Other policy changes affecting lone parents on IS and Jobseeker’s Allowance (JSA) have been introduced since 2008. Most of these can be controlled for relatively easily by introducing additional explanatory variables which identify those lone parents who are affected. By accounting for many of these policy changes, the eventual IA of the LPPC will also provide estimates (based on DiD) of the impact of these policies on lone parents on IS. However, some programmes that affect all lone parents cannot be controlled for in this way: any IA would have to assume that lone parents are affected by the same amount by these reforms (which would mean that any impact of these reforms are incorporated in the ‘common’ trend affecting all lone parents).

Some likely future policy changes – such as the introduction of the Work Programme – will have more significant implications for an eventual IA: the Work Programme will, in principle, change the nature of the policy regime for lone parents on JSA, and an extension of the LPPC to lone parents with younger children will directly affect the proposed comparison group. The most robust option would be to stop examining data when the Work Programme starts or LPPC is extended to more lone parents, but this would curtail the ability of an eventual IA to examine the medium- to long-run impact of LPPC on lone parents with children aged over seven. An alternative approach that deals with the Work Programme would be to estimate separate impacts of LPPC before and after the Work Programme. Alternative approaches that deal with the extension of LPPC would be to use lone parents with children aged under three as the comparison group, or to flip the definition of the treatment group and comparison group, using the lone parents whose youngest child is aged seven or over to estimate the impact of the LPPC on those whose youngest child is aged five to seven.

1 Introduction

From November 2008, additional obligations have been placed on lone parents claiming Income Support (IS), and new services have been offered. This report investigates the feasibility of estimating the net impact of these changes, hereafter referred to as ‘an impact assessment’ (IA).

An IA would form part of the the Department for Work and Pensions’ (DWP’s) overall evaluation programme on the effects of Lone Parent Obligations (LPO). This consists of a mixed methods approach including in-depth interviews with customers and staff, a large-scale survey of customers, and analysis of in-house and other data sources to assess the impact of the changes.

The overall aim of the evaluation is:

‘To explore the extent to which lone parent employment interventions provide an effective incentive to look for paid employment, alongside an effective package of support for workless lone parents to enable them to find, enter and sustain paid employment.’

The starting point for this study was that an impact assessment should cover as wide a range of outcomes as possible, and should assess short- and medium-run outcomes (where medium-run is defined as up to four years after a lone parent’s IS has ended), and this study shows to what extent such a wide-ranging impact assessment is or is not feasible. But just because something is feasible does not mean that it should be carried out: all research has costs, as well as benefits, and it takes time before one can estimate the impact of a policy on medium-run outcomes.

This report is arranged as follows: **Chapter 2** outlines the policy change, and discusses the populations of interest. **Chapter 3** discusses what methods could be used to estimate the impact of the lone parent policy change (LPPC). It argues why a difference-in-difference (DiD) approach, using lone parents with younger children as a comparison group, is an attractive method to evaluate the impact of the LPPC, and it provides evidence (from so-called pre-programme or placebo tests) on the suitability of such an approach; inevitably, some of the material in this chapter is technical in nature. **Chapter 4** discusses some detailed refinements to the suggested empirical approach. **Chapter 5** discusses the datasets that are likely to be of use, and what outcomes could be investigated in an eventual IA. **Chapter 6** discusses a number of confounding or complicating factors. **Chapter 7** concludes.

2 What is the policy change and what groups are of interest?

This chapter provides details of the policy changes that are part of the lone parent policy change (LPPC) (Sections 2.1 and 2.2), and discusses what groups in the population will be of interest to the eventual impact assessment (IA) (Section 2.3).

2.1 What is the policy change?

The key policy change, considered by this feasibility study, is that lone parents with a youngest child aged seven or over will no longer be entitled to claim Income Support (IS) solely on the grounds of being a lone parent.² A small minority of lone parents will be able to claim IS on other grounds, but the intention is that those lone parents who are able to work may claim Jobseeker's Allowance (JSA), and those with a health problem or disability may be able to claim Employment and Support Allowance (ESA). In addition to this, changes were made to the pre-employment support provided in the 12 months leading up to the date when entitlement to IS as a lone parent is removed, and increased post-employment support was introduced. These changes together are referred to as the LPPC.

Other policy changes also came into effect between 2008 and 2010 which affected lone parents receiving benefits. The main ones are:

- the introduction of ESA;
- changes to the JSA regime with the introduction of the Flexible New Deal (FND).

There are also likely to be policy changes in the future, such as the introduction of the Work Programme. These other policies, and their effect on a possible IA, are discussed in Section 6.2. The rest of this chapter gives more detail about the policies which form part of the LPPC, and therefore, fall within the scope of the impact assessment:

2.1.1 Withdrawal of eligibility to IS for lone parents with older children

Before November 2008, lone parents could claim IS until their youngest child reached 16. That age cut-off is gradually falling over a three-year period. In general terms, eligibility to claim IS on the grounds of being a lone parent was, or will be, withdrawn at the following times:

- for new and repeat claims by lone parents with a youngest child aged:
 - 12 or over, from November 2008;
 - ten or over, from October 2009;
 - seven or over, from October 2010;

² The Government announced in June 2010 plans to extend Lone Parent Obligations (LPO) to lone parents whose youngest child is aged five and six. Such a change would require primary legislation, and so the implications for this for an eventual IA are discussed in Chapter 5.

8 What is the policy change and what groups are of interest?

- for lone parents already receiving IS (referred to by DWP as ‘existing customers’) with a youngest children aged:
 - 14 or over, March to May 2009;
 - 12 or over, July to December 2009;
 - 11 or over, February to April 2010;
 - ten or over, June to October 2010;
 - nine or over, November 2010 to October 2011;
 - seven or over, from January 2011.

This report refers to the period between November 2008 and January 2011 as the ‘roll-out period’, and after January 2011 as the ‘steady-state’. The possible extension of LPO to lone parents whose children are aged five and six is discussed in Chapter 6.

The date at which lone parents already receiving IS cease to be eligible may also depend on the date of birth of their youngest child (different cohorts were or will be treated differently in some cases) and the date of their Work Focused Interview (WFI). Appendix C outlines in more detail when lone parents already receiving IS cease or ceased to be eligible, and Section 5.5 discusses how accurately one can replicate the date on which a lone parent lost their IS entitlement using the administrative data in the Work and Pensions Longitudinal Study (WPLS).

Some lone parents are not be affected by these changes, even if their youngest child is older than indicated in the list above. Lone parents who have other reasons for claiming IS (than being a lone parent) – for example, those who have children for whom the middle or highest rate of care component of Disability Living Allowance (DLA) is payable, or those who claim Carer’s Allowance, or who are fostering children – continue to be eligible for IS.³ The implication of these groups for the eventual IA is discussed in Section 5.5.

Having lost entitlement to claim IS as a lone parent, lone parents who are available to work will be entitled to claim JSA. Lone parents with a health problem or disability may be able to claim ESA. ESA is generally paid at the same rate as JSA during the assessment period. Afterwards, those placed in the Work Related Activity Group receive a work-related activity component in addition and those in the Support Group receive a support component in addition to the basic rate.

2.1.2 Preparing for the transition from IS to JSA/ESA

To prepare lone parents for the loss of IS eligibility a number of initiatives are in place to give lone parents information about the changes and help lone parents move closer to the labour market. These are:

³ The main groups of lone parents who will continue to be eligible for IS after the LPPC are foster carers, lone parents who are in receipt of middle or higher rate care component DLA for a dependent child, lone parents who are in receipt of Carer’s Allowance, and lone parents who are full-time students at the point the IS entitlement changes come into force, who are transitionally protected for the duration of the course. Furthermore, those few lone parents who had been claiming IS continuously since before April 2004 and who were receiving child additions to IS rather than the Child Tax Credit (CTC) could have their eligibility extended for one period of four weeks if the CTC claim for the lone parent moving onto JSA had not been processed at the time of the voluntary interview six weeks before the original IS end date.

2.1.3 Options and Choices events

These are group sessions designed to inform lone parents about the support available to them and the forthcoming changes to their benefit entitlement.

2.1.4 Final Year Quarterly WFIs

Quarterly WFIs (QWFIs) now take place in the year before a lone parent loses eligibility to IS.

Lone parents eligible for a QWFI are, or will be, those with a youngest child aged:

- 11 from 8 September 2008;
- ten to 11 from 1 February 2009;
- nine from 6 June 2009;
- eight from 25 October 2009; and
- six to seven from 3 January 2010.

An exception to this was the set of lone parents whose youngest child was 12 on or before 5 July 2009.

2.1.5 New in-work support

As part of the LPPC, additional support and guidance has been made available to lone parents moving into work. These are:

- national extension of In-Work Credit (IWC) (from April 2008);
- in one Jobcentre Plus district, there was a pilot scheme known as the In Work Retention Pilot that varied the frequency of IWC payments so that there are two lump sum payments towards the end of the 52 weeks. This took place between 1 July 2008 and 30 June 2010;
- In-Work Advisory Support from Jobcentre Plus Advisers;
- the national roll-out of the In Work Emergency Discretionary Fund;
- piloting the provision of support for Up-Front Childcare Costs in London between 28 April 2008 and 31 March 2010;
- Better-off in Work Credit, a payment that ensured that benefit claimants were at least £25 a week better off in full-time work than on an out-of-work benefit was piloted in one Jobcentre Plus region between October 2008 and October 2009 (although this should not be relevant for most lone parents on IS given the in-work support already available).

Some of these policies are not limited to lone parents affected by LPPC: they are, or were, also available to lone parents with younger children. As discussed in Section 6.2, the eventual IA will, therefore, need to account explicitly for these policies.

2.2 When does the lone parent policy change begin for a particular lone parent?

It should be stressed that, although the most important single date for a lone parent affected by the LPPC is the date on which they lose entitlement to IS as a lone parent, the eventual IA should not limit itself to examining how outcomes change only after that date. It is likely that the LPPC has an impact on lone parents' behaviour at least a year before they lose IS entitlement, for two reasons:

10 What is the policy change and what groups are of interest?

- First, the 12 months leading up to the loss of entitlement to IS as a lone parent is regarded as part of the LPPC, because that period sees lone parents facing a more intensive WFI regime, and most of the additional pre-employment support is concentrated in the final year of entitlement.
- Second, lone parents receiving IS may alter their behaviour in anticipation of losing their entitlement to IS in the near future. These changes are commonly referred to as ‘anticipation effects’, because the lone parent would be altering their behaviour because of something that will happen in the future. For example, a lone parent who considers that she will not be able to claim JSA indefinitely might decide to start work before she loses entitlement to IS if there is a risk that she might not be able to find such a good job offer later on when claiming JSA.

If there are anticipation effects, then considering only those lone parents who remained on IS until they lost entitlement would ignore some of the impacts of the LPPC. Furthermore, any impact assessment that attempted to estimate the effect of the LPPC by comparing outcomes before and after the reform would be biased if lone parents who leave IS in anticipation of the loss of IS entitlement were not representative of all lone parents affected by the LPPC. This is because in this case, differences between lone parents observed before and after the reform would reflect these compositional differences as well as the impact of the LPPC and wider economic trends.

As discussed in Section 6.1, in general, anticipation effects can be dealt with in an IA by altering the date on which it is considered the policy change first applies. Given the nature of the LPPC, it would be natural to take this date as being 12 months before the loss of entitlement to IS.

2.3 Who is the population of interest?

This section discusses the various groups which might be affected by the LPPC or which might be considered by the eventual IA.

The main group of lone parents who might be directly affected by the LPPC are those already on IS, who will have their eligibility to receive IS as a lone parent removed earlier than it would otherwise have been (DWP refer to this group as ‘existing customers’).

It is expected that the LPPC encourages some lone parents already claiming IS to leave benefit (whether IS or JSA) earlier than they would otherwise. This could be for a number of reasons:

- lone parents receiving IS or JSA may look for work sooner than they otherwise would have in anticipation of, or in response to, the JSA regime, with its requirement to participate in job search and accept certain job offers;
- these same factors could also induce lone parents receiving IS or JSA to leave benefit but not start work, perhaps because they have re-partnered;
- lone parents may find work more easily as a result of the new support being introduced in the run-up to the withdrawal of eligibility to IS.

As discussed in Section 2.1, the LPPC does not begin to affect lone parents only when they lose their entitlement to IS. The 12 months leading up to the loss of entitlement to IS is part of the LPPC, as it will feature a more intense WFI regime, and additional pre-employment support, and there are likely to be anticipation effects to the imminent loss of entitlement to IS as a lone parent.

The LPPC may also improve subsequent work retention among lone parents in this group who do leave work for a job, for a number of reasons:

- as a result of the additional in-work support provided during the first 26 weeks (or 52 weeks in the case of IWC) of work;
- because some lone parents are unable to claim JSA if they were to stop work (but would, in the absence of the LPPC, have been able to claim IS under the same circumstances).

There are also some potential impacts which are not considered further in this study.

First, as a result of the LPPC, some lone parents, depending on the date of birth (or age) of their youngest child, will be unable to claim IS. This will probably lead to them being less likely to be receiving any benefits at all – because not all of those who would have claimed IS in the absence of LPPC will be able to claim either JSA or ESA after the LPPC – and therefore, more likely to be in work or to re-partner. In fact, these impacts need not be confined to lone parents: some adults in couples with children, or some single women without children, for example, might alter their behaviour in the knowledge that they would not be able to claim IS as a lone parent. However, the implications of this for an eventual IA are not considered further in this report.

Second, theoretically, the LPPC may encourage some lone parents currently not receiving an out-of-work benefit to claim JSA in order to take advantage of the new opportunities for training available to claimants, or to take advantage of IWC, but the research team consider this extremely unlikely in practice.

Third, all lone parents, and others in the population, might be affected by the LPPC through spillovers, displacement and other general equilibrium effects on wages and prices (discussed further in Section 4.3).

2.4 Summary

This chapter has outlined the key features of the policies that make up the LPPC. The most important points for the eventual IA are that:

- two groups of lone parents are directly affected by the LPPC: those receiving IS who lose entitlement to it sooner than they would have done otherwise, and those who are no longer entitled to claim IS. Furthermore, all may be affected by general equilibrium effects, such as substitution, displacement or other spillover effects;
- the date at which a lone parent will cease to be eligible to IS depends on whether they are already claiming IS, the date of birth of their youngest child, and the date of their WFIs, but it does so in a predictable manner with virtually no personal adviser discretion. This predictability is of great help when designing an impact assessment;
- lone parents are being offered additional support to move closer to the labour market during the 12 months before they lose entitlement to IS as a lone parent, and those entering work are also receiving additional support. These changes are part of the LPPC and so the date on which the policy first affects a lone parent should be considered to be 12 months before the loss of their entitlement to IS; this will also allow for anticipation effects in response to the impending loss of entitlement to IS;
- other policy changes affecting lone parents have occurred since 2008, or will occur in the near future, and these will need to be accounted for in an eventual IA (discussed in Section 6.2);
- there is a long ‘roll-out’ period during which the age of youngest child cut-off is falling before the policy reaches a steady-state.

3 The empirical approach

This chapter reviews the econometric methods that could be used in an impact assessment (IA). It is necessarily more technical than the other chapters.

Section 3.1 sets out a framework for thinking about how to choose an evaluation technique, and a comparison group. Section 3.2 outlines methods that were considered but rejected, Section 3.3 argues why a difference-in-difference (DiD) approach, using lone parents with younger children as a comparison group, is an attractive method to evaluate the impact of the lone parent policy change (LPPC), and Section 3.4 outlines in detail which cohorts of lone parents could act as comparison groups. With these recommendations in mind, Section 3.5 reports the results from tests designed to assess the suitability of DiDs with lone parents with younger children as a comparison group. This was done by testing whether the ‘common trends’ assumption holds in the period before the LPPC was introduced (this is sometimes known as a placebo test, or pre-programme test).

The next chapter discusses some elaborations to the basic DiD approach, including what more can be gained by using a duration model, the advantages and disadvantages of evaluating the impact of the roll-out compared with the steady-state policy and why the IA will not be able to estimate the separate impact of the different elements of the LPPC.

3.1 Evaluation techniques and the role of the comparison group

The goal of an IA is to determine how the outcomes of treated or affected individuals were altered by the policy or reform in question. That is, an IA tries to compare the treated and untreated outcomes of a particular group of individuals, in a particular time period.

All empirical IA methods proceed by (implicitly or explicitly) contrasting the outcomes of the treated group to the outcomes of a suitably chosen comparison (or ‘control’ group). The key problem in an IA, which the comparison group is intended to solve, is that the counterfactual outcome of the treated group is not observed. For example, in evaluating the effect of LPPC, the outcomes of a group of lone parents who lose their entitlement to Income Support (IS) are observed (and receive a package of new services). What is not observed (and can never be observed) is the outcomes of exactly that group of lone parents in exactly that time period, had they not lost their entitlement to IS (and received a package of new services).

The role of the comparison group in an IA, therefore, is to serve as a basis for estimating the unobserved, counterfactual outcomes of the treatment group. This is true even in research designs in which it is not entirely obvious. For example, in a simple-difference (or time-series) design, the outcomes of the treated group are compared to the outcomes of the same individuals in an earlier (pre-treatment) period. Here, the same individuals effectively form the treatment and comparison group, and past outcomes of these individuals are used to impute what their outcomes in the period of interest would have been in the absence of the treatment. In a parametric duration model, in which treatment status (exposure to the policy or reform) appears as an explanatory variable, a comparison is being made between individuals exposed to the treatment and individuals not exposed to the treatment (or perhaps between the same individuals in periods before and after exposure). This comparison may be ‘adjusted’ by including other variables in the model, but, without variation in treatment status, the model could not be estimated.

In a randomised trial, individuals are randomly assigned to treatment and comparison groups, and so, if the sample is large and there is full compliance with the experimental design, the treatment and comparison group will be, on average, the same in all dimensions, including in untreated outcomes. The outcomes of the untreated comparison group are, therefore, a sound basis for estimating the unobserved untreated outcomes of the treatment group.

In settings such as the LPPC, where there has not been randomisation, the problem is that the comparison group may differ from the treatment group in important ways. As a result, the outcomes of the comparison group may be a poor guide to what the outcomes of the treatment group would have been in the relevant period, had they not been treated. For example, in the simple-difference design, if there are important trends in outcomes (due to changing macro-economic conditions, for example), the past outcomes of the treatment group may not be a good guide to what their outcomes would have been, in the absence of treatment, in the period of interest. In the particular case of the LPPC, the lone parents exposed to the treatment were different in the past in the critical respect that their children were younger. It is well recognised that the age-of-youngest child is an important determinant of labour supply among lone parents, and this means that past outcomes of the treated group do not offer a compelling counterfactual. Thus, the simple-difference or time-series research design is unattractive.

To choose among alternative evaluation strategies is to choose a comparison group that offers a compelling and convincing contrast, and possibly, also to choose a method of adjusting the outcomes of that comparison group to make the contrast more convincing. As the LPPC is being rolled out nationally, the choice of comparison groups is limited. The view of the research team is that any research design that involves comparisons with comparison groups comprised of individuals other than lone parents (for example, single women, or members of couples) is not credible: the labour supply behavior of these groups is just too different for such a research design to be compelling.⁴ Thus, attention was focused on research designs involving contrasts with lone parents who are not exposed to the LPPC. There are two groups:

- lone parents with a youngest child of a different age (including perhaps data on the same individuals at an earlier date);
- lone parents who are exempted from the change.

As noted above, a review of the particulars of the LPPC, and of the programme evaluation literature has led us to conclude that:

- the most compelling contrast available is between affected lone parents and lone parents that are not affected because of the age of their youngest child;
- it is necessary to use data from other (pre-LPPC) time periods to correct for those differences in the behavior of lone parents with children of different ages that are not related to the LPPC.

These elements suggest a DiD research design.

⁴ Some papers have used mothers in couples as comparison groups for lone mothers with a DiD approach (Gregg and Harkness, 2003). That may be appropriate when the population of interest is all mothers (as it was in Gregg and Harkness). However, in this application, the population of interest is parents receiving an out-of-work benefit, and there are considerable differences between lone mothers receiving an out-of-work benefit, and mothers in couples with children receiving an out-of-work benefit.

Such contrasts can be implemented through a standard DiD estimator (which can be implemented via linear regression, as will be explained below). It is also possible to set up the conditioning variables in a duration model (particularly variables capturing time and group effects) in such a way as to implement the contrasts described above. A duration model will typically involve additional assumptions relative to a standard DiD estimator, but enables more outcomes to be examined, for example the effect of the LPPC on the subsequent job duration, subsequent claims of out-of-work benefits, and, if the data permits, on earnings.

The next two sections briefly review some research designs which were deemed by the research team to be inappropriate for this particular impact evaluation, and the reasons for those conclusions (Section 3.2) and then discuss the difference-in-difference design, the possible threats to the validity of that design, and the reasons why it should work well in this context (Section 3.3).

Before turning to a review of possible research designs, Box 3.1 offers a formal statement of the evaluation problem: this is based on the *Potential Outcomes* framework which is often ascribed to Rubin (1974) (see, for example, Holland, 1986) but which dates back to the much earlier work of statisticians such as Fisher and Neyman.

Box 3.1 – A formal statement of the evaluation problem

The evaluation problem can be formalised as follows. Let Y_1 and Y_0 be random variables capturing the outcome for an individual if they do, and do not, receive treatment (in this case, the LPPC) respectively. The realisations of these random variables for individual i are given by Y_{1i} and Y_{0i} . The impact of treatment for this individual is given by:

$$\Delta_i = Y_{1i} - Y_{0i} \tag{1}$$

This framework makes a stable-unit-treatment-value assumption: the impact of an intervention may vary across individuals but is assumed to be constant for a particular individual (no general equilibrium effects, for example). In other respects it is quite general. First, Y_{1i} and Y_{0i} are allowed to capture unintended effects. Second, this framework allows for heterogeneity in the impact of the treatment. Third, while a single treatment has been assumed to keep the exposition simple, the argument applies where there are multiple treatments as well; a treatment can be thought of as ‘a policy regime’).

The essence of the evaluation problem is as follows. For those who receive treatment, only the treated outcome (Y_{1i}) is observed, and for those who are not treated only the untreated outcome (Y_{0i}) is observed. For each individual, the only observed outcome is:

$$Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i} \tag{2}$$

Where D_i is a dummy (0,1 indicator) variable indicating the incidence of treatment.

For no individual is more than one of Y_{1i} and Y_{0i} observed, and thus their difference is never known.

Suppose a researcher wants to estimate the average effect of the Lone Parent Obligations (LPO) on those who are exposed to the treatment ($D_i=1$). This is given by:

$$E [\Delta_i | D_i = 1] = E [Y_{1i} - Y_{0i} | D_i = 1] = E [Y_{1i} | D_i = 1] - E [Y_{0i} | D_i = 1] \tag{3}$$

where $E []$ denotes the expectation (or average), and $E [| D_i = 1]$ denotes a conditional expectation (the average if $D_i=1$). $E [Y_{1i} | D_i = 1]$ is easily estimated using data on outcomes of treated individuals. The problem is: how can $E [Y_{0i} | D_i = 1]$ be estimated when Y_{0i} is not observed for those individuals for whom $D_i = 1$?

One possibility is to use the outcomes of non-treated individuals as a measure of what treated individuals would have experienced had they not received treatment. These are then a comparison group and the comparison gives:

$$E[Y_{1i} | D_i = 1] - E[Y_{0i} | D_i = 0] = E[Y_{1i} - Y_{0i} | D_i = 1] + E[Y_{0i} | D_i = 1] - E[Y_{0i} | D_i = 0] = E[\Delta_i | D_i = 1] + (E[Y_{0i} | D_i = 1] - E[Y_{0i} | D_i = 0]) \quad (4)$$

Comparing the observed outcomes of those who are and are not exposed to the LPPC will yield a biased estimate of the average effect of treatment on the treated if the treated would have had different outcomes than non-treated in the absence of treatment. The second term on the right-hand-side of the equation above captures this bias, often called ‘selection bias’. For example, if the comparison group is composed of lone parents with children too young to be affected by the LPPC, then they will almost certainly have different exit rates from benefit and employment rates in the absence of the LPPC. The key issue in thinking about different possible estimators is how Y_{0i} varies across individuals.

3.2 Some alternative evaluation strategies

This section briefly reviews some research designs that the research team concluded were not appropriate for this particular impact evaluation, and the reasons for those conclusions.

3.2.1 Simple-difference, regression, and matching

The first possibility is a simple comparison of lone parents directly affected by the LPPC with comparison groups drawn from lone parents who are ineligible or exempt. As noted in Section 2.1, the former would be lone parents with a younger child, and the latter would include lone parents who were receiving a disability benefit or Carer’s Allowance, or who are fostering. It seems obvious that such comparisons would not be credible: the variables that distinguish these groups (age of youngest child, disability) certainly affect the outcomes of interest (employment, transitions off benefit, etc.) and so the outcomes of such lone parents would be a poor guide to the counterfactual untreated outcomes of the treated group.

A variant of this would be to use data on the treated lone parents in the past (ie when their children were younger) to estimate their counterfactual outcomes. This is the time-series or longitudinal version of the simple-difference design. However, as noted above, lone parents exposed to the treatment were different in the past in the critical respect that their children were younger, and age of youngest child is an important determinant of labour supply among lone parents. This would, therefore, not be a compelling comparison.

Finally, one might think of choosing an earlier cohort of lone parents with children of the same age as those going through the LPPC. However, such a comparison would capture both the impact of the LPPC and any change in outcomes due to changing economic circumstances over time (and such changes might be thought to be significant, given that the LPPC took place during a recession).

With rich data it is sometimes possible to ‘adjust’ the outcomes of a comparison group for differences between the treatments and comparison group. This can improve the estimate of the counterfactual drawn from the comparison group, and so lead to a more accurate estimate of the impact of a policy. This can be done with a variety of methods including regression, matching or reweighting. For example, a regression-based approach would assume that untreated outcomes depend upon a set of explanatory variables, or covariates, X_i :

$$Y_{0i} = \alpha X_i + u_i \quad (5)$$

These methods work well when two conditions are met:

- there must be **common support** in the variables that will form the basis for adjustment (the covariates, X_i). For example, to adjust for differences in age between the treatment and comparison groups, there must be individuals in the comparison group of the same age as each of the members of the treatment group. If all of the comparison group members are under 40, and all the treatment group members are over 40, then no credible adjustment for age differences is possible;
- regression and matching methods rest on **conditional mean independence**. This means that, having conditioned on the available covariates, the adjusted untreated outcomes of the treatment and comparison group must be (on average) the same. (In the notation above, this means that $E[u_i | D_i = 1] = E[u_i | D_i = 0]$.) Note that this can only be an assumption: the untreated outcomes of the treatment group are not observed. Nevertheless, it is possible to reflect on the plausibility of the assumption.

It is difficult to believe that either of these conditions are met by the LPPC:

- individuals in the potential comparison groups belong to those groups precisely because they have characteristics that are not shared by members of the treatment group. For example, if the comparison group of lone parents with younger children is used, there are no members of that group who have a youngest child the same age as the youngest child of members of the treatment group. This would mean that common support does not exist, and so it is not possible to adjust credibly for age of youngest child (as a regression covariate or matching variable);
- on the other hand, if this variable is not adjusted for, conditional independence is unlikely: because age of youngest child is an important determinant of labour supply, the untreated employment outcomes of the two groups are very likely to be quite different.

Put more simply, the potential comparison groups differ from the treatment group in dimensions that are very likely to affect the outcomes of interest; moreover, they are distinctly different from the treatment group in these dimensions, which precludes adjusting for these differences through regression or matching techniques.

3.2.2 Regression discontinuity

There are many circumstances in which treated and untreated individuals take on discretely different values of a key variable. In such cases, there is a 'cut-off' value, which separates the groups. In a hypothetical example, students with grades above a certain cut-off will receive the treatment of a scholarship, while students with lower grades are not treated; similarly, candidates or parties receiving more than 50 per cent of the vote (in a two-party system) will win an election, and candidates with less than 50 per cent of the vote will not.

Such circumstances often lend themselves to a regression discontinuity design. This is an old idea (Thistlethwaite and Campbell, 1960) which has recently received renewed and intense interest from economists. The basic idea is that individuals just below the cut-off may be very similar to individuals just above the cut-off, and so a reasonable estimate of the programme effect can be obtained by comparing the outcomes of individuals just below and above the cut-off.

This idea can be implemented in a number of ways (for a very good guide to practice see Imbens and Lemieux, 2007). Strengths of this method include the fact that, under certain conditions, the comparison of individuals just below and above the cut-off is as good as randomisation (in a way that can be made precise). Thus, under certain conditions there is a strong case for the internal validity of the method. Moreover, there are a range of fairly intuitive and easy-to-implement specification tests that can be applied (again see Imbens and Lemieux).

A weakness of the method is that it really only delivers an estimate of the policy effect at the cut-off. If, for example, the effect of a scholarship on the outcomes of students far above the scholarship cut-off is very different from the effect on students near the cut-off, then little can be learnt about the former group using a regression discontinuity (RD) design. In this sense, the RD design does not do well on external validity.

A RD design was used by de Georgi (2005) to evaluate the impact on New Deal for Young People (NDYP), by comparing unemployed young adults who were 24 (and therefore, affected) and 25 (not affected). Similarly, Mulheirn and Pisani (2008) examine the effect of the Working Tax Credit for childless adults without children over the age of 25 by comparing changes in labour market participation just above and just below the cut-off. Lemieux and Milligan (2007) look at the effect of higher unemployment benefits for those over 30 on labour market participation by comparing the employment rates of those aged just above and below 30. At first consideration it might seem that evaluation of the LPPC is an ideal setting for a RD design because there is a sharp cut-off in the age of youngest child in determining entitlement to IS as a lone parent. However, further reflection suggests that this is not the case. The key requirement for an RD design is that the probability of treatment is discontinuous in the forcing variable (here, age of youngest child). But this does not occur with the LPPC – although there is a sharp change in the benefit regime faced by lone parents when they lose entitlement to IS, because of anticipation effects, there is no discontinuity in the effect of the treatment at this point. This is because lone parents know that they are about to face this change and may not be able to time their response to it perfectly. For example, if a lone parent has decided to start work when they lose IS entitlement, it is unlikely they will be able to find a job immediately at this point and so they might actually start work shortly before or afterwards. This contrasts, for example, with the scholarship case: a student that just misses the scholarship cut-off does not anticipate receiving the scholarship a month later. Similarly, those who want to leave work when their benefits increase when they reach the appropriate age can simply quit their jobs at this point. However, if lone parents with a youngest child just below the cut-off age change their behaviour as a result of knowing that they will soon be treated, then the regression discontinuity design is invalidated. These effects are highly likely, and so the regression discontinuity design does not seem appropriate for an evaluation of the LPPC.

3.2.3 ‘The Norway Approach’

Pronzato and Morgstad (2008) estimate the effect of a Norwegian welfare reform by looking at the outcomes of mothers in couples who become lone parents before and after a reform is introduced. In order to control for wider economic trends, they use a DiD estimate in each period, comparing what happens to those who become lone parents with those who do not. This does not rely on the ‘common trends’ assumption so much as a pure DiD approach, but it does require that any differential trend in the outcomes of mothers who remain in couples and those who become lone parents is constant over time. However, there are two reasons why this approach is not recommended:

- this study uses register data on the whole Norwegian population. No such dataset exists in the UK, and there is no way of identifying women who become lone parents in the administrative data that is available;
- this approach estimates the effect of the programme on those women who become lone parents. These estimates are, therefore, unlikely to have external validity to the population of lone parents as a whole. Also, this approach cannot be used to estimate the effect of the LPPC on lone parents already claiming IS, who are one of the main groups affected by it.

3.3 The difference-in-difference design

Suppose data exists on individuals before the introduction of the treatment. In what follows, outcomes (such as employment) in the period after treatment are denoted as $Y_{1i,t}$ and $Y_{0i,t}$, and outcomes in the period before treatment as $Y_{1i,t-1}$ or $Y_{0i,t-1}$.

If the treatment does not have any effect before it occurs (as might be the case if the treatment is not anticipated), then $Y_{1i,t-1}$ and $Y_{0i,t-1}$ are equal. However, because of the same kinds of ‘selection bias’ reasons given above, it would not be wise to assume that the average pre-treatment outcomes of the treatment and comparison groups are the same. Thus:

$$E[Y_{0i,t-1} | D_i = 1] \neq E[Y_{0i,t-1} | D_i = 0] \quad (6)$$

The DiD design compares changes (that occur when the treatment is introduced) in outcomes for treated and untreated individuals. Formally, the DiD estimate is given by:

$$\begin{aligned} & E[Y_{1i,t} - Y_{1i,t-1} | D_i = 1] - E[Y_{0i,t} - Y_{0i,t-1} | D_i = 0] \\ &= E[(Y_{1i,t} - Y_{1i,t-1}) - (Y_{0i,t} - Y_{0i,t-1}) | D_i = 1] + E[Y_{0i,t} - Y_{0i,t-1} | D_i = 1] - E[Y_{0i,t} - Y_{0i,t-1} | D_i = 0] \\ &= E[(Y_{1i,t} - Y_{0i,t-1}) - (Y_{0i,t} - Y_{0i,t-1}) | D_i = 1] + E[Y_{0i,t} - Y_{0i,t-1} | D_i = 1] - E[Y_{0i,t} - Y_{0i,t-1} | D_i = 0] \\ &= E[\Delta | D_i = 1] + (E[Y_{0i,t} - Y_{0i,t-1} | D_i = 1] - E[Y_{0i,t} - Y_{0i,t-1} | D_i = 0]) \end{aligned} \quad (7)$$

where the third line uses the fact that $Y_{1i,t-1}$ and $Y_{0i,t-1}$ are equal, and the final line used the definition $Y_{1i} - Y_{0i} = \Delta_i = \Delta$ (for simplicity, assume that the treatment effect is equal for all treated individuals).

This is a good estimator of the true impact if:

$$(E[Y_{0i,t} - Y_{0i,t-1} | D_i = 1] - E[Y_{0i,t} - Y_{0i,t-1} | D_i = 0]) \approx 0 \quad (8)$$

This says that, in the absence of treatment, changes over time in the outcomes would be similar for the two groups (treated and untreated); this is normally referred to as the common trends assumption.

Equation (8), which says that the average **change** in untreated outcomes is the same in the two groups, contrasts with Equation (6), which says that the average **level** in untreated outcomes is not the same. If the common trends assumption is true, then the comparison group provides good information about the counterfactual because changes in the outcomes of the comparison group tell us how the outcomes of the treatment groups would have evolved in the absence of treatment.

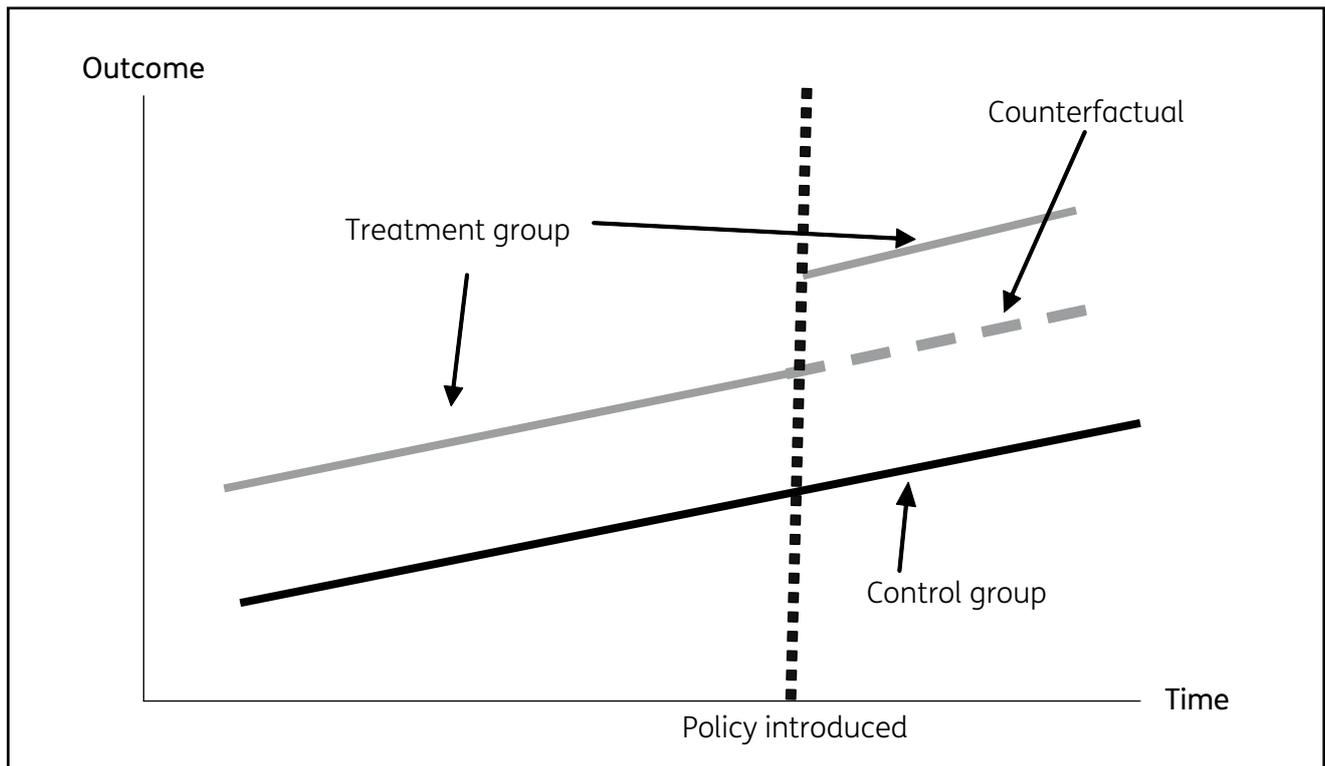
The common trends assumption will hold in practice if:

$$E[Y_{0it} | t, D] = \alpha_t + \gamma_D \quad (9)$$

In other words, it will hold if untreated outcomes depend on a ‘time effect’, and a ‘group effect’, but where the time effect is common to the groups (treatment and comparison) and the group effect (the difference between the treatment group and the comparison group) is constant through time. Crucially, there is no interaction effect, and this is what allows us to attribute any difference in the changes over time in the two groups to the treatment.

This is illustrated in Figure 3.1. Outcomes for the treatment and comparison groups have different levels, but would have the same trend if there were no treatment. The comparison group can, therefore, be used to infer what would have happened to the treatment group in the absence of treatment (the counterfactual). Subtracting the change experienced by the comparison group from the change experienced by the treatment group gives a good estimate of the effect of the treatment.

Figure 3.1 Comparison of the treatment and comparison groups after treatment is introduced



Noting that observed outcomes are $Y_{it} = Y_{0it} + \Delta D_{it}$, the ‘common trends’ assumption about the variation Y_{0it} implies that:

$$Y_{it} = \alpha_t + \gamma_D + \Delta D_{it} + u_{it} \tag{10}$$

Where $u_{it} = Y_{0it} - E[Y_{0it} | t, D]$ (and thus $E[u_{it} | D_{it} = 1] = E[u_{it} | D_{it} = 0]$).

In the simplest implementation of the DiD design, there are just two groups (treatment and comparison) and two time periods (before and after). The time and group effects can be captured then, by a single dummy for the treatment group, and a single dummy for the ‘after’ period. Thus:

$$E[Y_{0i} | t, D] = \alpha_t + \gamma_D = \alpha + \beta A_{it} + \gamma T_{it} \tag{11}$$

and:

$$Y_{it} = \alpha_0 + \alpha_1 A_{it} + \gamma T_{it} + \Delta D_{it} + u_{it} \tag{12}$$

where A_{it} is a dummy equal to one in the ‘after’ period and zero otherwise; T_{it} is a dummy equal to one if the individual belongs to the treatment group and zero otherwise; D_{it} is the dummy indicating receipt of treatment, and u_{it} is as above. The parameters of this regression can be related to the mean outcomes of different groups in different periods, as Table 3.1 illustrates.

Table 3.1 A simple difference-in-difference design

Mean (average) outcomes	Treatment group	Comparison group	Difference between groups
After	$\alpha_0 + \alpha_1 + \gamma + \Delta$	$\alpha_0 + \alpha_1$	$\gamma + \Delta$
Before	$\alpha_0 + \gamma$	α_0	γ
Difference over time	$\alpha_1 + \Delta$	α_1	Δ

Table 3.1 also illustrates an alternative formulation of the common trends assumption. Assuming that the time effects are common to groups is the same as assuming the group effects are constant through time. Thus, the DiD estimator can be thought of in two equivalent ways. Either:

- compare the change in the treatment group with the introduction of policy ($\alpha_1 + \Delta$), and use the comparison group to estimate the counterfactual time trend (α_1); or
- compare the treatment group and comparison group after the policy change $\gamma + \Delta$, and then use the pre-treatment period to estimate the difference between the groups which is unrelated to treatment (γ).

The latter intuition is a natural way to think about this approach to the LPPC evaluation. During the policy roll-out, outcomes of groups of lone parents affected by the policy can be compared with groups that are not affected. However, these groups differ in the age of the youngest child. Section 3.2 discussed alternative research designs and noted that if comparison groups are chosen that were very close to the treatment group in terms of age of youngest child (as in the regression discontinuity design), anticipation effects would confound the comparison, because the comparison group would anticipate treatment in the very near future. On the other hand, a simple comparison between affected lone parents and lone parents with much younger children would not be compelling because the labour supply behaviour of such a comparison group would be very different, and such differences were directly related to the age of youngest child, and so difficult to eliminate through controlling for other variables.

The DiD design offers a way out of this dilemma: lone parents affected by the LPPC can be compared with lone parents with sufficiently younger children to diminish the likelihood of anticipation effects, and then data from periods before the LPPC roll-out can be used to estimate the difference in outcomes between the treatment and comparison groups that results simply from the differences in the ages of their youngest children and that is unrelated to the LPPC. As in Table 4.1, the direct comparison of treatment and comparison groups confounds the treatment effect and group effects; but the common trends assumption implies that the group effects are time-invariant, so the pre-treatment period can be used to estimate the group effects; having done so, they can be removed from the treatment period differences to isolate the treatment effect.

3.3.1 Extending the DiD design: more than two groups or time periods

This sub-section considers extensions to the DiD design where there are more than two groups or more than two time periods (including situations, like LPPC, where the policy under investigation has different start dates for different groups).

The simple (two group, two period) DiD design generalises naturally to a setting – such as the evaluation of the LPPC – where there are multiple comparisons involving multiple groups (defined by the age of the youngest child), multiple time periods and multiple start dates for the policy. In such situations, standard DiD design continues to require the common trends assumption (ie that $E[Y_{0it} | t, D] = \alpha_t + \gamma_D$), but the regression $Y_{it} = \alpha_t + \gamma_D + \Delta D_{it} + u_{it}$ now contains a full set of time dummies, and a full set of group dummies (capturing α_t and γ_D) respectively.

The model can also be extended to allow for different treatment effects at different times or for different groups; this will be important when allowing for anticipation effects (below). A model with multiple treatment effects can be implemented by estimating a regression of the form:

$$Y_{it} = \alpha_t + \gamma_D + \sum_k \Delta^k D_{it}^k + u_{it} \quad (13)$$

where Δ^k is now the effect of policy k and D_{it}^k is a dummy indicating that individual i is exposed to policy k at time t . This sort of model was used in Brewer *et al.* (2009) to estimate the impact of the lone parent pilots, which had different start dates in each of the four sets of districts in which it was implemented.

If data is available for multiple time periods before the policy begins, then it is possible to estimate a trend-adjusted DiD model (see section IV.C.2 of Blundell and Costa Dias, 2008 or p444 of Blundell and Costa Dias, 2008; see also Bell *et al.*, 1999). A trend-adjusted DiD model allows the common trends assumption to be relaxed. If outcomes for the treatment and comparison group follow different trends, then the conventional DiD estimator will estimate the sum of the true impact of the policy plus a contribution from the diverging trends. But, given sufficient pre-programme data, and on the assumption that the difference in the trends would be constant in the absence of any policy change, the different trends can be estimated, and so the DiD estimator corrected (see section IV.C.2 of Blundell and Costa Dias, 2008 for full derivation).

3.3.2 Threats to the internal validity of the DiD design

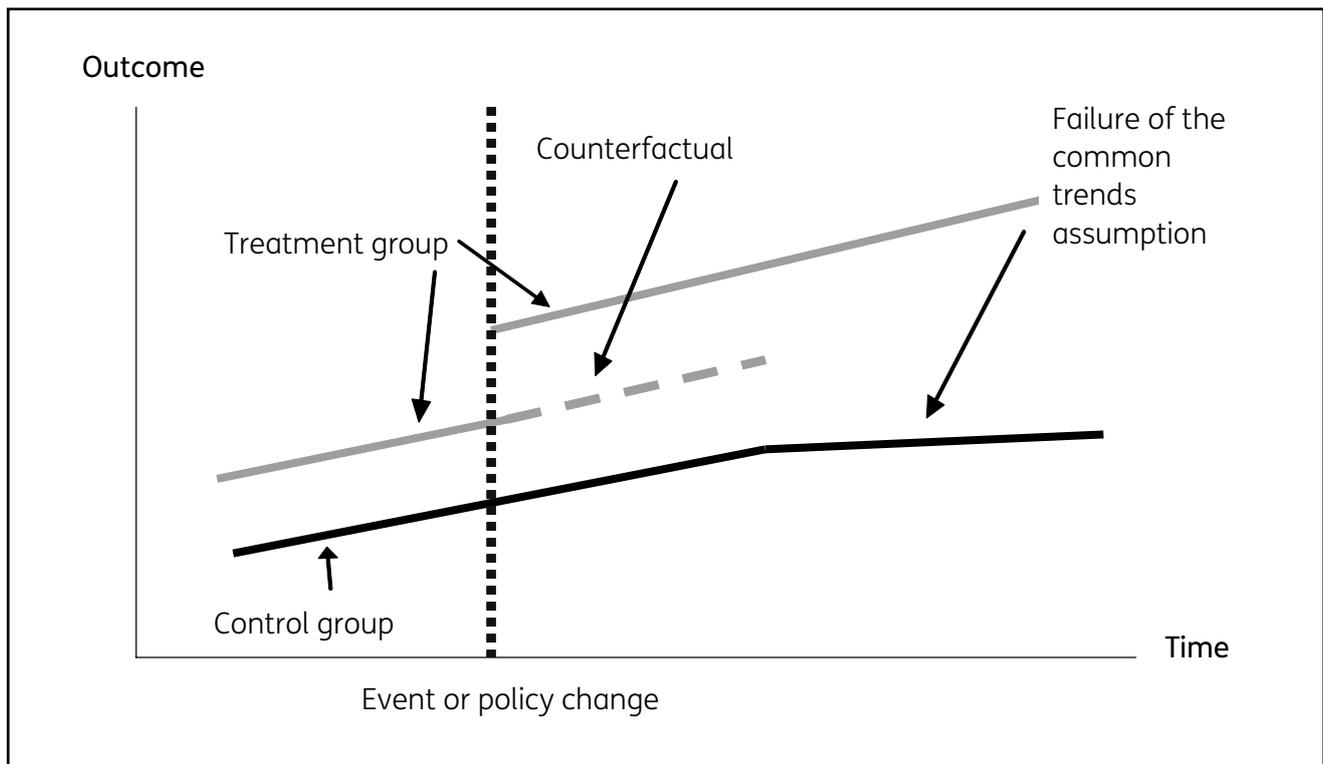
The key assumptions of the DiD design are:

- there are no anticipation effects;
- the common trends assumption.

As shown above, the formal derivation of the DiD approach rests on the absence of a treatment effect in the pre-treatment period (so that $Y_{1i,t-1}$ and $Y_{0i,t-1}$ are equal). Less formally, the immediate problem that anticipation effects cause is that treatment group outcomes are altered by the treatment in both 'pre' and 'post' treatment periods, so that differencing the data eliminates some part of the treatment effect. (The anticipation of treatment can cause further, more subtle problems, which are discussed below in the context of the common trends assumption.)

Because the LPPC was announced sometime before it was implemented, and because eligibility is very predictable (depending, in very approximate terms, on the age of the youngest child and calendar time), one might expect behavioural responses to occur before individuals are actually exposed to the obligation. As Section 2.1 discussed, in the case of the LPPC, it makes sense to define the treatment as beginning 12 months before the date on which the lone parent loses entitlement to IS as a lone parent, as this is when the quarterly WFIs and other services begin. It would also be sensible to allow the policy effect to vary according to whether loss of entitlement is anticipated or has actually occurred, and this can be done with a specification including multiple treatment dummies and multiple treatment effects (as described above).

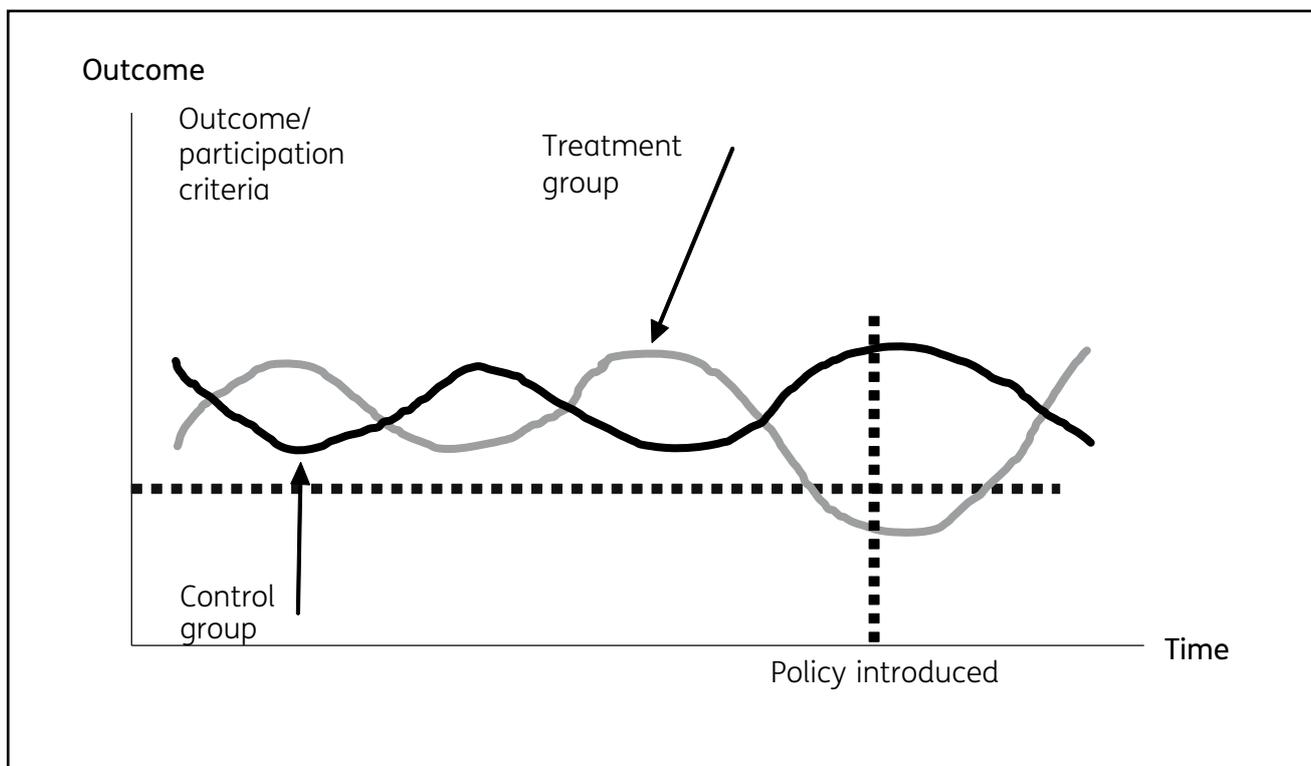
Therefore, the key concern is with the common trends assumption. Failure of the common trends assumption is illustrated in Figure 3.2.

Figure 3.2 Results of failure of the common trend assumption

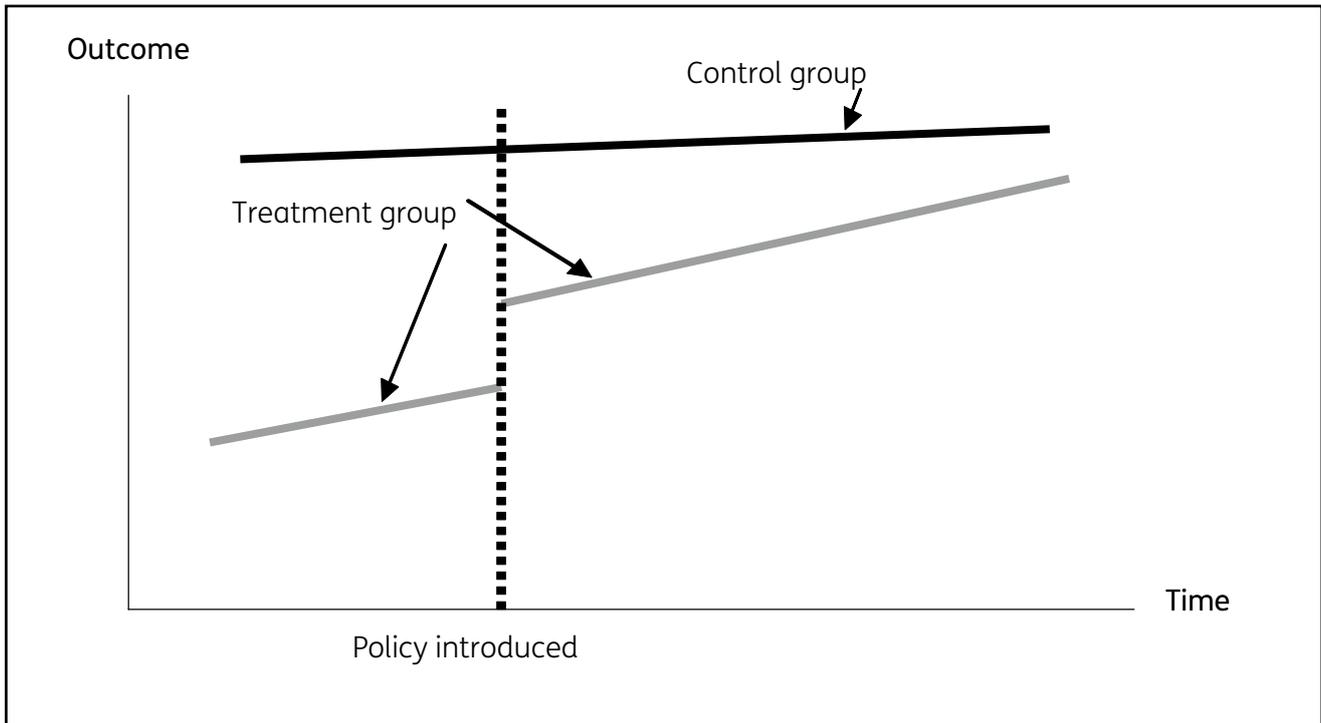
A reading of the literature suggests six key reasons why the common trends assumption fails to hold. These are discussed below, along with an assessment of whether any is likely to be relevant to the LPPC evaluation.

- 1 **Treatments and comparison group are drawn from different local labour markets.** In this case, the comparison might experience (for example) different increases in employment opportunities (if the economy of the local area from which the comparison group are drawn is growing more strongly than the local area from which the treatments are drawn). A well known illustration of this problem is Angrist and Krueger's reanalysis of Card's analysis of the Mariel Boatlift (Angrist and Krueger, 1999; Card, 1990). As treatments and comparison groups can be drawn from the same local labour markets, this should not be a problem for the LPPC evaluation.
- 2 **The common trends assumption will fail if programme eligibility depends on lagged outcomes, and those lagged outcomes are mean-reverting.** For example, if incomes vary from year to year, and programme eligibility depended on income being below a threshold, then the treatment group will be composed of individuals with temporarily low incomes (that would be expected to rise even in the absence of treatment) and the comparison group will be disproportionately drawn from individuals with temporarily high incomes (which can be expected to fall). This is a version of 'Ashenfelter's dip' (Ashenfelter, 1978). It will lead to a spurious estimate of a positive treatment effect (and is shown visually in Figure 3.3). Another well-known example of this is when schools are eligible for a programme on the basis of low test scores (Chay *et al.*, 2005). Again, this is not relevant for the LPO evaluation because eligibility for treatment does not depend on lagged outcomes. By definition, eligibility for the LPPC depends upon being on IS, but this will also be true for the suggested comparison group: the difference between treatment and comparison groups will be based on the date of birth of the youngest child.

Figure 3.3 Failure of the common trend assumption if programme eligibility depends on mean-reverting lagged outcomes



- 3 The outcome is bounded, and the comparison group is near the bound (see Figure 3.4). Eissa and Liebman (1996) used childless single women as a comparison group for female lone parents when examining the impact of the US earned income tax credit on employment rates. One potential problem is that, in the US, the employment rates of single women exceeded 90 per cent. This means that there is little scope for childless single women to increase their employment rates, but lone parents, with lower employment rates, do have greater scope to respond to macroeconomic developments; essentially, a situation with outcomes near a bound makes the common trends assumption much less plausible (see Figure 3.4). As noted above, the suggested comparison group to evaluate the LPPC comprises lone parents who have children of different ages. These groups have baseline employment or benefit receipt rates that, while a little different from the treatment group, are far below 100 per cent (and nowhere near as dissimilar as single women with and without children). Thus, this problem seems irrelevant for the LPPC evaluation.

Figure 3.4 Result if the outcome is limited and the comparison group is near it

- 4 **Differential effect of coincident policy changes.** If there are other policy changes occurring around the time of the treatment in question, and if those changes have a differential impact on treatment and comparison groups, then the change over time observed for the comparison group will be a poor guide to what the counterfactual time path of outcomes would have been for the treatment group, in the absence of the treatment in question. This is a potential problem in the evaluation of the LPPC because of the large number of concurrent policy innovations. Nevertheless, this problem can be handled empirically, and this is discussed further in Section 6.2.
- 5 **Group composition changes, including those associated with the anticipation of treatment.** The DiD design assumes that the composition of the group is constant. If the composition of the treatment or comparison group changes over time, it is difficult to disentangle the effects of the treatment from the effects of changing composition. One way that this can happen is if group membership is defined by variables that are not completely time invariant, and, in particular, if group membership might be affected by the anticipation of treatment. If compositional changes are driven by the anticipation of treatment, then they will affect treatment and comparison groups differently, so that the trends in the two groups will be different (see also Section 6.1).

The nature of the Work and Pensions Longitudinal Study (WPLS) data, combined with the fact that the LPPC is likely to be anticipated, mean this is a potential problem for the evaluation of LPPC that makes use of WPLS data. The WPLS contain only claimants, and so the groups will be defined as those lone parents receiving IS. If earlier anticipation of treatment (that is, anticipation more than 12 months before) induces a non-random subset of lone parents to leave IS before treatment, then the composition of the treatment group will change over time. This would invalidate the ‘common trends’ assumption (for further discussion of this type of problem, see Blundell and Costa Dias, 2008). This problem can be overcome empirically by assuming the LPPC starts to affect lone parents more than 12 months before they lose IS entitlement and selecting a sample of lone parents on IS, say, 24 months before they lost IS entitlement. This would ensure that selection into the sample was unaffected by the LPPC.

- 6 **Spillover effects affecting the comparison group.** The DiD design requires that the comparison group are completely unaffected by the existence of the policy. If the LPPC increases labour supply among the treatment group and this affects wages and the availability of jobs more generally, this assumption will be violated and the DiD estimator will be incorrect; it will incorrectly classify these negative spillover effects affecting the comparison group as positive treatment effects affecting the treatment group. This is likely to be a problem with an eventual IA as the treatment and comparison groups are drawn from the same labour market and are likely to be substitutable for each other from the point of view of employers. However, the size of any such spillover effects should be small, given that the treatment group is small relative to the overall size of the labour market, and given the likely size of the policy effects.

3.4 Choosing a comparison group

Section 3.3 argued that the most suitable comparison group is lone parents whose children are too young to be directly affected by the LPPC. Appendix B lists exactly which comparison groups (defined by the dates of birth of the youngest child) would be in each phase of the roll-out. There are several important points to note:

- there is a trade-off between examining outcomes for a longer period of time or allowing for longer anticipation effects, and reducing the difference between the treatment and comparison groups in terms of the age of the youngest child. The calculations in Appendix B assumed a maximum period of interest of four years, which would include two years of anticipation effects and two years of after effects, or one year of anticipation effects and three years of after effects.
- there is also a trade-off between the size of the comparison groups for the stock phases, and the difference in age of youngest child between the treatment and comparison groups.
- there are larger differences in the age of the youngest child between the treatment and comparison groups at the start of the roll-out than towards the end, or when evaluating the steady-state policy.
- lone parents whose youngest child was born between October 2004 and October 2006 are the comparison group for a large number of phases of lone parents affected by the LPPC. If a shock affects this group but not the groups affected by the treatment at this time, then the evaluation of a large number of the phases of the roll-out will be invalidated.

It is also important to consider which period should form the ‘before’ period for any DiD approach. There is a concern in this case because Work Focused Interviews (WFIs) were rolled out gradually by age of the youngest child between April 2001 and April 2004, meaning that the treatment and comparison groups were also being treated differently over time. There are two possible solutions to this problem:

- only the period after April 2004 (and/or the very short period before April 2001) could be used as the ‘before’ period. This would ensure that there was no difference in the WFI regimes faced by the treatment and comparison groups, but it might be the case that there was not enough time between then and the introduction of the LPPC to provide robust evidence on the relationship between labour market outcomes and age of youngest child in the absence of the LPPCs;
- alternatively, the different WFI regimes could be explicitly modelled in the DiD model. This would involve adding dummy variables to the model indicating which WFI regime a lone parent was subject to at a particular time. Using this methodology would involve implicitly assuming that the effect of a particular WFI regime was uniform, both over time and age of youngest child. But it would allow more data to be used to estimate the relationship between labour market outcomes and age of youngest child in the absence of the LPPCs.

3.5 Assessing whether common trends holds between lone parents affected by the LPPC and lone parents with younger children

Sections 3.3 and 3.4 argued that, in principle, DiD method, using lone parents with younger children as a comparison group, would be an appropriate method to use in an eventual IA.

This section assesses empirically the suitability of DiD with lone parents with younger children as a comparison group. It does this by testing whether the common trends assumption holds in the period before the LPPC was introduced, by applying the DiD methods to data from the period before the LPPC was introduced (this is sometimes known as a placebo test, or pre-programme test).

The argument for performing these pre-programme tests is as follows: if data has been chosen from a period where the policy regime did not change differentially for lone parents with younger and older children, then it should be the case that the DiD method finds no significant evidence of a policy effect. If, however, the DiD method does find significant evidence of a policy effect where none should exist, then this suggests that the method or the chosen comparison group are inappropriate.⁵

To implement these pre-programme tests, it was necessary to replicate as closely as possible the groups who would be affected by the LPPC (as defined by the age of their children), and compare these with the same comparison groups that would be used by an eventual IA, but choosing lone parents observed before LPPC began. This was done as follows:

⁵ It should be borne in mind however, that some potential problems cannot be detected by this procedure. In particular, pre-programme data cannot be used to assess threats to the ‘common trends’ assumption that are relevant only to the treatment period, such as points 4 and 5 in Section 3.3.2. Therefore, the common trends assumption may fail during the treatment period even if it was shown to have held during the pre-treatment period. A particular concern for this IA is that the LPPC is being implemented during a recession, and there is no pre-programme data to test whether common trends held during previous recessions.

- using the WPLS, all lone parents who are on IS on five particular dates (1 April of 2001, 2002, 2003, 2004 and 2005) were selected;
- for each of these five samples, it was determined which lone parents would be affected by each phase of the LPPC if it had been introduced one year after each of these dates (e.g. 1 April 2002 for those initially observed on benefit on 1 April 2001)⁶. This calculation was based on the date of birth of the child who was the youngest on the date on which they are selected;⁷
- for each of the five samples, an appropriate comparison group was selected from among the lone parents with younger children on IS on the same date. (See Section 3.4 and Appendix B for details of appropriate comparison groups for each of the phases of the roll-out of the LPPC);
- the outcomes of the lone parents in each group (according to the measure of benefit receipt and employment in the WPLS) are then tracked over the next three years (two years in the case of those selected on 1 April 2005).

This section first presents analysis based on simple average outcomes, and then presents analysis based on regression models to implement the DiD models.

3.5.1 Assessing the suitability of DiD with lone parents with younger children as a comparison group: analysis using simple averages

This section compares, graphically, how quickly lone parents observed at different points in time with differently aged youngest children leave benefit and start work. This provides a visual assessment of the common trends assumption.

Figure 3.5 compares lone parents whose youngest child is aged between 11 and 14 on the date they are sampled (so they are aged between 12 and 15 on the day the hypothetical LPPC comes into effect one year later; this is the group that were affected by the phase 1 of the roll-out of the LPPC) with those whose youngest child is aged two to three on the day they are first sampled (so they are aged three to four on the day the hypothetical LPPC comes into effect one year later; this is the group that is recommended to be used by an eventual IA as the comparison group). Clearly, all are on benefit at the time they are first sampled and then some leave (and some then start new benefit claims).⁸

Benefit outcomes in the period from April 2001 to April 2007 were similar for the two groups, although they differ by around one percentage point at most points in time. If anything, the outcomes for the treatment group (those whose youngest child is aged between 12 and 15) deteriorate slightly relative to those of the comparison group (those whose youngest child is aged three or four).

⁶ We refer to this hypothetical policy change as a 'hypothetical LPPC'.

⁷ This means that we include lone parents who would be affected if the hypothetical LPPC were introduced a year after they are initially observed based on the age of their youngest child at the point they are initially observed (i.e. those whose youngest child was aged six or over when we select our sample). This means that we effectively ignore the arrival of any additional children in the 12 months before the hypothetical LPPC starts. This is because the WPLS only records the arrival of additional children for those who remain on benefit. Using age of youngest child as measured in the WPLS on the date the LPPC starts to select the sample would, therefore, lead to bias. Also, having more children to delay being affected by the LPPC may be a behavioural response that we wish to measure.

⁸ The benefits we include here are the three main out-of-work benefits for lone parents during this period (IS, Jobseeker's Allowance (JSA) and Incapacity Benefit (IB)).

Figure 3.6 shows the same analysis comparing lone parents whose youngest child is aged ten to 11 with those whose youngest child is aged three to four (note that all figures use the same comparison group). This is broadly the group that has been affected by the second stage of the LPO roll-out. Once again, benefit outcomes are broadly similar for lone parents with differently aged youngest children, although are lower by up to 3.5 percentage points for those with children aged ten and 11 by the end of the period. However, those with older children tend to be slightly less likely to leave benefit at all times, and outcomes for those with older children tend to deteriorate slightly over time relative to those with younger children.

Figure 3.7 performs the same analysis for lone parents whose youngest child is aged seven to nine, comparing them once again with lone parents whose youngest child is aged three or four. This is, broadly speaking, the group which will be affected by the third phase of the roll-out of LPO. Outcomes for lone parents whose youngest child is aged seven to nine are particularly close to those of lone parents whose youngest child is aged three or four (as one might expect since they are close in age). However, lone parents whose youngest child is aged seven to nine are slightly slower to leave benefit than those whose youngest child is aged three or four (their outcomes are about one percentage point lower for most of the period), and that their outcomes deteriorate slightly over time relative to those of lone parents whose youngest child is aged three or four, becoming up to 3.5 percentage points lower by the end of the period in question.

Figures 3.8 to 3.10 show the differences in outcomes between lone parents with differently aged youngest children over time. The difference in outcomes between the two groups moves in a similar way over the three years for each cohort, as required by the common trends assumption, although outcomes for those with older children deteriorates relative to those with younger children in later years.

Figure 3.5 Comparing trends of benefit outcomes for lone parents with differently aged youngest children (12 to 15 compared to three to four)

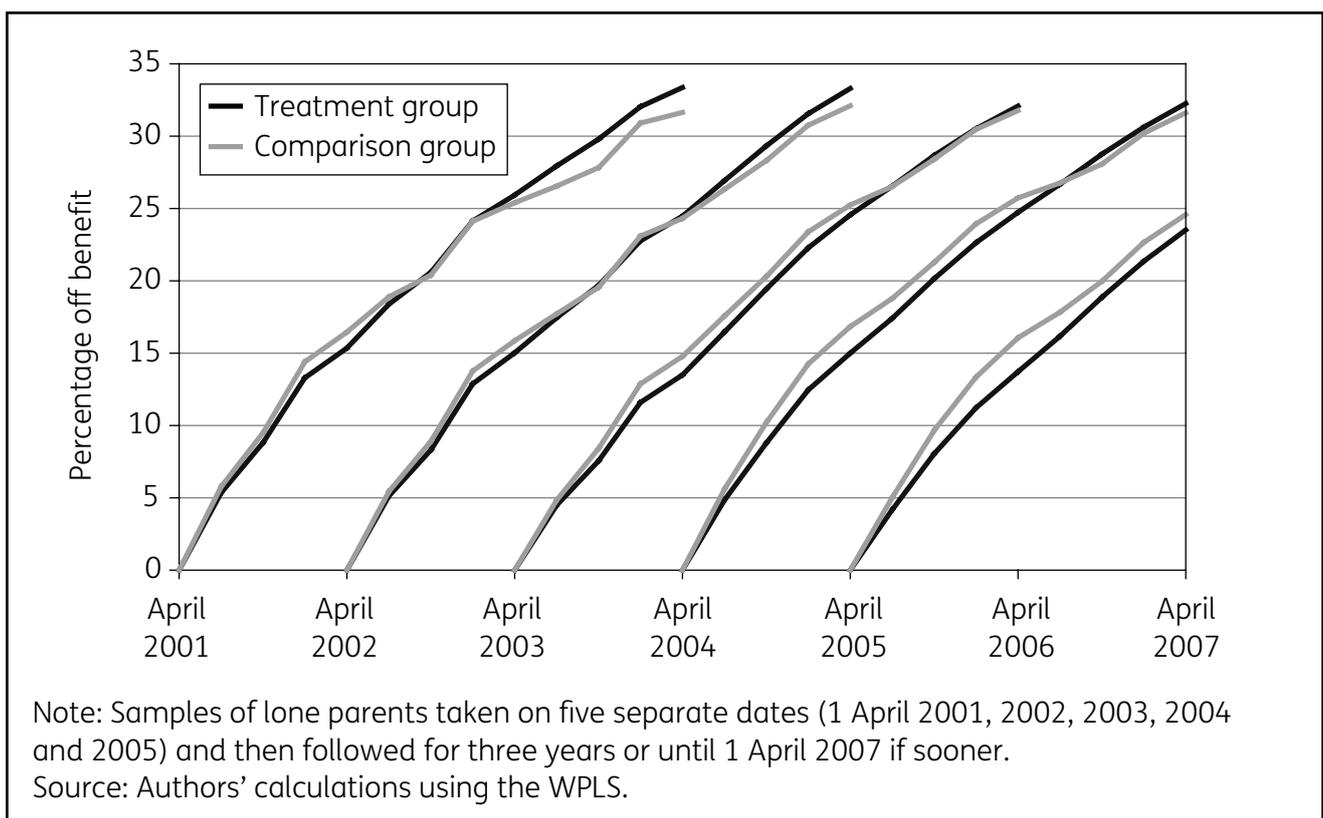


Figure 3.6 Comparing trends of benefit outcomes for lone parents with differently aged youngest children (10 to 11 compared to three to four)

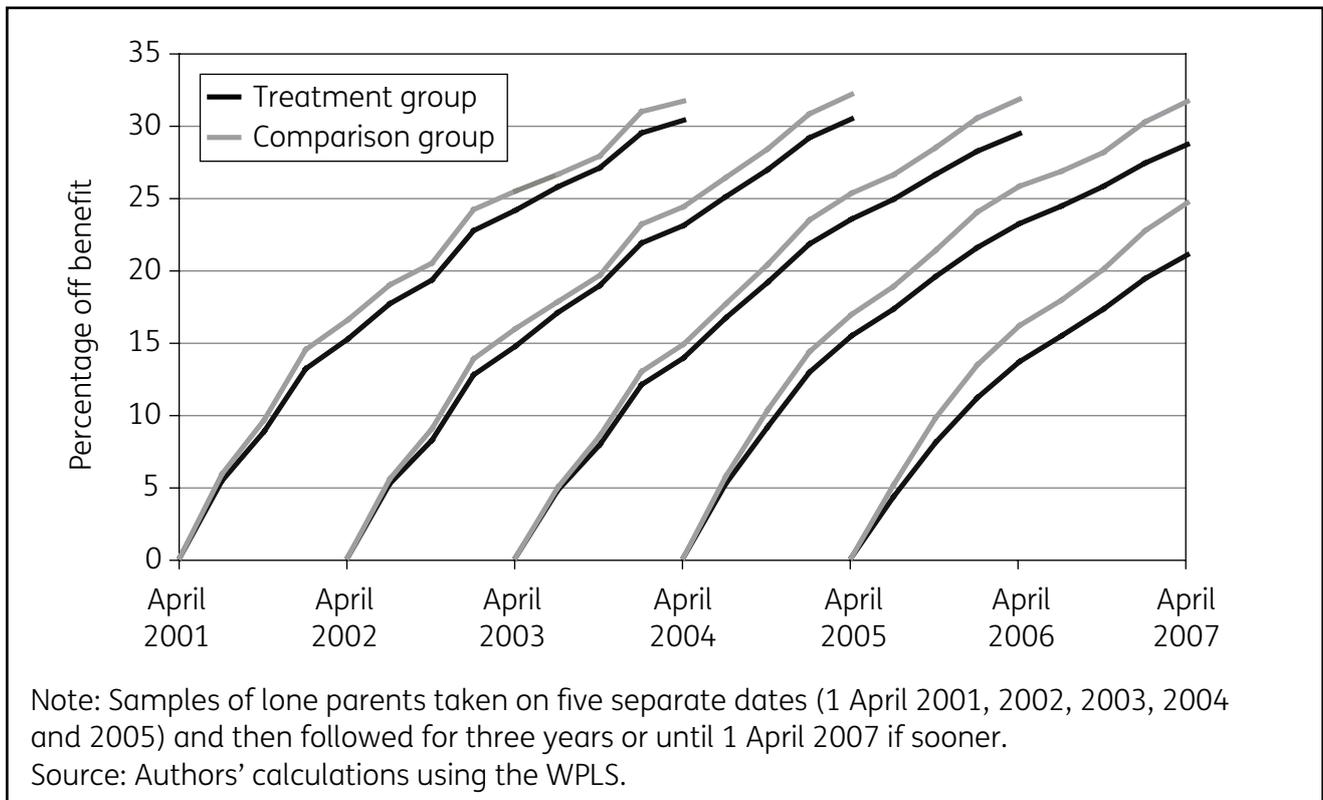


Figure 3.7 Comparing trends of benefit outcomes for lone parents with differently aged youngest children (seven to nine compared to three to four)

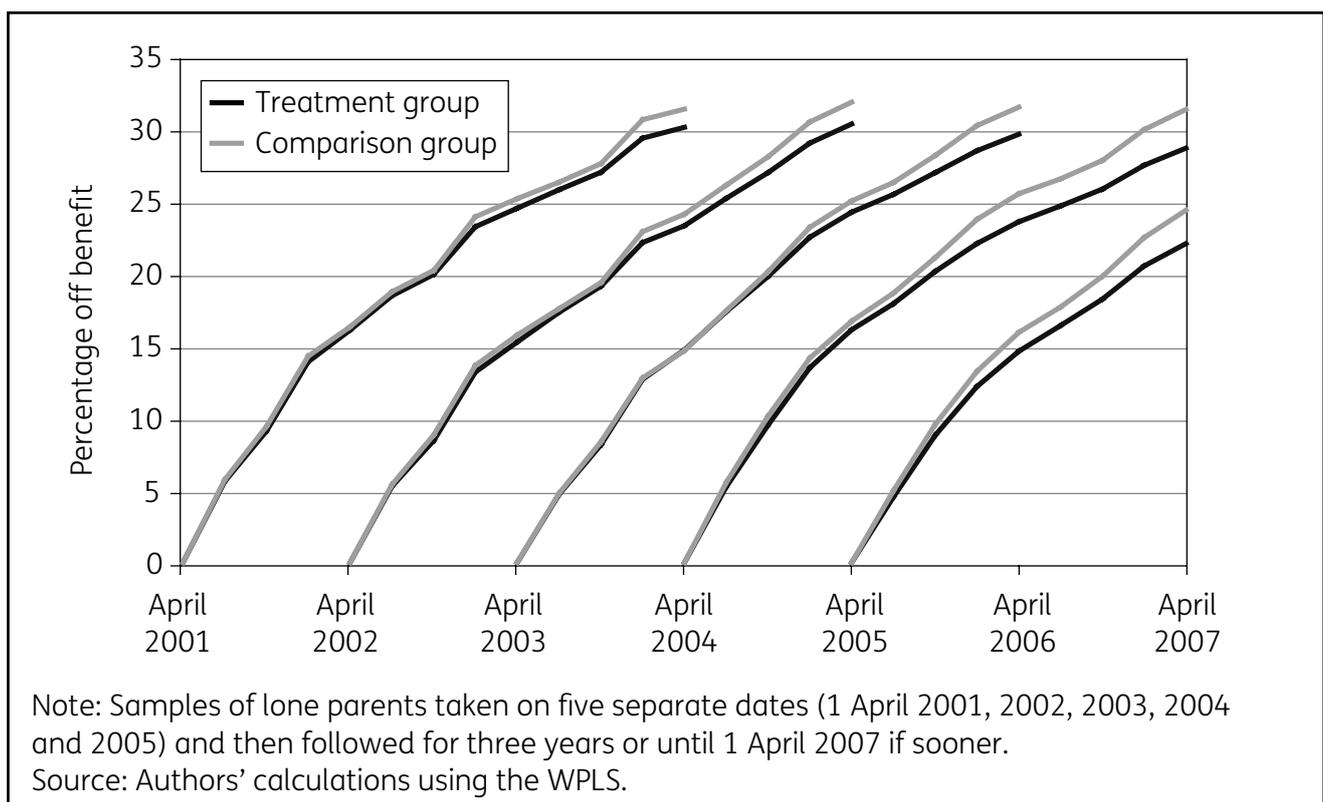


Figure 3.8 Comparing differences in benefit outcomes for lone parents with differently aged youngest children (12 to 15 compared to three to four)

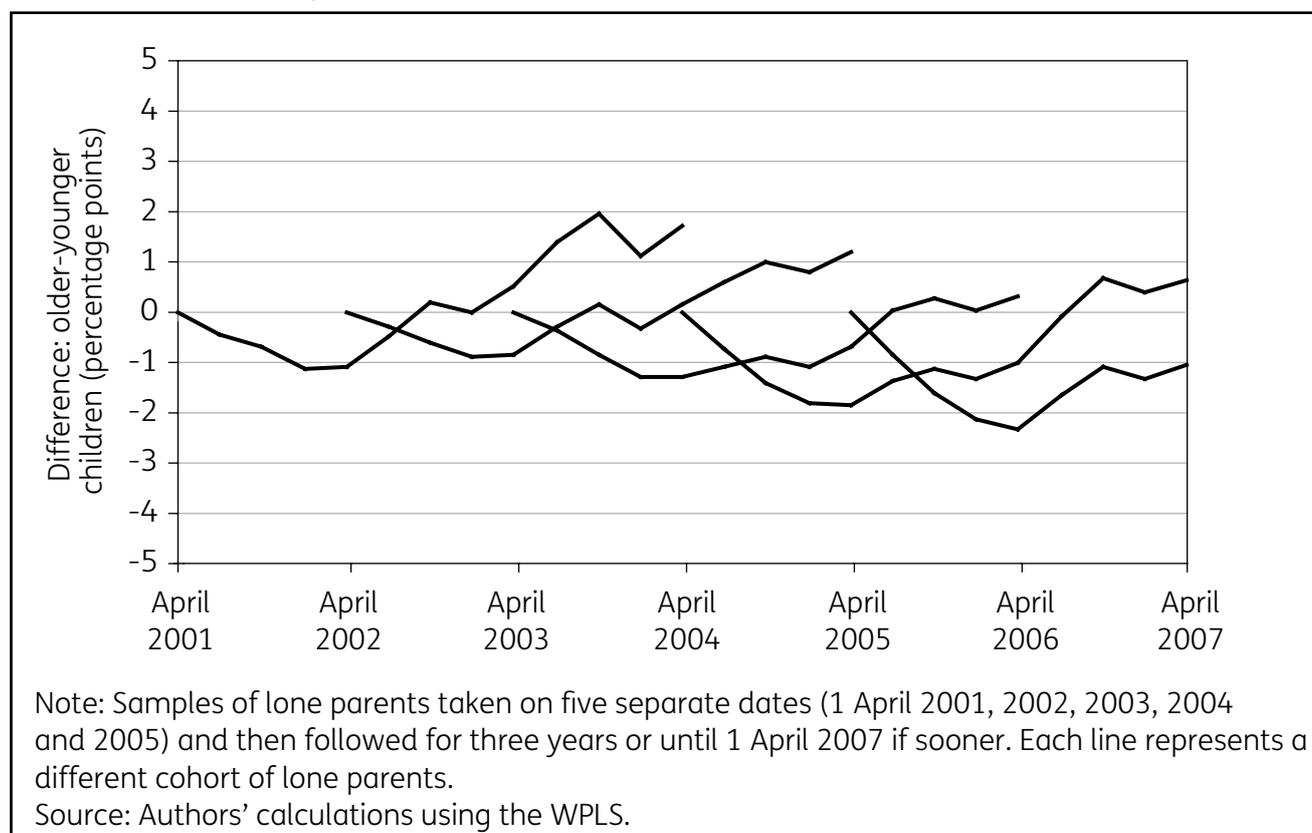


Figure 3.9 Comparing differences in benefit outcomes for lone parents with differently aged youngest children (10 to 11 compared to three to four)

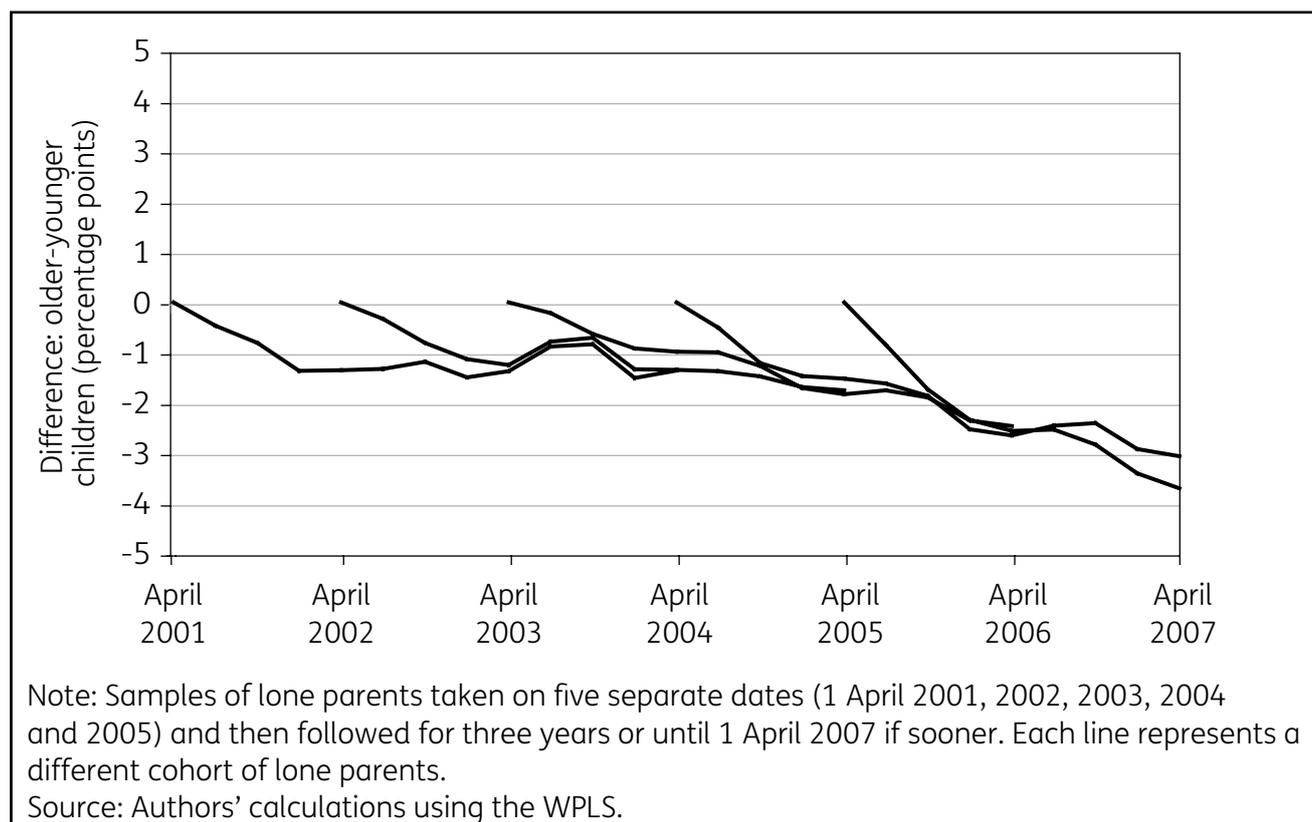
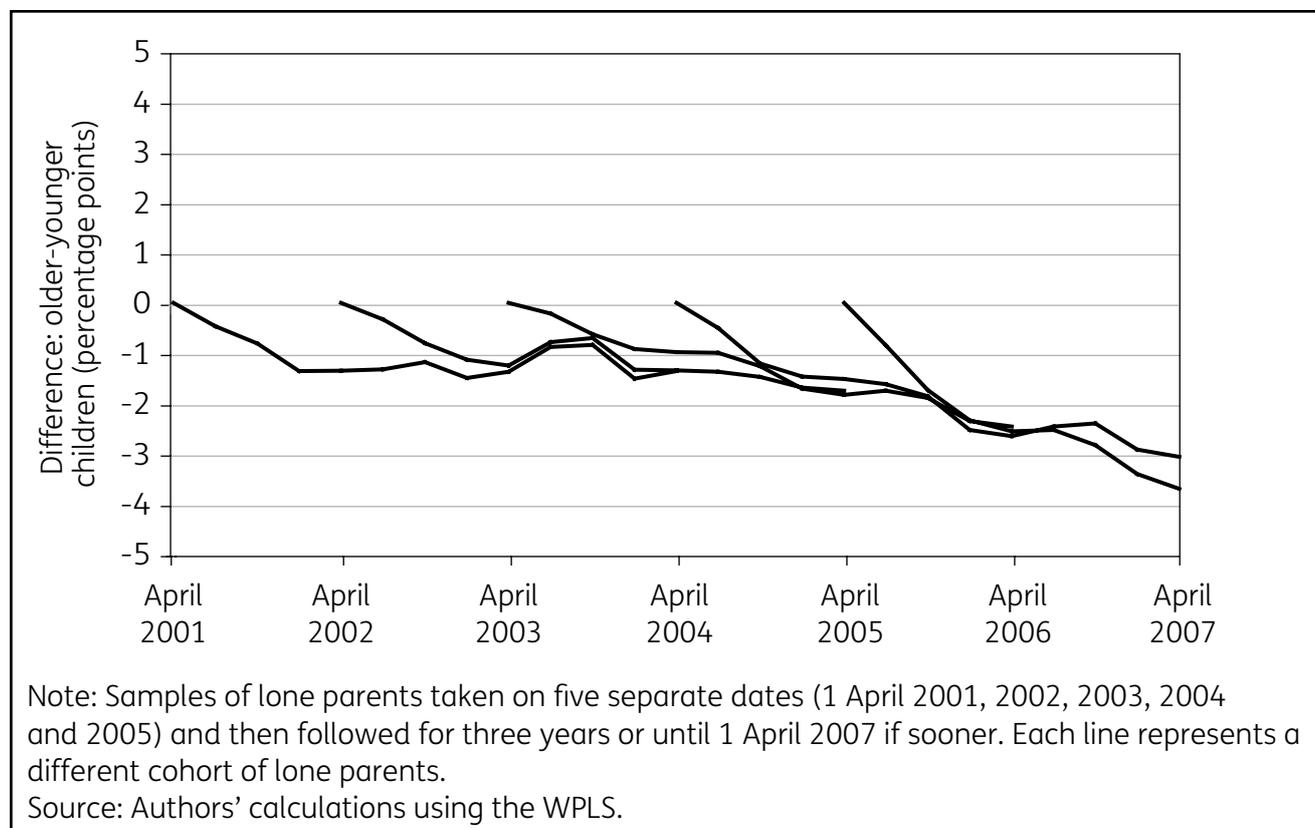


Figure 3.10 Comparing differences in benefit outcomes for lone parents with differently aged youngest children (seven to nine compared to three to four)



Figures 3.11 to 3.13 take the same samples of lone parents who are on IS on 1 April 2001, 2002, 2003, 2004 and 2005 and show how many of them are in work at different points in time over the next three years. The outcome measure is the work measure in the WPLS, which is based on estimated dates when jobs started and ended using information provided by employers to Her Majesty's Revenue & Customs (HMRC) for income tax purposes. There are two key problems with this data:⁹

- it does not necessarily contain information on all jobs, particularly those where the individual earns less than the income tax threshold;
- and there are many instances where the end date of an employment spell is not known.

The result of this is that there are an implausibly large number of lone parents who appear to be on benefit and in work at the same time. Nevertheless, if this error is assumed to remain of a similar magnitude over time (and there is no way of testing whether this assumption is valid), a DiD analysis can still provide a valid estimate of the effect of the LPPC on the number of lone parents in work.

⁹ Many analysts using the work records in the WPLS have made these comments: previous experience of IFS researchers can be found in Brewer, M., Browne, J., Chowdry, H. and Crawford, C. (2010), *The lone parent pilots after 24-36 months: the final impact assessment of In-Work Credit, Work Search Premium, Extended Schools Childcare, Quarterly Work Focused Interviews and New Deal Plus for Lone Parents*, DWP Research Report No. 606.

Figure 3.11 compares lone parents whose youngest child is aged 12 to 15 a year after they are initially sampled with those whose youngest child is aged three or four at the same point in time. Outcomes for lone parents with older children worsened slightly relative to those with younger children. This represents a slight deviation from the common trends assumption.¹⁰

Figure 3.12 shows the same analysis comparing lone parents whose youngest child is aged ten or 11 with those whose youngest child is aged three or four. Work outcomes move roughly in parallel for the two groups, although there is some evidence that outcomes of lone parents with older children worsen relative to those of lone parents with younger children. This represents a slight deviation from the common trends assumption.

Figure 3.13 shows the same analysis comparing lone parents whose youngest child is aged seven to nine with those whose youngest child is aged three or four. The difference in outcomes between lone parents with differently aged children remains broadly the same over time, although there is some evidence of outcomes for lone parents with younger children improving relative to those of lone parents with older children.

Figures 3.14 to 3.16 highlight the differences in outcomes between lone parents with older and younger children. Ignoring the break in the series caused by a change in the methodology used by the WPLS in July 2003, the difference in outcomes between the two groups over the three years moves in a similar way for each cohort, although with a small deterioration in the outcomes of lone parents with older children relative to those of lone parents with younger children.

¹⁰ As mentioned above, there is an implausibly large number of lone parents who are ‘in work’ at the same time as they are initially sampled as being on IS. There is also a discontinuity in the data in July 2003 caused by changes in the way it was collected. This problem appears to be worse for lone parents with older children than those with younger children. However, this is not a violation of the common trends assumption (this only requires that outcomes for the two groups change in the same way over time, not that they are identical in the absence of the LPPC).

Figure 3.11 Comparing trends of work outcomes for lone parents with differently aged youngest children (12 to 15 compared to three to four)

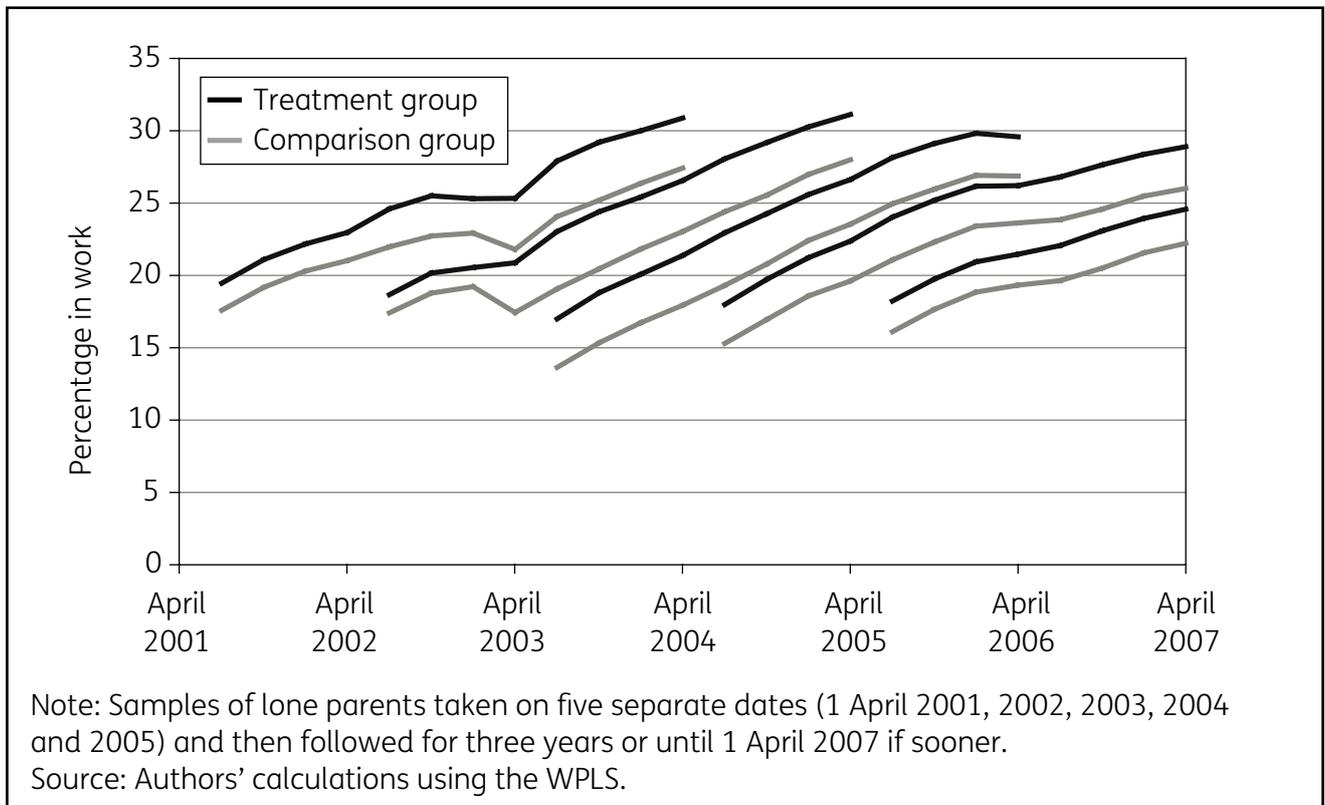


Figure 3.12 Comparing trends of work outcomes for lone parents with differently aged youngest children (10 to 11 compared to three to four)

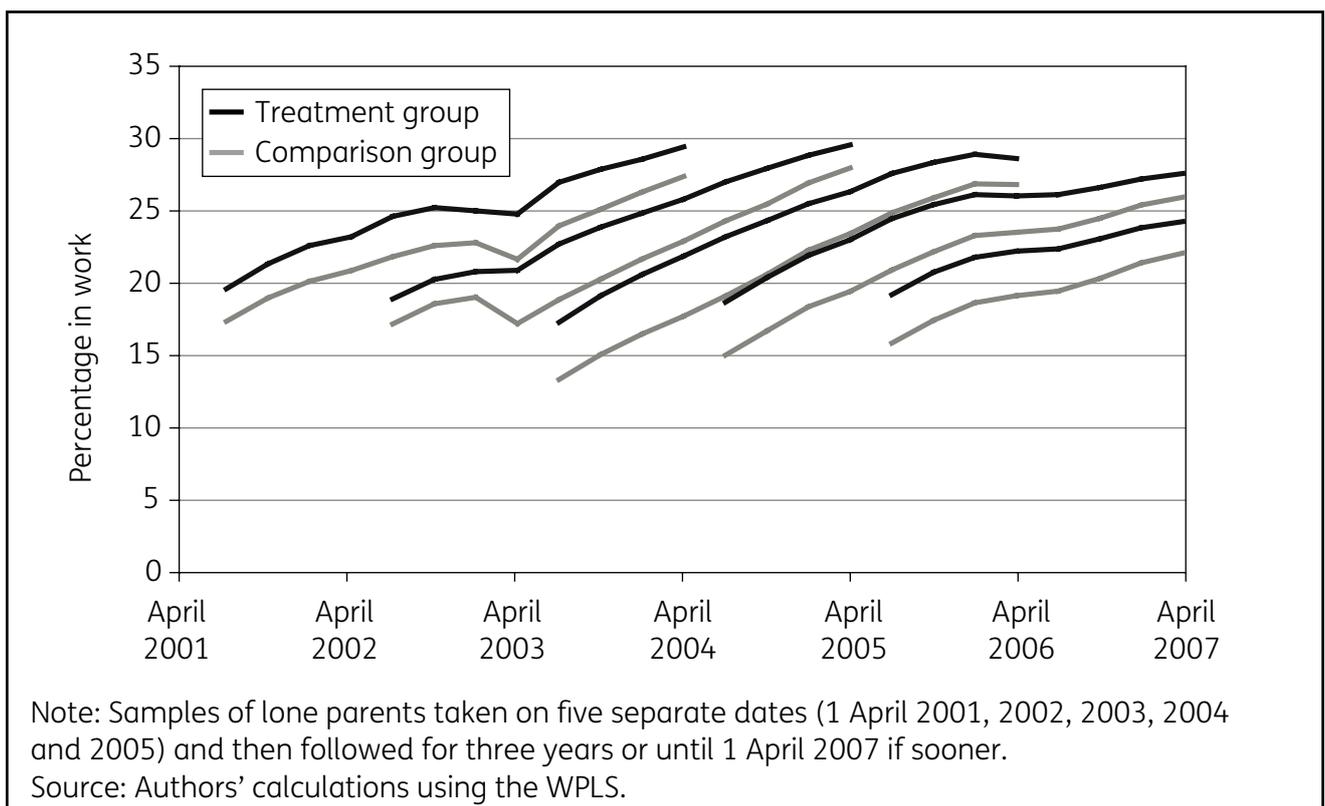


Figure 3.13 Comparing trends of work outcomes for lone parents with differently aged youngest children (seven to nine compared to three to four)

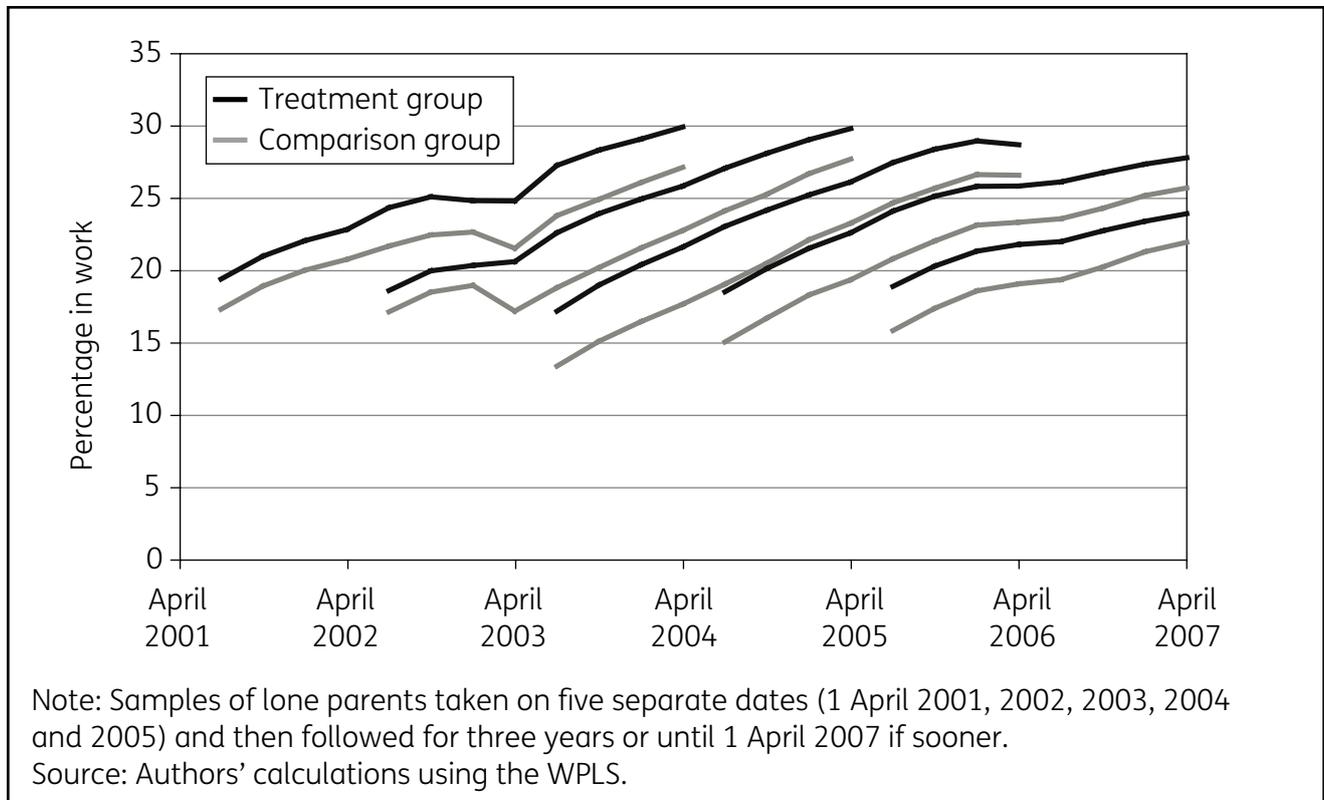


Figure 3.14 Comparing differences in benefit outcomes for lone parents with differently aged youngest children (12 to 15 compared to three to four)

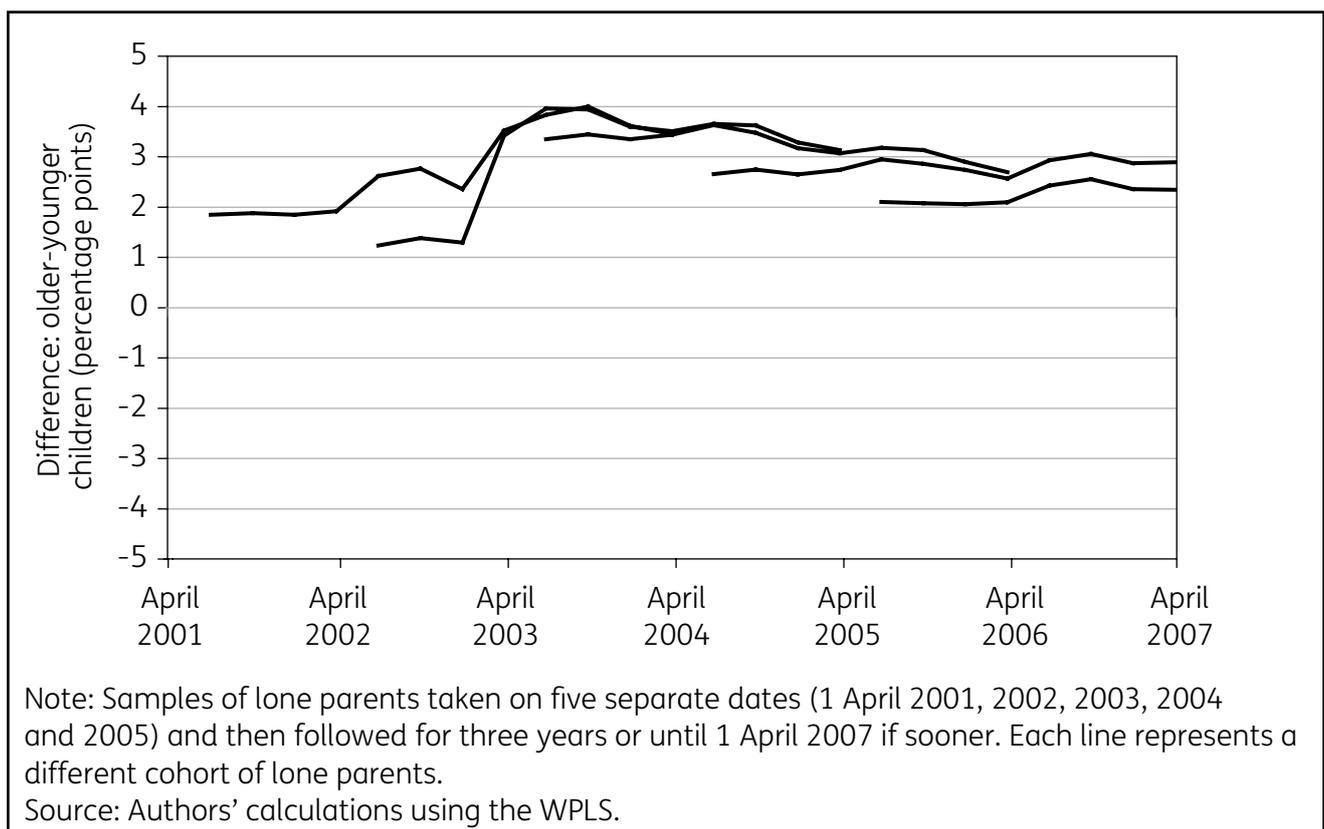


Figure 3.15 Comparing differences in benefit outcomes for lone parents with differently aged youngest children (10 to 11 compared to three to four)

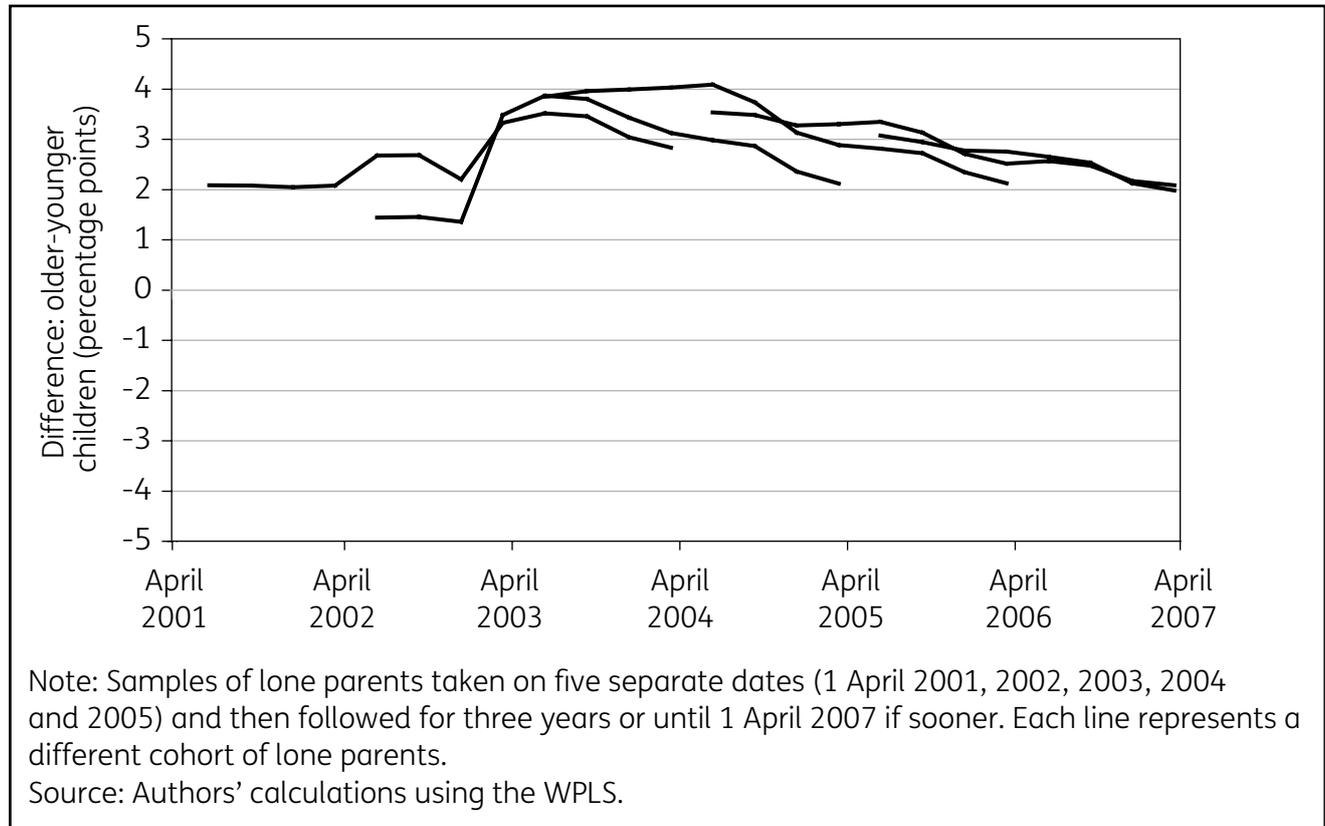
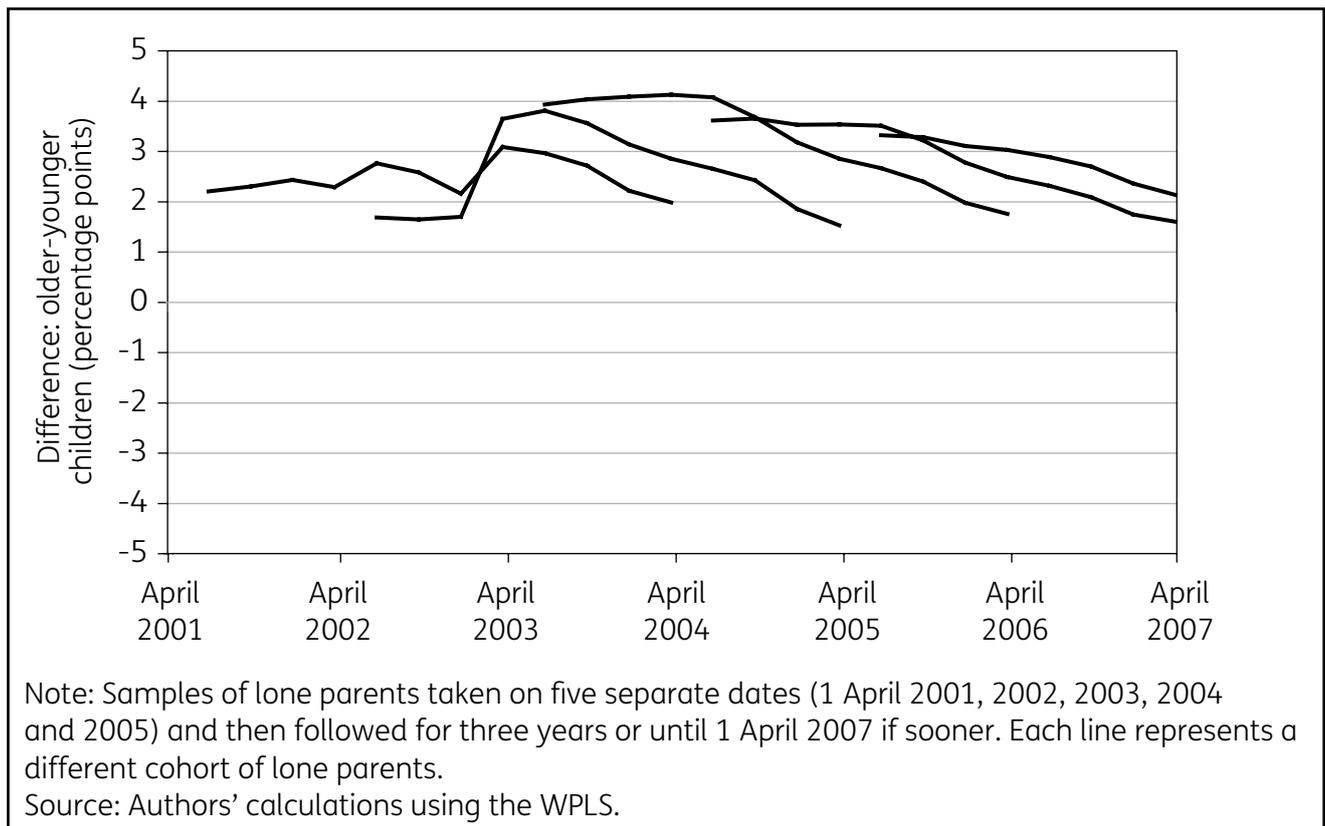


Figure 3.16 Comparing differences in benefit outcomes for lone parents with differently aged youngest children (seven to nine compared to three to four)



This analysis of simple average outcomes shows that lone parents with differently aged youngest children tend to leave benefit at fairly similar rates, on average, and that lone parents with older children are more likely to be in work at all times. In both cases though, outcomes for lone parents with older children deteriorated slightly over time relative to those of lone parents with younger children. This represents a slight deviation from the common trends assumption.

However, it is not clear whether the common trends assumption fails or not from just examining these figures, for two reasons: First, it may be the case that these apparent differences are in fact the result of changes in observable factors among lone parents. Second, any differences may not be statistically significant. To test formally the common trends assumption and control for other factors, regression models need to be estimated. The next section discusses how to conduct these tests correctly.

3.5.2 Correct inference in DiD models

To test whether the common trends assumption is plausible in a regression framework, models of the following form were estimated (on the sample described above):

$$y_{it} = \alpha_t + \gamma_G + \Delta \cdot D_{it} + X_{it}^r \beta + u_{it} \quad (14)$$

where y_{it} is the outcome of interest, α_t is a time effect, γ_G is a group effect (in other words, an age of youngest child effect), Δ is the effect of the hypothetical LPPC, D_{it} is a dummy indicating that individual i is a lone parent whose youngest child is old enough for them to be affected by the hypothetical LPPC and they are observed in a period after the hypothetical LPPC was introduced, and u_{it} is an error term. If the Δ parameter was significantly different from zero, this would imply that the common trends assumption had failed, since this would show that trends in outcomes between lone parents with differently-aged children were different during a period when there was no policy change that affected lone parents differently by age of youngest child.

But an important issue when estimating and interpreting such regression models, though, is how to estimate correctly the standard errors of the impact of the hypothetical policy (in other words, how to test whether the Δ parameter is statistically significantly different from zero). Standard inference using ordinary least squares regression assumes that each observation (i.e. each lone parent) is randomly drawn from the same distribution, and is independent from all other observations. This clearly does not hold in this instance, where lone parents in the sample are from one of four distinct groups – those lone parents with older and younger children, who are observed before or after the hypothetical policy change has taken place. In reality, therefore, there is likely to be a component of the error term that is common to all lone parents whose youngest child is a particular age observed in a particular time period.¹¹

In the case where all members of a particular group are affected by a common error, the model becomes:

$$y_{it} = \alpha_t + \gamma_G + \Delta \cdot D_{it} + X_{it}^r \beta + \nu_{Gt} + \varepsilon_{it} \quad (15)$$

¹¹ This problem was initially studied by Moulton (1986).

where v_{Gt} is a shock that is common to all lone parents whose youngest child is a certain age in a particular time period, and ε_{it} is an idiosyncratic shock that is particular to each lone parent.¹² Having a large sample of lone parents in each group (where group is defined by age of youngest child) and each time period will reduce the standard errors on the coefficients of interest caused by the latter error, but will do nothing to eliminate the shocks common to a particular group and time period. Assuming that the error term is idiosyncratic (i.e. not accounting for the fact that a part of the error term is likely to be common to all lone parents in a particular group and time period) will, therefore, underestimate the size of the standard errors of the parameters of the model, and lead to incorrect inferences being drawn.

One way of accounting for the fact that the error term in the model is of the form specified in (15) is to ‘cluster’ the standard errors to account for the fact that error terms are likely to be correlated between lone parents in the same group and time period. The standard errors that are produced by this procedure will provide valid inference so long as the component of the error term that is common to all members of a group in a particular time period (represented by v_{Gt} in equation (15)) are uncorrelated between groups and time periods. Unfortunately, this is unlikely to be the case. The error terms for a particular group are also likely to be serially correlated – that is to say that if lone parents whose youngest child is a particular age are affected by a common shock in period t , the common shock to all lone parents whose youngest child is the same age in period $t+1$ is likely to be similar.¹³ Therefore, just as the standard errors from the model need to be adjusted as a result of the correlation of error terms within groups at a particular time period, they also need adjusting to account for serial correlation in these error terms themselves. The simplest way of dealing with this problem is to ‘cluster’ standard errors at the group, rather than the group and time period level.¹⁴ This allows for the error term to be correlated within all lone parents whose youngest child is a particular age over time, including the serial correlation between lone parents observed at different times.

Bertrand, Duflo and Mullainathan (2004) show that, while this technique yields standard errors that give correct inference when the number of clusters is large, the standard errors are still likely to be biased downwards if there are only a small number of clusters (as is the case here), meaning that the common trends assumption will be over-rejected. This is because with few clusters either the serial correlation or the correlation between individuals in a particular cluster will be underestimated. There are ways in which this bias can be corrected which could be investigated further and implemented by an eventual IA.¹⁵ These methods are at the cutting-edge of modern

¹² It is also possible that the shocks are common to parents with children born in a particular time period. We find this a little less plausible. In any case, such a situation would be easier to deal with than a situation when shocks are common to parents with children of a given age, because the actual policy impacts should vary with the age (not date of birth) of the children. Cameron *et al.* (2010) offers a suggestion on what to do if data is subject to clustering in more than one dimension, but these methods may not be appropriate in a case like ours where the groups sizes are small.

¹³ Also, in a panel model where the same lone parents are tracked over time, it is possible that the error terms of particular cohorts of lone parents (where cohorts are defined as all lone parents whose youngest child was born in the same month) will have a common component, and that this common error component will be serially correlated over time. This could be similarly dealt with by ‘clustering’ standard errors at the cohort level.

¹⁴ This method was first suggested by Arellano, M. (1987).

¹⁵ For example, Bell and McCaffrey (2002) present a method of adjusting standard errors obtained from clustering by a method known as bias-reduced linearisation. Cameron *et al.* (2008) show that several bootstrap methods work well with small numbers of clusters.

econometrics, and the literature has not yet come to a consensus on the best way in which to correct for such bias; any eventual impact assessment should examine the ways in which standard errors could be adjusted to take account of this bias.

3.5.3 Assessing the suitability of DiD with lone parents with younger children as a comparison group: analysis using linear regression

This section reports the results of implementing the pre-programme tests with linear regression methods, and where standard errors were clustered at the level of the group (see Section 3.5.2).

The analysis in Section 3.5.1 was based on sampling lone parents with differently aged youngest children, and tracking their outcomes for three years, having split the lone parents whose children were old enough for them to be affected by the hypothetical LPPC, into three groups (those whose youngest child was aged 12 to 15, ten to 11 and seven to nine). In this section, separate regressions are run for each phase of the LPPC.

The precise procedure for selecting the sample for each regression was as follows:

- lone parents were sampled who are on IS on each of 1 April 2001, 2002, 2003 or 2004;¹⁶
- for each phase, lone parents were selected whose youngest child is the same age (in months) as those lone parents who will be affected by that phase of the LPPC, as were those lone parents whose youngest child is the same age as those lone parents in the suggested comparison group for that phase. For example, to form a sample for a regression for the Phase 1a stock, lone parents were sampled whose youngest child is aged between 14 years two months and 15 years 11 months on the day the hypothetical LPPC is introduced, as are those whose youngest child is aged between two years two months and three years 11 months (the suggested comparison group for eventual impact assessment of this phase) on the same day;
- the outcomes of interest are whether they are on an out-of-work benefit (that is to say IS, JSA or IB) three years after initially sampled on IS, whether they are in work (according to the measure in the WPLS) on the same day, what proportion of the three years after they are initially sampled on IS they spend on an out-of-work benefit (defined as above), and what proportion of this time they spend in work (according to the measure in the WPLS).

The regression estimated were as follows:

$$y_{it} = \sum_t \alpha^t + \gamma_G + \sum_t \Delta^t \cdot D_{it} + X_{it}^t \beta + v_{Gt} + \varepsilon_{it} \quad (16)$$

This is similar to equation (15) except that it is amended to allow for more than two time periods. There are now multiple time effects and multiple group-time effects; these allow the common trends assumption to be tested by comparing outcomes of the 2001 cohort with those who are initially sampled in each of 2002, 2003 and 2004. Essentially, the 2001 cohort is treated as the group unaffected by any hypothetical LPPCs, and the subsequent cohorts are treated as affected by a hypothetical LPPC.

The tables that follow report the time effects (represented by dummy variables for each cohort, α^t), group effects (represented by a dummy variable for those lone parents who are affected by the hypothetical LPPC, γ_G) and group-time effects. Each regression corresponds to a particular phase of the LPPC. The coefficients on the other demographic variables are not reported (represented by the vector β), but each regression included local area variables from different sources, information on

¹⁶ We dropped the 1 April 2005 sample as we do not observe them for a full three years in our data (which stops on 31 March 2007).

individual lone parents reported in the WPLS, and various policy variables.

The local area variables included are:

- local employment rates from the 2001 census;
- index of multiple deprivation (IMD) scores, childcare availability;
- the qualifications of lone parents who are unemployed from the 2001 census; and
- tenure variables for lone parents from the 2001 census.

The variables from the WPLS are:

- number of children;
- sex;
- age of youngest child (linearly);
- initial benefit amount;
- history variables – amount of time in the 18 months prior to initial sampling spent in work, and amount of time in the 18 months prior to initial sampling spent on benefit;
- whether a lone parent is entitled to a disability premium at the point they are sampled or at some point in the past;
- whether a lone parent is simultaneously claiming another benefit at the same time as IS (IB, Disability Living Allowance (DLA), Attendance Allowance (AA), Carer's Allowance, Bereavement Benefit or Widows Benefit); and
- whether a lone parent has participated in the New Deal for Lone Parents.

Dummy variables to take into account whether lone parents are affected by the following policies were also included:

- WFIs;¹⁷
- Quarterly WFIs (QWFIs) for all lone parents whose youngest child is aged 14 or over;
- Extended Schools Quarterly WFIs for lone parents whose youngest child is aged 12 or over in areas where Extended Schools Childcare is operating;
- the New Deal Plus for Lone Parents; and
- In-Work Credit (IWC).

As discussed in Section 3.5.2, it should be noted that the researchers believe the standard errors being used in this analysis are probably downward-biased (i.e. too small) because standard errors were clustered by group, and there are only a small number of groups. This bias may lead one to judge incorrectly that coefficients are statistically significantly different from zero, and thus incorrectly rejecting the common trends assumption.

¹⁷ These were rolled out by age of youngest child between April 2001 and April 2004, during the period we study here. There is a danger that any differential trend between lone parents with differently aged youngest children is in fact picking up the effect of WFIs on lone parents with younger children. Assuming that the effect of WFIs is the same over time, allowing for more than two time periods should enable us to separately identify the effect of WFIs from the common trend.

Tables 3.2 to 3.4 show that, for the outcome of being off benefit three years after initial sampling (two years after the hypothetical LPPC was introduced), some of the parameters that test whether the common trends assumption fails are statistically significant at the one per cent level, and some of these deviations from the common trends assumption are large, particularly when comparing lone parents whose youngest children are aged 12 or over with the chosen comparison group.¹⁸ There is, therefore, substantial evidence that the common trends assumption did not hold during this period, as the divergences from this assumption during the pre-programme period were both large and statistically significantly different from zero. However, two caveats to this finding are needed: First, most of the parameters are small in magnitude (around one percentage point), meaning that estimates of the effect of the LPPC would be subject to this degree of inaccuracy had it been introduced during this period, and are not significantly different from zero at the one per cent level. Second, as discussed in Section 3.5.2, the standard errors will be biased downwards due to the small number of clusters, meaning that the common trends assumption could be being incorrectly rejected.

Looking at the other parameters in the model, there is some evidence that outcomes improved over the period in question, as the α parameters are large and positive, and that outcomes were better for lone parents with older children, as γ_G is large and positive in most models.

For the other benefit outcome - the proportion of time lone parents spend on benefit in the three years after they are initially sampled on IS - Tables 3.5 to 3.7 show that there are fewer instances of the common trends assumption being rejected at the one per cent significance level. The deviations from common trends also tend to be smaller (at typically less than one percentage point). Again, outcomes for both groups tended to improve over time, and lone parents with older children tended to have better outcomes.

¹⁸ The one per cent significance criterion is, arguably, the correct one to use in this instance (rather than the conventional five per cent criterion): the sample sizes are so large that the power of the statistical tests being carried out is large (that is to say, the chance of not rejecting the common trends assumption when it does not in fact hold is small) even when a higher significance level is used. In choosing which significance level to use, we have to make a trade-off between making Type I and Type II errors. A Type I error occurs when we reject a hypothesis that is true. In this instance, this would mean rejecting the common trends assumption when this is a reasonable assumption. This is less likely to occur when we use a low significance criterion (i.e. one per cent rather than five per cent). A Type II error occurs when we fail to reject a hypothesis that is false. In this instance, this would mean not rejecting the common trends assumption when it does not hold. This is more likely to occur when we use a low significance criterion. Increasing sample size reduces the probability of making a Type II error, but does not change the probability of making a Type I error. It is, therefore, sensible when working with a large sample to reduce the significance criterion, as this allows us to use the increased sample size to lower the probability of both types of error.

Table 3.2 Outcome: off benefit three years after initially observed on IS, Phase 1

	Phase 1a stock	Phase 1a flow	Phase 1b stock	Phase 1b flow
γ_G	7.8	-82.2**	2.8	17.2**
Δ^{2002}	1.0*	1.2**	1.1**	-0.1
Δ^{2003}	-0.3	3.1***	1.0***	1.3*
Δ^{2004}	-0.4***	2.1***	0	-1.2*
α^{2002}	0.9	1.7*	1.2**	2.1**
α^{2003}	1.5	1.8	2.1**	2.1*
α^{2004}	2.5	3.0*	3.2**	3.5*
Observations	988,581	335,461	868,803	215,366

Notes:

* significant at the ten per cent level.

** at the five per cent level.

*** at the one per cent level. Parameters as defined in equation (16).

Table 3.3 Outcome: off benefit three years after initially observed on IS, Phase 2

	Phase 2a stock	Phase 2a flow	Phase 2b stock	Phase 2b flow
γ_G	3.2	11.8***	-8.1**	-5.7
Δ^{2002}	0.7**	0.7***	1.1**	1.2***
Δ^{2003}	0.3*	0.8***	0.4**	1.0***
Δ^{2004}	-1.2*	-1.3**	-1.0**	-0.2
α^{2002}	1.4	1.9**	1.6*	2.3*
α^{2003}	2.5	2.2*	2.5*	2.3*
α^{2004}	3.6	3.6*	3.8*	3.8
Observations	43,796	345,715	375,190	177,808

Notes:

* significant at the ten per cent level.

** at the five per cent level.

*** at the one per cent level. Parameters as defined in equation (16).

Table 3.4 Outcome: off benefit three years after initially observed on IS, Phase 3

	Phase 3a stock	Phase 3a flow	Phase 3b stock	Phase 3b flow
γ_G	1.1	4.4**	4.5**	7.9**
Δ^{2002}	0.9***	-0.6***	1.9**	1.3***
Δ^{2003}	-0.6**	-0.2**	1.8***	1.3**
Δ^{2004}	-1.4**	-1.4**	0.3*	-0.1
α^{2002}	1.4	1.8*	1.3*	1.8*
α^{2003}	2.6	2.2*	2.4*	2.2*
α^{2004}	3.8	3.7	3.5*	3.6*
Observations	573,446	564,741	710,385	608,520

Notes:

* significant at the ten per cent level.

** at the five per cent level.

*** at the one per cent level. Parameters as defined in equation (16).

Table 3.5 Outcome: proportion of three years spent off benefit, Phase 1

	Phase 1a stock	Phase 1a flow	Phase 1b stock	Phase 1b flow
γ_G	-23.4**	-1	4.1*	9.7**
Δ^{2002}	0.8***	0.6**	1.0*	0.4*
Δ^{2003}	0.1	0.5	0.9***	0.7**
Δ^{2004}	-0.3	0	0	-0.7*
α^{2002}	-0.3**	0.2***	-0.1	0.3
α^{2003}	0.1**	0.4**	0.6	0.7
α^{2004}	1.5**	1.6***	1.8	2.1
Observations	988,581	335,461	868,803	215,366

Notes:

* significant at the ten per cent level.

** at the five per cent level.

*** at the one per cent level. Parameters as defined in equation (16).

Table 3.6 Outcome: proportion of three years spent off benefit, Phase 2

	Phase 2a stock	Phase 2a flow	Phase 2b stock	Phase 2b flow
γ_G	1.1	7.7**	-2.6*	-0.3
Δ^{2002}	0.8**	0.9**	1.0**	0.9**
Δ^{2003}	0.6**	1.0**	0.7**	0.7**
Δ^{2004}	-0.6**	-0.1*	-0.1*	0
α^{2002}	0	0.3	0.1	0.5
α^{2003}	0.8	0.7	0.7	0.9
α^{2004}	2.1	1.9	2	2.2
Observations	643,796	345,715	375,190	177,808

Notes:

* significant at the ten per cent level.

** at the five per cent level.

*** at the one per cent level. Parameters as defined in equation (16).

Table 3.7 Outcome: proportion of three years spent off benefit, Phase 3

	Phase 3a stock	Phase 3a flow	Phase 3b stock	Phase 3b flow
γ_G	3.1**	4.2**	3.3**	6.6***
Δ^{2002}	0.7**	-0.3*	1.1**	0.8**
Δ^{2003}	-0.2	0.0	1.5**	1.0***
Δ^{2004}	-0.6**	-0.6**	0.2*	0.0
α^{2002}	0	0.2	-0.1	0.2
α^{2003}	0.9	0.7	0.6	0.6
α^{2004}	2.1	2	1.9	1.8
Observations	573,446	564,741	710,385	608,520

Notes:

* significant at the ten per cent level.

** at the five per cent level.

*** at the one per cent level. Parameters as defined in equation (16).

Tables 3.8 to 3.10 look at whether a lone parent is in work three years after they are initially observed on IS. They show that work outcomes tend to be better for lone parents with older children, and that outcomes in later periods tend to be better (at least when comparing the group that would be affected by the third phase of the LPO roll-out). Most deviations from the common trends assumption tend to be less than one percentage point in magnitude, although some are larger (but all are less than two percentage points). The common trends assumption is again rejected at the one per cent level in a few instances. Generally, outcomes for lone parents with older children tended to improve relative to those with younger children between 2001 and 2003, but then fell back relative to those of lone parents with older children in 2004.

Finally, Tables 3.11 to 3.13 look at the proportion of time spent in work during the three years after lone parents are initially sampled on IS. Deviations from common trends are relatively large at between one and two percentage points (although generally lower for Phase 3, where the treatment and comparison groups are closer in terms of the age of their youngest child) and the common

trends assumption is rejected at the one per cent level in a few instances. Outcomes were better for lone parents with older children during the period in question, and that outcomes for lone parents with older children tended to improve relative to those of lone parents with younger children between 2001 and 2003, but then fell in 2004.¹⁹

Table 3.8 Outcome: in work three years after initially observed on IS, Phase 1

	Phase 1a stock	Phase 1a flow	Phase 1b stock	Phase 1b flow
γ_G	8.0	-14.2*	6.5*	10.3***
Δ^{2002}	1.3**	0.8**	0.7	0.0
Δ^{2003}	-0.3	0.4	1.6**	0.7*
Δ^{2004}	-1.7*	-1.2*	0.0	-1.6***
α^{2002}	-0.2	0.3	-0.2	0.4*
α^{2003}	0.0	-0.7	0.1*	-0.5**
α^{2004}	0.6	-0.2	0.7	0.0
Observations	988,581	335,461	868,803	215,366

Notes:

* significant at the ten per cent level.

** at the five per cent level.

*** at the one per cent level. Parameters as defined in equation (16).

Table 3.9 Outcome: in work three years after initially observed on IS, Phase 2

	Phase 2a stock	Phase 2a flow	Phase 2b stock	Phase 2b flow
γ_G	6.7*	5.8**	-2.8	14.4
Δ^{2002}	0.2	0.9**	0.2	0.4
Δ^{2003}	-0.2	1.0*	0.4	0.3
Δ^{2004}	-1.9***	-0.8*	-1.1*	-0.7**
α^{2002}	0.3	0.6**	0.4	0.2
α^{2003}	0.1	-0.2	-0.3	-0.4
α^{2004}	0.9	0.5	0.6	-0.1
Observations	643,796	345,715	375,190	177,808

Notes:

* significant at the ten per cent level.

** at the five per cent level.

*** at the one per cent level. Parameters as defined in equation (16).

¹⁹ These dates all refer to the years in which the lone parents are initially observed on IS.

Table 3.10 Outcome: in work three years after initially observed on IS, Phase 3

	Phase 3a stock	Phase 3a flow	Phase 3b stock	Phase 3b flow
γ_G	4.5**	4.6**	4.4***	8.0***
Δ^{2002}	0.1	-0.3	0.9*	0.7*
Δ^{2003}	-0.2	0.2	1.3*	1.5*
Δ^{2004}	-1.7**	-1.1*	-0.4**	-0.1
α^{2002}	0.3*	0.4	0.2**	0.4**
α^{2003}	0	-0.4*	0	-0.4**
α^{2004}	0.7	0.3	0.6	0.3
Observations	573,446	564,741	710,385	608,520

Notes:

* significant at the ten per cent level.

** at the five per cent level.

*** at the one per cent level. Parameters as defined in equation (16).

Table 3.11 Outcome: proportion of three years spent in work, Phase 1

	Phase 1a stock	Phase 1a flow	Phase 1b stock	Phase 1b flow
γ_G	-0.2	3.8	5.1*	5.6***
Δ^{2002}	1.6**	0.8*	1.4*	0.6*
Δ^{2003}	1.1	1.2	2.3**	1.9**
Δ^{2004}	-0.8	-1.1*	0.0	-0.8***
α^{2002}	-1.7**	-0.8*	-1.5**	-0.8***
α^{2003}	-1.2***	-1.3*	-1.0*	-1.3**
α^{2004}	0.9**	1.0*	1.2	1.0
Observations	988,581	335,461	868,803	215,366

Notes:

* significant at the ten per cent level.

** at the five per cent level.

*** at the one per cent level. Parameters as defined in equation (16).

Table 3.12 Outcome: proportion of three years spent in work, Phase 2

	Phase 2a stock	Phase 2a flow	Phase 2b stock	Phase 2b flow
γ_G	3.4	-2.6*	0.9	10.1
Δ^{2002}	1.0*	0.9**	1.1**	0.6
Δ^{2003}	1.4*	1.9**	1.8**	1.3*
Δ^{2004}	-1.0***	-0.6*	-0.3	-0.7***
α^{2002}	-1.2**	-0.7*	-1.1**	-0.7*
α^{2003}	-1	-1.1	-1.4*	-1.0***
α^{2004}	1.5	1.4	1.2	1.1
Observations	643,796	345,715	375,190	177,808

Notes:

* significant at the ten per cent level.

** at the five per cent level.

*** at the one per cent level. Parameters as defined in equation (16).

Table 3.13 Outcome: proportion of three years spent in work, Phase 3

	Phase 3a stock	Phase 3a flow	Phase 3b stock	Phase 3b flow
γ_G	5.0**	4.5**	4.0***	7.1***
Δ^{2002}	0.9*	0.4	0.9*	0.6*
Δ^{2003}	1.2*	1.3*	2.0**	1.9**
Δ^{2004}	-0.8**	-0.7*	-0.2**	-0.2*
α^{2002}	-1.2***	-0.9**	-1.2**	-0.8***
α^{2003}	-1.2*	-1.2***	-1.1**	-1.2**
α^{2004}	1.3	1.2	1.2*	1.2
Observations	573,446	564,741	710,385	608,520

Notes:

* significant at the ten per cent level.

** at the five per cent level.

*** at the one per cent level. Parameters as defined in equation (16).

3.5.4 Overall assessment of the suitability of the DiD method, using lone parents with younger children as a comparison group

Section 3.5.3 shows that there is evidence that the common trends assumption did not hold during the period from 2001 to 2007. However, it is the assessment of the researchers that an impact assessment using the DiD methodology could be informative in telling us about the effect of the LPPC on those directly affected, for the following reasons.

First, there are some important points to note about the pre-programme tests in Section 3.5.3:

- the standard errors used will probably be biased downwards (i.e. too small) due to the small number of clusters in the model, meaning that the common trends assumption could be being incorrectly rejected;

- even with this likely bias, most of the parameters tested are not significantly different from zero at the one per cent level;
- where there are statistically significant divergences, most are small in magnitude, at less than one percentage point.

Second, an IA that used the DiD methodology could be amended to reflect the evidence of a pre-existing deteriorating trend in outcomes among lone parents with older children by using a trend-adjusted DiD method, as discussed in Section 3.3.1 (this would be equivalent to exploring whether the estimated impacts are robust to changing the model specification in order to take account of the apparent deteriorating trend in outcomes among lone parents with older children). When doing this, it would help if as long a period of pre-programme data as possible is used.

Ultimately, though, the suitability of this approach depends on the size of impact that an eventual IA is expected to find (or the questions which policy-makers most want answering): if the impact of the LPPC is expected to be smaller than the two percentage points by which the trends of the treatment and comparison groups diverge during the pre-programme period, then it seems unlikely that such impacts could be detected reliably. But if an eventual IA is intended to test whether the LPPC had impacts as large as five or ten percentage points (compared with no effect), then the DiD method (and the trend-adjusted DiD method) could provide robust conclusions.

Finally, it should be noted that even if the common trends assumption was considered reasonable during the period from 2001 to 2007, this is no guarantee that it would have continued to hold after the LPPC was introduced in the absence of any policy change. This is of particular concern in this instance as the period considered ended some time before the start of the recession into which the LPPC was introduced.

3.6 Summary

This chapter has argued that:

- the recommended empirical strategy is to use a DiD approach. This will compare the outcomes of lone parents directly affected by the LPPC with those of lone parents whose children are sufficiently young so that they are not affected; this comparison will then itself be compared to the difference in outcomes of similar groups of lone parents observed before the LPPCs. The idea is that differences in outcomes of lone parents with differently aged youngest children before the LPPC begins are informative about what the differences in outcomes of lone parents with differently aged youngest children would have been in the absence of the LPPC;
- there is a trade-off between examining outcomes for a longer period of time, or allowing for longer anticipation effects, and reducing the difference between the treatment and comparison groups in terms of the age of the youngest child. For some of the early phases of the roll-out, the difference in ages of the treatment and comparison groups will be large;

- the key assumption needed for a DiD estimate to be valid is known as the common trends assumption. In the case of the LPPC, common trends would fail if the labour market behaviour of lone parents with younger children changed in a way which was different from the change in the labour market behaviour of lone parents with older children. There is some evidence that the common trends assumption did not hold during the period from 2001 to 2007. But the divergences are usually small, and may not be statistically significant, and they could be accounted for by using a trend-adjusted DiD model. However, even if this assumption holds in the pre-programme period, there is no guarantee that it will still hold after the LPPC is introduced. A particular concern for this IA is that the LPPC is being implemented during a recession, and there is no pre-programme data to test whether common trends held during previous recessions. Over and above this, though, the main threats to the DiD approach are anticipation effects, and the impact of other contemporaneous policy changes, and these are discussed in Chapter 6;
- naïve analysis using the DiD method often produces estimated standard errors which are too small, and the conventional ways of addressing these (with clustered standard errors) may not be appropriate in the particular case of the LPPC. An eventual impact assessment should examine the ways in which standard errors can be estimated correctly in this situation, but it should be noted that these methods are at the cutting-edge of modern econometrics and the literature has not yet come to a consensus;
- the overall assessment of the research team is that, if the actual impact of LPPC is expected to be smaller than two percentage points, then it seems unlikely that such impacts could be detected reliably. But if an eventual IA is intended to test whether the LPPC had impacts as large as five to ten percentage points (compared with no effect), then a DiD or a trend-adjusted DiD model should provide robust answers.

4 The empirical approach: further refinements and detailed considerations

This chapter discusses some elaborations to the basic difference-in-difference (DiD) approach: Section 4.1 discusses what more can be gained by using a duration model, Section 4.2 discusses the advantages and disadvantages of evaluating the impact of the roll-out compared with the steady-state policy and Section 4.3 discusses why the impact assessment (IA) will not be able to estimate the separate impact of the different elements of the lone parent policy change (LPPC) (although it will be able to provide a profile of the impacts over time).

4.1 Dynamic selection and duration models

The research team believe that a version of the DiD methodology is an appropriate method for estimating the overall effect of the LPPC on one population of interest (lone parents who are receiving Income Support (IS) as the treatment starts, as discussed above). A future IA may, however, wish to estimate particular impacts that, by their nature, introduce further selection problems. Specifically, there may be an interest in examining outcomes for sub-groups of lone parents, where these sub-groups are defined by the choices that they make: the most obvious would be employment-contingent outcomes for those who enter work.

For example, suppose a future IA aims to estimate the effect of the LPPC on the duration of subsequent employment spells amongst lone parents previously on IS. Further suppose that one could find a comparison group of lone parents on IS that is, on average, just like the treatment group (this is what a randomised trial would deliver). This would allow us to estimate the effect of the LPPC on the transition into work by a straightforward comparison of means. However, this would not allow us to estimate the effect of the LPPC on subsequent employment durations. The reason is as follows: if the original treatment group and comparison group are, on average, the same but the LPPC leads to a larger fraction of the treatment group gaining employment, then the employed treatment group members will not necessarily be the same, on average, as the employed comparison group members.

Although some of these differences can be reflected in the data on individuals (i.e., it is observable), some of it may not (i.e. it is unobservable). For example, suppose that the key unobserved characteristic is intrinsic motivation: the most motivated find employment. Intrinsic motivation varies across individuals, but is on average the same in the treatment and comparison group (that is what makes it a good comparison group). Suppose that the $X\%$ most motivated lone parents in the comparison group find work, and the $Y\%$ most motivated individuals in the treatment group find work; because the treatment raises motivation and job-finding probabilities, Y should be greater than X . The average intrinsic motivation of employed members of the treatment group will then not (in general) be equal to that of the employed members of the comparison group: if the treatment provides additional motivation to find work, then individuals in the treatment group with lower intrinsic motivation will find work, and this in turn means that the average intrinsic motivation among employed members of the treatment group will be lower than the average intrinsic motivation of employment members of the comparison group. A simple comparison of employment durations between the treatment group and comparison group would confound the true effect

of the policy on employment durations with these differences in the intrinsic motivation of the employed members of the two groups (and usually lead to an estimate of the treatment effect which is too small).

This is an example of a dynamic selection problem (Ham and Lalonde, 1996). Initial randomisation – or the initial suitability of the comparison group – breaks down if researchers wish to study dynamic outcomes, such as the duration of subsequent spells of employment and non-employment. The DiD estimator, and elaborations of that estimator, typically cannot overcome such dynamic selection problems. Instead, researchers have to model explicitly the way that outcomes depend both on treatment status and unobserved characteristics of treated and untreated individuals. This can be done in a duration (or survivor) model. This will typically involve making stronger assumptions about functional form and the nature of the unobserved heterogeneity than in a conventional DiD estimator, but these stronger assumptions will allow us to make richer statements about the impact of the programme. For example, rather than just determining the overall effect on employment rates, researchers will be able to parse such effects into impacts on job finding rates and on job durations. It will also allow us to separate people affected by the programme into the marginals – those people who are induced to change behaviour – and non-marginals – those whose behaviour was unaffected. Having data with multiple spells for individuals allows for slightly less strong assumptions about the nature of the unobserved heterogeneity.

There is a wide literature using duration models to examine the effect of additional obligations on the length of benefit claims (an early example for the UK is Dolton and O'Neill (1996)). It is also possible to control for dynamic selection bias in order to examine subsequent work outcomes, as Ham and Lalonde (1996) show, looking at work outcomes of women after they participate in training schemes. Other examples include Card and Hyslop (2005), who use a duration model to estimate the impact of a time-limited work subsidy for lone parents on welfare in Canada (similar to In-Work Credit (IWC)), and Bewley *et al.* (2008), which was part of the impact assessment of Pathways to Work. Researchers at the Institute for Fiscal Studies (IFS) have recently been working on such a model, results from which form part of the draft report on the LPP pilots²⁰.

Typically (but not always), duration models examine transitions between mutually exclusive states. In the case of the LPPC, the states of interest could be:

- receiving an out-of-work benefit;
- not receiving an out-of-work benefit and in work for at least 16 hours per week; and
- not receiving an out-of-work benefit and working for fewer than 16 hours per week (including not working at all).

A duration model is based on data on transitions from one state to another: in this model, there would be three states and hence six possible transitions. A more complicated model would expand the out-of-work benefit state to reflect whether lone parents were claiming IS, Jobseeker's Allowance (JSA) or Incapacity Benefit (IB), increasing the number of states to five and the number of transitions that needed to be modelled to 16. The likelihood of lone parents making these transitions would be modelled as a function of observable and unobservable characteristics, as well as being affected by the LPPC.

²⁰ Brewer *et al* (2009).

A duration model can control for dynamic selection bias by explicitly modelling the unobserved differences between lone parents that lead to the dynamic selection bias. The form that these unobserved differences take would have to be specified in a parametric duration model, but the fact that they would be correlated for a given lone parent over time, including when they were in different states, would mean that – given the assumptions of the model – the impact of dynamic selection bias could be separated from the genuine effect of the LPPC.

4.2 Estimating the impact of the roll-out of the lone parent policy change as well as the steady-state

In principle, it is feasible to estimate the impact of the roll-out policy using the methods described in Section 3.3. As that explained, the fact that there are many different groups of lone parents (ie the different phases of the roll-out) affected by policy changes at different times means that a simple DiD estimator (which allows for two groups – those affected and those unaffected by a policy change – and two time periods – before and after a policy change) would be inappropriate. But it is easy to modify the DiD estimator to incorporate several dates on which policy changes, and several groups.

There are certain features that make it more attractive to estimate the impact of the roll-out than the steady-state:

- Because the roll-out features policy changes for different groups of lone parents (or the different phases) on different dates, estimates of the impact of the roll-out based on the DiD method should be more robust to failures of the common trends assumption due to group-specific shocks. For example, if, hypothetically, there was a single implementation date for the LPPC, and lone parents with younger children started to have better outcomes just as the LPPC was implemented, and those with older children did not – due to some change in the economy – then a DiD estimator would give an underestimate of the impact of the LPPC. In reality, there is more than one implementation date, so even if the common trends assumption was violated in this way, the DiD estimate would underestimate the impact on one phase of the roll-out, but there would not necessarily be a problem with the impact on other phases.
- Anticipation effects – over those already allowed for by the DiD estimator – should be less of a problem during the roll-out than the steady-state, because the policy has been in existence for less time (which places an absolute limit on anticipation effects), and because any peer effects will be relatively small. In the steady-state though, lone parents may know about the LPPC a long time before they are themselves affected by it, giving a greater scope for anticipation effects.
- An IA of the steady-state has the problem that lone parents may be affected by the LPPC more than once, particularly if they have another child after they first lose their entitlement to IS. This might mean that some lone parents in the comparison group have been affected by the LPPC, which might violate the common trends assumption.
- Some of the other policy changes (discussed in Section 6.2) will not have been implemented during the roll-out.

But there are some features that make it less attractive to estimate the impact of the roll-out than the steady-state:

- As discussed in Section 5.5, working out exactly when lone parents will lose entitlement to IS is more complicated in the roll-out than in the steady-state, because it depends on the date of Work Focused Interviews (WFIs), for some phases. If an estimate made by researchers of when lone parents will lose their entitlement to IS (based on the information available to them in the Work and Pensions Longitudinal Study (WPLS)) is inaccurate, or if the implementation of the roll-out of the LPPC differs from the plans (and no systematic information is available on these discrepancies), then this will introduce some measurement error, and reduce the accuracy of estimates of the impact of the roll-out. By comparison, the relative simplicity of the steady-state rules, and the fact that there will be fewer lone parents making the transition from IS to JSA each month in the steady-state, make it more likely both that the exact conditions of the LPPC can be replicated in the WPLS, and that the implementation of the LPPC is as planned.
- As discussed in Section 3.4, the treatment and comparison groups during the roll-out will have a larger difference in the age of youngest child than in the steady-state.

On balance, though, the first set of arguments seem more significant than the second set, and so the internal validity of estimates of the LPPC from the roll-out is likely to be greater than those based on the steady-state.

However, it must be recognised that the impact of the roll-out of the LPPC is likely to be different from the impact of the steady-state of the LPPC for the following reasons:

- the roll-out affects different groups of lone parents (i.e. those whose youngest child is older) than the steady-state, and there is no reason why the impact of the LPPC should be the same on lone parents receiving IS regardless of the age of their youngest child. The roll-out, therefore, represents the only opportunity to estimate the impact on the LPPC on lone parents whose youngest child is aged 12 or more;
- the roll-out will take place at a different time (and therefore, under different economic conditions) from the steady-state. For example, the fact that the first phases of the roll-out are being implemented at a time when employment is falling and unemployment rising (and rising fast) may mean that, even if the LPPC is successful in increasing labour supply, there may not be enough labour demand, meaning that there might be no impact on work outcomes (Section 3.3 discussed what implication the recession has on the internal validity of a DiD estimator);
- the roll-out of LPPC might be implemented differently from the steady-state. In particular, given the considerable recent rise in JSA claimants, one might expect that Jobcentre Plus advisers would have less time to help lone parents during the early phases of the roll-out than they might in the steady-state;
- the LPPC may cause cultural changes among lone parents that happen some time after the LPPC is first introduced. These could not then be picked up by an IA of the roll-out, but would be in any IA of the steady-state.

In other words, the external validity of the estimates of the LPPC based on the roll-out may be lower than those based on the steady-state. But whether these represent significant disadvantages depends on whether there is interest in understanding the impact of the roll-out in itself, rather than merely using the estimates from the roll-out as a guide to the likely impact of the steady-state policy.

This report was asked to consider the feasibility of estimating the impact of the roll-out. The research team believe it is feasible, but just because something is feasible does not mean that it should be done. There are, of course, various benefits from estimating the impact of the roll-out which should be compared with the cost of undertaking it. One major benefit is that estimates of the impact of the roll-out would be available significantly earlier than estimates of the impact of the steady-state policy.

4.3 Separating the impact of the different elements of the lone parent policy change

An eventual IA based on the DiD method will be able to estimate the impact of the LPPC on individual lone parents at different points in time, measuring ‘time’ relative to the date at which the lone parent loses eligibility to IS as a lone parent. It cannot, however, determine which elements of the LPPC are causing these impacts.

It might seem reasonable to call the impacts measured after the date at which the lone parent lost eligibility to IS as ‘after effects’, to call impacts measured during the 12 months leading up to losing eligibility to IS as ‘pre-employment support impacts’, and to call impacts measured more than 12 months before losing eligibility to IS as ‘anticipation effects’, these should be thought of only as suggestive labels, rather than the outcome of a precise decomposition (the only thing for certain is that any impacts on lone parents before they are affected by any aspect of the LPPC were ‘anticipation effects’). For example, it would not be possible to separate formally any impact of the LPPC on lone parents in the 12 months leading up to losing eligibility to IS between that genuinely due to the support elements, and that due to the anticipation of losing entitlement to IS, and it would not be possible to separate any impact of the LPPC on lone parents after they had lost entitlement to IS into that due to pre-employment support and that due to in-work support.

A related question is whether the IA could separate the impact of different elements of the pre-employment and in-work support. Prerequisites for achieving this would be variation between lone parents in which services were received, and accurate data merged into the WPLS about which lone parents received which elements. The research team are currently unsure whether such administrative data will be available to researchers, but even if it was, and there was variation between lone parents in which services/elements were received/accessed, using such variation could not reliably infer the different impacts of different elements of LPPC. The reasoning behind this conclusion is identical to that which has led other authors to argue that it is not possible to distinguish between the different aspects of the Pathways to Work package. Adam *et al.* (2009) investigated whether one could separate the impact of the Choices package from that of other aspects of Pathways to Work, and concluded that:

‘...[P]articipation in Choices is voluntary, so it is difficult to know how far different outcomes for participants and non-participants are caused by Choices and how far they reflect pre-existing differences in the type of people who choose to participate. Using propensity score matching techniques, we account for differences between participants and non-participants in a very large set of background characteristics; we thus compare outcomes for Choices participants with those for non-Choices participants who are observably similar in many dimensions. But it remains likely that there are important differences in the unobserved characteristics of the two groups, and it is impossible to know how far the difference in outcomes between the two groups is a result of these unobserved pre-existing differences rather than a result of participating in Choices.’

The authors say that this reflects *'the intrinsic difficulty of evaluating programmes based on voluntary participation when there is no exogenous variation in the availability of the programme'*. Exactly the same argument applies to the LPPC: because all elements of the LPPC are available to all lone parents about to lose their entitlement to IS, any variation between lone parents in which services were accessed will reflect choices made by individual lone parents and their personal advisers, and it is likely that these choices reflect existing differences between lone parents, and that these differences themselves will partially determine subsequent outcomes.

In the absence of variation in the availability of the different elements of the LPPC, no quantitative method could formally separate the overall impact of the LPPC into that due to anticipation effects, support effects and after effects, or separate the different elements of the support package. This would remain the case even were additional data to the WPLS available. Therefore, any IA of the LPPC would give a single estimate of the effect of the whole of the LPPC. It is possible that qualitative research could be informative on this, but this would depend on how well lone parents are able to distinguish between the different parts of the LPPC, and how well they could consider how they would have behaved in the absence of one or more of them.

4.4 Summary

This chapter has argued that:

- estimating the impact of the LPPC on work-contingent outcomes (such as earnings or job retention) of lone parents is more difficult than estimating the impact on, say, initial flows off benefit. This is because if the LPPC led to a larger fraction of the treatment group gaining employment, then the employed treatment group members will not necessarily be the same, on average, as the employed comparison group members, and a simple comparison would confound the true effect of the policy with the differences in the employed members of the two groups, and usually lead to an estimate of the treatment effect which is too small. This is an example of a dynamic selection problem. The DiD estimator, and elaborations of that estimator, typically cannot overcome such dynamic selection problems, but duration models can be estimated in a way which may control for dynamic selection bias;
- there are different problems to contend with when examining the impact of the roll-out and the steady-state policy. On balance, the internal validity of estimates based on the roll-out is probably higher than those based on the steady-state, but estimates of the roll-out of the LPPC may be different from (and therefore, a poor guide to) estimates based on the steady-state. Whether these represent significant disadvantages depends on whether there is interest in understanding the impact of the roll-out in itself, rather than merely using the estimates from the roll-out as a guide to the likely impact of the steady-state policy;
- no quantitative method could robustly separate the overall impact of the LPPC into that due to anticipation effects, support effects and after effects, or separate the different elements of the support package, even were additional data available in the WPLS. Any IA of the LPPC would give an estimate of the effect of the whole of the LPPC.

5 Outcomes of interest and likely sources of data

This chapter discusses the main sources of data that could be used in an eventual impact assessment (IA), and what outcomes it will (and will not) be possible to examine with these datasets (Section 5.1). Section 5.2 discusses the problems in estimating the impact of the lone parent policy change (LPPC) on new and repeat claimants, Section 5.3 discusses how well an IA could inform a cost-benefit analysis, Section 5.4 gives an informal guide as to what impacts could be detected accurately using Work and Pensions Longitudinal Study (WPLS) data and Section 5.5 discusses the technical issue of identifying which lone parents are affected by the LPPC.

5.1 What outcomes could be addressed by an impact assessment?

The main sources of data identified by the research team that are relevant for the eventual IA are listed in Table 5.1. There are two types:

- administrative data (ie data held by the Department for Work and Pensions (DWP) based on individuals' benefit, tax credit, employment and earnings records, and hereafter referred to as the WPLS');
- repeated large-scale household surveys (hereafter referred to as 'household surveys').

These vary in terms of:

- the number of lone parents covered (the sample size);
- which lone parents are covered;
- the variety of data they contain on the population of interest;
- whether they are longitudinal or cross-sectional.

They also vary considerably in their ease of access and use, but this is not discussed further here; it is assumed that some way will be found for researchers to access the WPLS in any eventual IA.

Table 5.2 assesses whether an eventual IA would be able to assess the impact of the LPPC on a wide range of outcomes. Some general points about Tables 5.1 and 5.2 are:

- the nature of the policy intervention (the LPPC is designed primarily to help and encourage lone parents who are receiving out-of-work benefits to move into work) means that longitudinal data is a great help, as it allows researchers to identify the directly affected population, and then to track their outcomes over time. This is one of the most important strengths of the WPLS. Of the longitudinal datasets, the British Household Panel Survey (BHPS) has too small a sample, and *Understanding Society*, its successor, will not be informative on the pre-LPPC cohorts; the sample size of the Labour Force Survey (LFS) is more promising, but the longitudinal aspect covers only 15 months of data, meaning only short-run impacts of the LPPC can be estimated from the LFS;
- to be of use in a difference-in-difference (DiD) framework, data needs to be available on a comparison group. Given the group suggested in Section 3.3 – lone parents with younger children – this means that the LPO Survey will not be able to inform directly the eventual IA. It is, therefore, not mentioned in Table 5.2;

- to be able to identify the directly affected population not already receiving Income Support (IS) (see Section 2.3), a dataset needs to sample from the whole population of lone parents, and include measures of which benefits lone parents are receiving. This is achieved comprehensively only by the household surveys, but Section 5.2 discusses the use of the WPLS to estimate the impact of the LPPC on repeat claimants;
- most of the household surveys have sample sizes which will be too small to estimate small impacts precisely. It is, therefore, possible that an eventual IA will be unable to show conclusively that the LPPC has had any effect at all even if there genuinely is an impact.

The impact of the LPPC on work entry can be easily measured in a DiD framework for the whole directly affected population, but the impact of the LPPC on outcomes of those lone parents who enter work – such as on their earnings or duration in employment – can be estimated only if the issue of dynamic selection bias is addressed. As discussed in Section 4.1, dynamic selection bias arises when examining these outcomes because the LPPC is likely to increase the number of lone parents in work and therefore, change the characteristics of lone parents in work. Simply comparing lone parents in work after the LPPC has started with those who started work before the LPPC will confuse the effects of the LPPC with effects caused by the change in the composition of lone parents in work before and after the LPPC is introduced. This problem is a fundamental feature of the nature of these outcomes – namely, that they are only observed for the sub-group of the population that makes the transition from benefit to work – rather than a result of any limitation in the data available.

Table 5.1 Summary of datasets

	Sample size of lone parents	Cross-sectional or longitudinal	Can it identify directly affected population?	Can it identify indirectly affected population?	Is there data on comparison group?	Particular strengths	Particular weaknesses
Administrative data							
WPLS	100 per cent sample of lone parents ever on IS, and their benefit/work/ tax credit history and future	Longitudinal, with data held in spell format	Yes	Partially (see Section 5.2)	Yes	Longitudinal, large sample size	Limited number of outcomes
Household surveys							
Family Resources Survey (FRS)/ Households Below Average Income (HBAI)	c2,000 a year	Annual cross-section	Yes	Yes	Yes	Measure of net income and poverty status used in official statistics	Not longitudinal
LFS	c4,200 a quarter, although earnings available for only 40 per cent of sample	Rotating panel: each household interviewed for five quarters	Yes	Yes	Yes	Large sample for a household survey. Focus on labour market outcomes	Not all variables available in single version of LFS. Little on income. Short panel
General Household Survey (GHS)	c700 a year	Rotating panel: each household interviewed for four years	Yes	Yes	Yes	None	Small sample
Family Expenditure Survey/Expenditure and Food Survey/ Living Costs and Food Survey	c550 a year	Annual cross-section	Yes	Yes	Yes	Measures expenditure	Small sample

Continued

Table 5.1 Continued

	Sample size of lone parents	Cross-sectional or longitudinal	Can it identify directly affected population?	Can it identify indirectly affected population?	Is there data on comparison group?	Particular strengths	Particular weaknesses
BHPS and Understanding Society	BHPS c300 a year; Understanding Society: c2,000 a year	Longitudinal with annual interviews	Yes	Yes	Yes	Longitudinal, and wide range of outcomes	Larger sample of <i>Understanding Society</i> not available before 2009

Table 5.2 Outcomes

Outcome	Feasible to estimate for directly affected population?	If not, feasible to estimate for other population?	Most suitable datasets. Other comments
Main outcomes			
Take-up of New Deal for Lone Parents (NDLP) or other Jobcentre Plus provision	Yes	n/a	WPLS, if these other services are recorded in WPLS and are available to comparison group and have been available over time (such as NDLP, In-Work Credit (IWC))
Flows off benefit	Yes, including flows between benefits, and including anticipation effects	n/a	WPLS, LFS (only within 12 month window)
Job entry	Yes	n/a	WPLS (subject to noisiness of P45/P46 measure of work), LFS (only within 12-month window)
Entry to job of 16+ hours	Yes	n/a	WPLS (if tax credit data provides a reliable measure of work of 16+ hours), LFS (only within 12-month window)
Sustained employment/job retention	Yes, depending on definition. Dynamic selection bias may apply, depending on definition ¹	n/a	WPLS (if tax credit data provides a reliable measure of work of 16+ hours and/or subject to noisiness of P45/P46 measure of work)
Household income and child poverty (including composition of poor, and child poverty rates for subgroups)	No (FRS will give estimates of household income and poverty status of directly affected population, but only when they are on IS, which is of limited interest)	Yes, but would need to have large impact, given sample size	FRS/HBAI.
Mentioned in international literature review ²			
Benefit on-flows			
Flows from IS to other benefits	Yes (including at what stage)	n/a	WPLS
Flows on to IS	n/a	Possibly	WPLS (see Section 5.2), LFS (only within 12-month window)
			Continued

Table 5.2 Continued

Outcome	Feasible to estimate for directly affected population?	If not, feasible to estimate for other population?	Most suitable datasets. Other comments
Benefit off-flows	Yes (for individual benefits and for all benefits, including anticipation effects)	n/a	WPLS
Proportion of benefit caseload doing mini-jobs	Yes	n/a	LFS (only within 12-month window) ³
Job entries	Yes (including at what stage)	n/a	WPLS, LFS (only within 12-month window)
Whether job part-time or full-time	Yes, although dynamic selection bias may apply depending on definition	n/a	WPLS (if tax credit data in WPLS records whether individuals are working 16-29 or 30+ hours), LFS (only within 12-month window)
Hours worked	Yes, although dynamic selection bias may apply depending on definition	n/a	WPLS (if tax credit data in WPLS records self-reported hours worked), LFS (only within 12-month window)
Increase in labour supply/overall employment rate	n/a ⁴	Yes	LFS (only within 12-month window), FRS
Recidivism	Yes, although dynamic selection bias may apply, depending on definition	n/a	WPLS
Displacement, substitution	n/a	No: see Section 5.3	n/a
Characteristics of job (industry, occupation, shift work, managerial, permanency)	Yes	Yes	LFS (only within 12-month window)
Weekly earnings and hourly wages	Yes	n/a	LFS (only within 12-month window and sample size is 40 per cent of that indicated in Table 5.1). Given sample size, would need to have large impact to be estimated precisely. See Appendix A for discussion of measuring earnings in WPLS

Continued

Table 5.2 Continued

Outcome	Feasible to estimate for directly affected population?	If not, feasible to estimate for other population?	Most suitable datasets. Other comments
Income from benefits	Potentially	Yes	For directly affected population: in theory, WPLS (see Appendix A) For population: FRS (would need to have a large impact to be detected)
Income from tax credits	Potentially	Yes	For directly affected population: in theory, WPLS (see Appendix A) For population: FRS (would need to have a large impact to be detected)
Other sources of income	No (FRS will give estimates of household income and poverty status of directly affected population, but only when they are on IS, which is of limited interest)	Yes	FRS (would need to have a large impact to be detected)
'Income gains', financial gains to work	Yes, but sample sizes will be small	n/a	LFS (together with tax and benefit calculations) It is conceivable that WPLS data could be used to construct a partial measure of gains to work, but it would require the accurate data on benefit income, earnings and tax credit income to be available, as discussed in Appendix A
Spending patterns	No	Yes, but sample sizes will be small	EFS
Impact of individual policy elements	No (see Section 5.3: this is a methodological barrier, not a data limitation)	No	This is a methodological barrier, not a limitation of the datasets
Deterrent effect of sanctions	No	No	This is a methodological barrier, not a limitation of the datasets: would need qualitative research
Proportion entering 'unknown' destination	Yes	n/a	WPLS ⁵

Continued

Table 5.2 Continued

Outcome	Feasible to estimate for directly affected population?	If not, feasible to estimate for other population?	Most suitable datasets. Other comments
Family formation/characteristics; number or proportion of lone parent households	Yes	Unsure	WPLS, LFS (only within 12 month window)
Persistent poverty	No	No	None (<i>Understanding Society</i> started too late to provide data on pre-LPPC lone parents)

- ¹ 'Proportion of time spent in employment' is defined for all directly affected lone parents, and therefore, dynamic selection bias is irrelevant. A variable 'Job lasts more than 13 weeks' is defined only for lone parents who work, and therefore dynamic selection bias is relevant.
- ² Finn and Gloster (2010).
- ³ The National Benefits Database (NBD) measures earned income if this is above the earnings disregard and is reported to DWP. The research team suspect that this will under-estimate instances of mini-jobs.
- ⁴ The impact on the employment rate could be estimated with WPLS as 'number of additional jobs/lone parent population', but this would miss impact on new/repeat claimants, and ignore any substitution and displacement.
- ⁵ The LFS might be informative about what the nature of the 'unknown' destinations in WPLS is because LFS is a longitudinal dataset, there will be no 'unknown' destinations (except where there is sample attrition).

5.2 Estimating the impact of the lone parent policy change on new and repeat claimants

Section 2.3 outlined that lone parents are affected by the LPPC in two ways:

- Group 1: lone parents already on IS, who have their eligibility to receive IS as a lone parent removed earlier than it would otherwise have been (DWP refer to this group as ‘existing customers’);
- Group 2: lone parents not on IS, some of whom are no longer eligible to start an IS claim if they wish to do so.

In order for an IA to analyse the effect of the LPPC on group 2 (lone parents who are no longer eligible to claim IS), it is necessary to have data on lone parents who are not claiming an out-of-work benefit but might start a claim (in other words, it is not sufficient to have data only on those lone parents who make a new or repeat claim; it is necessary also to have data on those lone parents who could have made, but did not, a new or repeat claim).

The WPLS contains data on all individuals who have claimed a DWP benefit at any point since 1999 that relates to the time that they are receiving a DWP benefit. By construction, it does not contain information on these individuals when they are not on benefit (other than being informative of the fact that they are not receiving a DWP benefit). Therefore, it cannot be used to determine which individuals are lone parents during periods when they are not receiving a DWP benefit, nor is it informative about the age of their youngest child. This clearly makes it impossible to identify which former benefit recipients are directly affected by the LPPC by being no longer eligible to claim IS.

However, if these individuals were still eligible for and claiming tax credits (more than 90 per cent of lone parents are eligible for tax credits, and take-up rates are high), then administrative data on these individuals should be available through Her Majesty's Revenue & Customs' (HMRC's) tax credit data and contained in the WPLS. If this data was of good quality, it would, therefore, be possible to use administrative data to estimate the effect of the LPPC on the probability of lone parents making a repeat claim. The reliability of this data could be checked in the following ways:

- how many IWC recipients also appear to start a WTC claim at around the same time as they start their IWC claim. This would give some indication as to how reliable eligibility to WTC is as a measure of whether a lone parent is in work (because all IWC recipients must be working at least 16 hours a week, the same condition as for eligibility to WTC);
- how many lone parents starting or ending an IS claim can be tracked before or after (respectively) that date using the tax credit data, either as a lone parent or a member of a couple.²¹

This method would not allow the WPLS to be informative about the impact of the LPPC on new claimants, though, because the WPLS only contains data on those individuals who have claimed a DWP benefit since 1999. Whether this omission is a problem depends on how many lone parents whose youngest child is seven or over have never received a DWP benefit since 1999, and on how accurately the tax credit data is at recording who is a lone parent. Calculations, based on the WPLS, suggest that, of all claims of IS which were begun by lone parents whose youngest child was aged seven or over in 2006/07, 36 per cent of individuals had not previously received IS. This suggests that the WPLS will not be informative about the impact of LPPC on new and repeat claimants.

²¹ Note that Brewer *et al.* (2009) showed that there were more lone parents receiving support for their children through either IS or tax credits in 2006/07 than were thought to live in the UK.

If new claimants did appear to be a significant group, an alternative data source would need to be found for this analysis. As Table 5.2 stated, the longitudinal LFS could be used for this purpose: this provides data on lone parents' work and benefit status at quarterly intervals for five quarters. Looking at whether lone parents not initially on benefit start a benefit claim in the year they are followed would enable an IA to estimate whether the LPPC had any effect on the propensity of lone parents to start new (or repeat) benefit claims. This would again use a DiD methodology, comparing the proportion of lone parents with children above and below the age limit not on benefit starting a new benefit claim before and after the reform. In this case, the difference between the treatment and comparison groups in the age of their youngest child would not need to be so large because it is less likely that there would be large anticipation effects of the LPPC on this group. The only constraint would be that the comparison group should still be eligible for IS at the end of the comparison period. Since the comparison period need not be more than a year (as the outcome of interest is whether a lone parent starts a new IS claim or not), the difference between the treatment and comparison groups in terms of the age of their youngest child should only be one year.

5.3 Substitution, displacement and other issues relevant to a cost-benefit analysis

Table 5.3 shows the research team's opinion on the feasibility of the IA estimating various outcomes that are important for a cost-benefit analysis:

Table 5.3 Feasibility of estimating various outcomes needed for a cost-benefit analysis

Outcome	Feasible to estimate?
Job entry rate	IA would estimate change in job entry rate due to LPPC
Benefit exit rate	IA would estimate change in benefit exit rate due to LPPC
Cost per additional job	Additionality estimate from IA, cost of policy from elsewhere
Cost per additional benefit exit	Additionality estimate from IA, cost of policy from elsewhere
Net fiscal benefit from additional jobs	Additionality estimate from IA, fiscal benefit per additional job from elsewhere
Net fiscal benefit	No
Fiscal benefits/cost ratio	No
Net economic benefits	No
Impact on earnings in non-additional jobs	Dynamic selection bias a problem here; requires duration model. Chapter 5 discusses whether impact on earnings can be measured accurately
Impact on hours in non-additional jobs	Dynamic selection bias a problem here; requires duration model. Chapter 5 discusses whether impact on hours can be measured accurately
Spells rather than individuals	Additionality can be estimated in this way
Deadweight	Will estimate additionality, and this allows deadweight to be calculated
Additional time spent in employment	IA could estimate additional time in employment due to LPPC
Additional time spent off benefit	IA could estimate reduced time on benefit due to LPPC
Job duration	WPLS can look at length of employment spells, which are not necessarily the same as job spells
Wages	LFS (only within 12-month window). Given sample size, would need to have large impact to be estimated precisely See Appendix A for discussion of measuring earnings in WPLS
Participation in other interventions	Can use administrative data to examine effect on participation in NDLP, etc.
Wider impacts	No
Substitution and displacement effects	No
Sensitivity analysis	No

The research team consider that it would not be possible to estimate the impact of the LPPC on groups not directly affected by it. Since the LPPC affects all areas of the country at the same time, everyone in the country would be potentially affected by substitution and displacement effects (whether through employment or wages or prices). This means that there is no obvious group who could act as a comparison or comparison group. Methods which have been used in previous evaluations are not applicable given the design of the LPPC (see Box 5.1).

Box 5.1 – Estimating spillover effects in UK welfare-to-work evaluations

The Pathways to Work evaluation (Adam *et al.*, 2008) examined the effects of that policy on claimants of other benefits who were not directly affected by that policy change using a DiD approach. This was possible because Pathways to Work was only introduced in some Jobcentre Plus districts, and so claimants of other benefits in areas where Pathways to Work was not operating could be used as a comparison group (as they could be thought of as operating in different labour markets). Since the LPPC is being introduced in all areas of the country simultaneously, this strategy could not be pursued in this case.

Blundell *et al.* (2004) attempt to look for spillover effects on the comparison group in a DiD evaluation of the New Deal for Young People (NDYP) by using a different comparison group that they believe is less substitutable for the treatment group in the labour market. They argued that if they found a smaller impact of the policy using this alternative comparison group, then that would suggest that there were spillover effects that harmed the original comparison group. However, using a comparison group that is less similar to the treatment group would make the ‘common trends’ assumption less believable, threatening the internal validity of the evaluation. Blundell *et al.* use those who are slightly older than the treatment group as their main comparison group, and those who are slightly older again as an alternative comparison group. This strategy cannot be pursued for the LPPC because of the inability to distinguish between spillover effects and anticipation effects.

Other than this, the eventual IA should be able to come up with the month-by-month estimates of the additionality necessary for a cost-benefit analysis to be performed. As discussed in Section 4.1, dynamic selection bias becomes a problem when examining the impact of the LPPC on earnings, hours and job duration in work, but a duration model would be able to get around this problem under various assumptions.

5.4 How precisely can any impacts be estimated?

DWP officials have estimated how many lone parents will be affected by the different phases of the LPPC. Appendix D lists the number who will reach the point where their IS entitlement ends: as it is recommended that any IA also examines outcomes in the 12 or 24 months leading up to this point, then the sample sizes are likely to be a little larger than this. It shows that the smallest number of lone parents affected by any phase is around 10,000.

But naïve estimates of the size of impact that could be reliably detected with such sample sizes are likely to be too optimistic, for the reasons set out in Section 3.5. As that section concludes, the assessment of the research team is that the impacts estimated from the WPLC will need to be around five to ten percentage points to be considered robust and meaningful.

5.5 Identifying which lone parents are affected by the lone parent policy change, and when

Chapter 2.1 noted that not all lone parents lose their entitlement to IS earlier than they otherwise would have as a result of LPPC. Some of these groups can be identified using the WPLS (e.g. those claiming other benefits, asylum seekers and those with children eligible for DLA), but this is not possible for all of the groups.²²

For these groups, wrongly assuming that they were affected by the LPPC would mean that a DiD estimate would underestimate the true impact of the LPPC (this is a standard result where a measurement error in the treatment variable leads to attenuation bias). How substantial this bias is will depend on what fraction of lone parents are exempted and where this cannot be deduced from administrative data.

Carrying out an IA will also require researchers to identify the date when a lone parent would have lost their entitlement to IS, based on data available to the researchers. As discussed in Section 2.1, the actual date will depend on the age and date of birth of their youngest child and, in some cases, on the date their WFI is due.

To identify the affected population, it is absolutely crucial for this IA that age of youngest child is recorded accurately in the WPLS.

Even if the date of birth of the youngest child can be recorded accurately, identifying the date when a lone parent would have lost their entitlement to IS is more complicated for those who lose their entitlement to IS after a WFI. Lone parent advisors record each meeting that is booked on the Labour Market System, noting its type and whether the meeting was actually attended. This data can be merged in to the WPLS, which would make it possible to obtain a reasonable estimate of when the next WFI for a particular lone parent was due. This is, of course, dependent on the WFI regime being implemented as it is intended, and on all lone parents having WFI dates recorded accurately.

One concern here relates to lone parents who are given a four week extension to their IS claim because of problems processing their CTC claim. (This would only be relevant to IS claims that have been ongoing since before 2004 and where lone parents were having this support paid as child additions to IS.) It appears that these cases will not be identified in the administrative data, and this will, therefore, create some inaccuracy in determining when a lone parent was affected by the LPPC. As discussed above, this may lead to the DiD estimate of the LPPC being lower than the actual impact through an attenuation bias.

Finally, a related problem would arise if lone parents are not moved off IS precisely when they ought to be according to the rules (summarised in Appendix B). If these rules are not adhered to in practice, then this will lead to errors in determining when lone parents are actually affected by the

²² Obviously, we would identify them if they remained on IS after they would otherwise have been forced to leave it had the exemption not existed. But this technique would not identify otherwise identical lone parents in the comparison group, or those observed before the policy came into effect, as would be required to perform a consistent DiD analysis. One possibility would be to perform a Bloom adjustment (Bloom, 1984), where the impact estimates are divided by the proportion of lone parents who are actually affected by the LPPC. But this requires us to know the proportion of lone parents who are exempted, which we have already said is not possible using the data available to researchers. Therefore, it does not seem likely that this approach is feasible.

reform. One way to view this is that it is another form of measurement error, and would, as above, lead to the DiD estimate of the LPPC being lower than the actual impact through an attenuation bias. Alternatively, the DID estimate can be thought of as estimating the impact of an ‘intention to treat’ (ITT): that means that it estimates the intended impact of the LPPCs, but acknowledges that the policy is not always implemented precisely as intended.

Finally, analysis done on datasets other than the WPLS could be more problematic. Other datasets do not tend to publish children’s dates of birth (to reduce the risk of individuals being identified), and there is no other source of data on WFI dates. This means that analysis on these datasets will probably need to use age of youngest child measured in years as a cruder indicator of which lone parents were directly affected by the LPPCs.

5.6 Summary

- Two groups of lone parents are directly affected by the LPPC: those receiving IS who lose entitlement to it sooner than they would have done otherwise, and those who are no longer entitled to claim IS. Furthermore, all may be affected by general equilibrium effects, such as substitution, displacement or other spillover effects.
- For the first group, administrative data can be used to measure outcomes relating to benefits and, if tax credit data is of sufficient quality, work outcomes. If information on earnings and tax credit payments were made available within the WPLS, then administrative data could be used to partially estimate the effect of the LPPC on family income. This would not be sufficient to estimate the effect of the LPPC on child poverty as currently defined, however.
- The WPLS may be able to be used to estimate the impact of the LPPC on repeat claimants, but cannot, through its nature, be used to estimate the impact of the LPPC on new customers.
- The impact of the LPPC on benefit, work, income and poverty outcomes could also be measured using household surveys, such as the LFS and FRS, if administrative data on these was of low quality or unavailable. However, small sample sizes would mean that impacts would have to be large in order to be detected reliably.
- It would not be possible for an IA to separate out the effects of the different elements of the LPPC, or examine displacement or substitution effects on the population as a whole. This is because of methodological barriers rather than inadequacies of the available data.
- Provided the date of birth of the youngest child based on Child Benefit records can be recorded accurately in the WPLS, it should be possible to replicate in the WPLS the rules for when lone parents lose their IS entitlement with a reasonable degree of accuracy, and so any loss of accuracy should be minimal. Other datasets will necessarily be less accurate in determining precisely when a lone parent lost entitlement to IS.

6 Potential confounding factors

This chapter discusses technical or methodological problems that could affect an eventual impact assessment (IA). Where possible, it suggests modifications to the basic empirical strategy, outlined in Chapter 3, that would help to overcome these problems. Section 6.1 discusses how to deal with anticipation effects and Section 6.2 discusses how an IA could deal with other policy changes (both those that have taken place since 2008, and those that might happen in the future).

6.1 Anticipation effects

‘Anticipation effects’ refer to a situation where a lone parent changes their behaviour because they expect to be affected by a policy change in the future

Section 2.2 argued that the lone parent policy change (LPPC) should be thought of as starting to affect a lone parent 12 months before the loss of Income Support (IS) entitlement. This means that any changes in behaviour occurring in those 12 months would be considered as a response to the LPPC. However, there may be anticipation effects to the LPPC before this point if lone parents are aware of the loss of IS entitlement before they have their first Final Year Quaterly Work Focused Interview (FYQWFI). If an IA failed to allow for these effects, it would bias the estimated impact of the LPPC based on a difference-in-difference (DiD) estimate in two ways:

- the DiD estimate would not capture some of the genuine impact of the LPPC (i.e. be biased downwards);
- the anticipation effects would alter the composition of the treatment group, and this would mean that a DiD estimate would be capturing the impact of the LPPC and any change in the characteristics of the lone parents caused by these anticipation effects (see also Chapter 4).

It would be possible for the eventual IA to check whether there are anticipation effects earlier than 12 months before the loss of IS entitlement. This could be done straightforwardly by changing the date at which it is considered the LPPC first affects a lone parent. For example, if it were suspected that lone parents knew about the future loss of entitlement to IS and would alter their behaviour as a result, up to 24 months before losing that entitlement, then the IA could define the lone parents who are 24 months (rather than 12 months) away from losing eligibility to IS as ‘directly affected by the LPPC and already receiving IS’. This would mean that the IA accounted for any effect of the policy in the period between one and two years before IS eligibility was removed. As a consequence of this, the comparison group would have to be re-chosen (so that their youngest children were a further 12 months younger); Appendix C makes this explicit for the roll-out stage of LPPC.

In general, it is impossible to be sure that one has gone back far enough to eliminate the possibility of any anticipation effects before that time, although it might be possible to argue for some of the first phases that there could be no anticipation effects earlier than a certain date (for example, although the Government announced its desire to make this change in 2007²³, there was little publicity until the policy actually started in November 2008²⁴).

²³ See, for example, http://news.bbc.co.uk/nolpda/ukfs_news/hi/newsid_6904000/6904520.stm

²⁴ See, for example, http://news.bbc.co.uk/nolpda/ukfs_news/hi/newsid_7742000/7742714.stm

6.2 Other policies

This section discusses relevant changes to policy which took place just before, or during, the period over which the LPPC was being rolled out, or which may take place in the near future.

In general terms, other policy changes will confound an IA if they affect the treatment and comparison groups differently. In such a case, the ‘common trends’ assumption underpinning the DiD methodology would not hold (as trends in the outcomes of the treatment and comparison groups would be different in the absence of the LPPC as a result of these other policies). Therefore, in general, other policy changes need to be accounted for in any IA. Broadly speaking, there are two ways in which this can be done:

- if the effect of these policies is the same on the treatment and comparison groups, ignoring the existence of these policies would not bias the estimates in the DiD model or duration model. In this case the additional policies form part of the common trends affecting both groups and so do not invalidate the DiD estimator (see Section 3.3);
- if the effect of these policies on the treatment group is not identical to that of the comparison group, then the policies can be modelled explicitly in a DiD regression or duration model by including an additional explanatory variable which indicated those lone parents who were affected by this additional policy. Note that this would only work if there was some variation in who was affected in both the treatment and the comparison groups (i.e. some lone parents in both the treatment and the comparison groups would have to be unaffected by this policy). This approach would, therefore, work for policies that are going to be piloted in certain areas or made available to selected lone parents receiving IS or Jobseeker’s Allowance (JSA), but not for those that will be immediately available across the whole country. It also requires that the effect of the LPPC does not vary systematically by area.

Neither of these techniques will work, though, if a policy is introduced which affects only lone parents with younger children, and is introduced in all areas of the country at the same time. In this case, the DiD estimator cannot be used to estimate the effect of the LPPC, as the comparison group will be affected by something that does not affect the treatment group, violating the common trends assumption. The DiD estimator would then be estimating the net effect of the LPPC minus the effect of the other policy. The most obvious example of this is if LPPC is extended to lone parents with children aged five to seven.

The rest of this section discusses some particular other policies in more detail.

6.2.1 Policies which have already taken effect

JSA and Flexible New Deal

In April 2009, the JSA regime changed, with a policy known as Flexible New Deal (FND), which affected the support available to all JSA customers. This initially applied in certain Jobcentre Plus districts, with the remaining districts affected from April 2010.

A simple way to reflect this would be to add an additional explanatory variable to the DiD or duration model to indicate which lone parents were subject to FND. Such a framework could potentially estimate both the impact of the LPPC and the additional impact of FND. This would essentially be a DiD evaluation of FND involving the districts not initially operating FND as comparison areas.

However, if it were thought that FND alters the nature of the LPPC, then it should be possible to estimate the impact of the LPPC separately in the FND areas, thus giving two separate estimates of the effect of the LPPC, depending on whether it is in operation in conjunction with the FND. Obviously, this could only be done in the period where some Jobcentre Plus districts were not operating FND, and so it would not be able to estimate medium- to long-run outcomes.

It is planned that a new Work Programme will be in place nationally by summer 2011 and FND will be phased out and folded into the Work Programme.

Employment and Support Allowance

Employment and Support Allowance (ESA) replaced Incapacity Benefit (IB) for new claimants from October 2008, just before the LPPC began.

It will not be possible to identify separately the impact of ESA from that of removing the entitlement to IS, because there will not be any lone parents who had entitlement to IS removed before their youngest child is 16 before ESA replaced IB for new claimants. A separate problem arises from the fact that the disability conditions for claiming ESA are stricter than those for IB, which meant that lone parents on IS with a work-limiting disability facing withdrawal of their entitlement had an incentive to start an IB claim before IB was closed to new claimants in October 2008 (rather than wait until the end of their IS entitlement and make an ESA claim, which might be rejected). This would be an anticipation effect of the LPPC that occurred quite a long time before the withdrawal of IS entitlement in some cases. As stated in Section 6.1, moving the effective start date of the programme further before the reform would allow for these effects to be taken into account.

In-Work Credit roll-out

In-Work Credit (IWC), a payment of £40 a week (£60 in London) for the first year of work for lone parents who have previously been receiving IS or JSA for at least a year, has been available nationally since April 2008. It was previously available in certain Jobcentre Plus districts, covering around 45 per cent of lone parents receiving IS. Therefore, this change only affects lone parents in districts that did not previously have IWC.

The extension of IWC affects both the treatment and comparison groups, so failing to account for this would lead to the DiD approach estimating the impact of the LPPC plus any differential impact of IWC on the treatment group relative to the comparison group. However, as with FND, adding an additional explanatory variable to the DiD or duration model to indicate which lone parents are potentially eligible to IWC (i.e. those who have been on benefit for at least a year and are living in a pilot area prior to April 2008 or anywhere from April 2008, working 16 hours or more a week) would give correct estimated impacts of the LPPC, and of the national roll-out of IWC.

Child poverty pilots

Ending child poverty: everybody's business, published alongside the Budget report 2008, announced the details of a suite of pilots to develop new ways to tackle child poverty.²⁵ These pilots are aimed more at reducing child poverty rather than increasing lone parent employment, although policies such as improving childcare provision affect both. However, any IA ought to take account of these policies to avoid a situation where effects of the child poverty pilots are wrongly attributed by the DiD method to the LPPC. As with FND and the national roll-out of IWC, this can be done very simply by adding additional explanatory variables to indicate which lone parents live in these local

²⁵ HM Treasury (2008). *Ending child poverty: everybody's business*. London: The Stationery Office, http://www.hm-treasury.gov.uk/d/bud08_childpoverty_1310.pdf

authorities at a time when these pilots were running; doing this would allow the DiD model to estimate correctly the impact of the LPPC and estimate separately both the impact of the child poverty pilots on lone parents receiving IS.

6.2.2 Policies which have yet to take effect

The Work Programme

At the time of writing, it is planned that a new Work Programme will be in place nationally by summer 2011. The Work Programme will replace the FND. It will support a range of customers and will be more personalised, with the support supplied being determined by the needs of the individual.

In the context of an IA of the LPPC, the Work Programme is best thought of as a change to the treatment: rather than JSA and FND, lone parents who claim JSA will experience JSA and the Work Programme. One approach would be for an eventual IA to stop analysing data at the point at which the Work Programme begins; this would curtail the ability of an eventual IA to examine the medium- to long-run impact of LPPC on lone parents with children aged over seven. Alternatively, an eventual IA could take account of this by estimating separate impacts of LPPC before and after the Work Programme (something which will be easier to do if the Work Programme is rolled out geographically in stages).

Extension of LPPC to lone parents whose youngest child is aged five and over

In the Emergency Budget in June 2010, the Government announced its intention to extend the LPPC to lone parents whose youngest child is aged five and six. This is expected to be introduced in early 2012. Because this affects the proposed comparison group, this policy, if implemented, will have a substantial impact on any IA of LPPC.

As details of this proposed extension are unclear at the time of writing, it is not possible to say definitively what impact it would have on an eventual IA. But there are three broad approaches (two of which are similar to the options discussed above for dealing with the Work Programme). First, an eventual IA could stop examining data at the point at which the LPPC was extended (or, if anticipation effects are suspected, an earlier point at which it is suspected lone parents with children under seven started to alter their behaviour significantly). This would curtail the ability of an eventual IA to examine the medium- to long-run impact of LPPC on lone parents with children aged seven and over. Second, an eventual IA could use lone parents with even younger children as the comparison group, such as those whose youngest children were aged under 36 months. Third, an eventual IA could decide to stop estimating the impact of LPPC for lone parents whose youngest child was aged seven or more at the time when it was extended to those aged five and six, and instead estimate the impact of LPPC for lone parents whose youngest child was aged five and six using lone parents whose youngest child was aged seven or more as the comparison group; in other words, the definition of the treatment group and comparison groups would be flipped. This would be a reasonable thing to do provided that the policy regime for lone parents whose youngest child was aged seven or more was relatively stable at that time: the key assumption required for this to be appropriate would be that the impact of LPPC on lone parents whose youngest child was aged seven or more was stable during the time when it was extended to those whose youngest child was aged five and six.

6.3 Summary

There are several potential confounding factors that could affect various parts of an eventual IA. In particular:

- anticipation effects can be dealt with by changing the effective start date of the programme in the model. The drawback is that this increases the differences in ages of the children in the treatment and comparison groups, making the common trends assumption underlying the difference-in-difference approach less plausible;
- other policy changes affecting lone parents on IS and JSA have been introduced since 2008. Most of these can be controlled for relatively easily by introducing additional explanatory variables which identify those lone parents who are affected. By accounting for many of these policy changes, the eventual IA of the LPPC will also provide estimates (based on DiD) of the impact of these policies on lone parents on IS. However, some programmes that affect all lone parents cannot be controlled for in this way: any IA would have to assume that lone parents are affected by the same amount by these reforms (which would mean that any impact of these reforms are incorporated in the common trend affecting all lone parents).
- some likely future policy changes – such as the introduction of the Work Programme – will have more significant implications for an eventual IA: the Work Programme will, in principle, change the nature of the policy regime for lone parents on JSA, and an extension of the LPPC to lone parents with younger children will directly affect the proposed comparison group. The most robust option would be to stop examining data when the Work Programme starts or LPPC is extended to more lone parents, but this would curtail the ability of an eventual IA to examine the medium- to long-run impact of LPPC on lone parents with children aged seven and over. An alternative approach that deals with the Work Programme would be to estimate separate impacts of LPPC before and after the Work Programme. Alternative approaches that deal with the extension of LPPC would be to use lone parents with children aged under three as the comparison group, or to flip the definition of the treatment group and comparison group, using the lone parents whose youngest child is aged seven or over to estimate the impact of the LPPC on those whose youngest child is aged five and six.

7 Conclusions and recommendations: what is and is not feasible in an impact assessment of the lone parent policy change

This report has investigated the feasibility of undertaking an impact assessment (IA) of the policy change known as ‘Lone Parent Obligations and New Services for Lone Parents’ (together called ‘lone parent policy change’, or LPPC).

Any impact assessment of the LPPC will have to overcome many difficulties. Some of these are due to its design: the policy is being rolled out very quickly in a uniform way across the country. Some of these reflect its timing: it is being rolled out during a recession, and there has not been a comparable recession for over 15 years. Some of these reflect the wider policy environment: other policy changes took place just before or after LPPC which will affect lone parents receiving Income Support (IS), and larger changes are planned for 2011 and 2012.

However, the research team’s assessment is that an IA for some groups of affected clients and for some outcomes is, in principle, feasible using the difference-in-difference (DiD), or trend-adjusted DiD, model approach. The main limitations are that an eventual IA will probably be limited to those outcomes which can be measured with administrative data (the Work and Pensions Longitudinal Study (WPLS)), that any DiD estimator is always subject to a number of untestable assumptions, and that it seems unlikely that the DiD approach could reliably detect small impacts (say less than two percentage points).

The rest of this chapter concludes on the form of IA that seems to hold the most promise.

7.1 Empirical methods

The research team’s recommended empirical strategy is to use a DiD approach to examine the impact of the LPPC on those lone parents receiving IS who will lose entitlement to it sooner than they would have done otherwise. A DiD approach effectively compares the outcomes of lone parents directly affected by the LPPC with those of lone parents whose children are sufficiently young so that they are not affected; this comparison will then be compared to the difference in outcomes of similar groups of lone parents observed before the LPPC. The idea is that differences in outcomes of lone parents with differently aged youngest children before the LPPC begins are informative about what the differences in outcomes of lone parents with differently aged youngest children would have been in the absence of the LPPC.

The key assumption needed for a DiD estimate to be valid is known as the ‘common trends’ assumption. In the case of the LPPC, common trends would fail if the labour market behaviour of lone parents with younger children changed in a way which was different from the change in the labour market behaviour of lone parents with older children. There is some evidence that the

78 Conclusions and recommendations: what is and is not feasible in an impact assessment of the lone parent policy change

common trends assumption does not always hold during the period from 2001 to 2007, but the divergences are usually small, may not be statistically significant, and they could be accounted for in an eventual IA with a trend-adjusted DiD model, especially if a long period of pre-programme data is used.

Some outcomes, such as the duration on benefit and the number of lone parents starting work, could be examined using either a conventional DiD regression or a duration model. However, a DiD estimator cannot be used to estimate the effect of the LPPC on work-contingent outcomes such as job duration or earnings (assuming suitable data are available) due to dynamic selection bias. A duration model, though, offers a promising way of examining these outcomes.

However, it is becoming increasingly accepted that naïve analysis using the DiD method often overstates the degree of reliability of the key parameters (or, equivalently, understates the uncertainty), by producing estimated standard errors which are too small. The conventional ways of addressing this problem (with clustered standard errors) may not be appropriate in the particular case of the LPPC; an eventual IA should, therefore, examine the ways in which standard errors can be estimated correctly in this situation, but it should be noted that these methods are at the cutting-edge of modern econometrics and the literature has not yet come to a consensus.

7.2 Choice of datasets

The most promising dataset for an eventual IA is the WPLS, containing administrative data collected by the Department for Work and Pensions (DWP) and Her Majesty's Revenue & Customs (HMRC). This dataset has many advantages over survey data: it allows the date on which a lone parent loses eligibility for IS to be identified accurately; it is longitudinal, so IS recipients can be followed on to other benefits or into work; and it has a large sample of lone parents directly affected by the LPPC. It should be possible to use the WPLS to estimate the impact of the LPPC on key benefit and employment outcomes, provided the tax credit data in the WPLS is of sufficient quality. There are two main downsides to the WPLS compared with survey data: it does not contain any information on the lone parents who are deterred (or prevented) from claiming benefits thanks to the LPPC, and it does not measure other outcomes which may well be of interest, such as the hourly wage, hours worked per week, a comprehensive measure of family income, or any outcome for children. The impact of the LPPC on benefit, work, income and poverty outcomes could also be estimated using household surveys, such as the Labour Force Survey (LFS) and Family Resources Survey (FRS), but small sample sizes would mean it would be unlikely that one could be confident that the LPPC had had any impact at all.

7.3 Detailed implementation issues for an impact assessment

The LPPC includes additional support offered to lone parents in the 12 months before they lose entitlement to IS as a lone parent (as well as after they have entered work). For this reason, and to allow for anticipation effects in response to the impending loss of entitlement to IS, the date on which the policy first affects a lone parent should be considered to be at least 12 months before the loss of their entitlement to IS. Longer anticipation effects can be dealt with by changing the effective start date of the programme in the model. But there is a three-way trade-off between examining outcomes for a longer period of time, allowing for longer anticipation effects, and reducing the difference between the treatment and comparison groups in terms of the age of the youngest child. For some of the early phases of the roll-out, the difference in ages of the treatment and comparison groups will be large. A small difference in ages is desirable to make the common trends assumption more plausible, but a larger difference is needed in order to ensure that the comparison group is completely unaffected by the programme.

There is a 27 month roll-out period between November 2008 and January 2011 before the policy reaches a steady-state (although the current Government has announced its intention to extend the LPPC to lone parents whose youngest child is aged five and six after this date). In principle, different phases of the roll-out and the steady-state can be examined separately. There are different problems to contend with when examining the impact of the roll-out and the steady-state policy. On balance, the internal validity of estimates based on the roll-out is probably higher than those based on the steady-state, but estimates of the roll-out of the LPPC may be different from (and therefore, a poor guide to) estimates based on the steady-state. Whether these represent significant disadvantages depends on whether there is interest in understanding the impact of the roll-out in itself, rather than merely using the estimates from the roll-out as a guide to the likely impact of the steady-state policy. For example, the roll-out provides the only opportunity to estimate the impact of the LPPCs on lone parents whose youngest child is aged ten or more.

An IA could estimate the size of impacts at a particular point in time relative to when a lone parent would have lost entitlement to IS, but cannot give reasons why these impacts occurred. It would not be possible for an IA to decompose the effect of the various elements of the LPPC, nor examine displacement or substitution effects on the population as a whole. This is because of methodological barriers rather than the inadequacies of the available data.

Other policy changes affected lone parents on IS and JSA at the time Lone Parent Obligations (LPO) was introduced. Programmes affecting all lone parents equally can be ignored in an IA: the impact of these reforms would be incorporated in the common trend. Other programmes can be controlled for relatively easily by introducing additional explanatory variables into the DiD analysis which identify those lone parents who are affected. But some likely future policy changes will have more significant implications for an eventual IA: the introduction of the Work Programme will, in principle, change the nature of the policy regime for lone parents on JSA, and an extension of the LPPC to lone parents with younger children will directly affect the proposed comparison group. The most robust option would be to stop examining data when the Work Programme starts or LPPC is extended to more lone parents, but this would curtail the ability of an eventual IA to examine the medium- to long-run impact of LPPC on lone parents with children aged seven and over. An alternative approach that deals with the Work Programme would be to estimate separate impacts of LPPC before and after the Work Programme begins. Possible alternative approaches that deal with the extension of LPPC would be to use lone parents with children aged under three as the comparison group, or to flip the definition of the treatment group and comparison group, in order to use the lone parents whose youngest child is aged seven or over to estimate the impact of the LPPC on those whose youngest child is aged five and six. This will need to be explored when more details are available about the introduction and roll-out of these two policies.

Appendix A

Constructing outcome measures from the Work and Pensions Longitudinal Study

The Work and Pensions Longitudinal Study (WPLS), can tell us about whether, on any given day, an individual is:

- receiving any Department for Work and Pensions (DWP) benefit (and details of which one);
- on a DWP programme (such as New Deal for Lone Parents (NDLP));
- in work, according to P45/P46 information from Her Majesty's Revenue & Customs (HMRC) in the WPLS.

It also tells us:

- each individual's annual income (P14 earnings data);
- various pieces of information derived from HMRC's administrative data on tax credits;
- dates of meetings with personal advisers, such as Work Focused Interviews (WFIs);
- when lone parents were receiving In-Work Credit (IWC).

The information in the WPLS is held as 'spells', and so it is possible to construct measures of:

- **transitions** (i.e. job starts, benefit exits, flows between benefits);
- (job or benefit) **durations** (although estimating the impact of the lone parent policy changes (LPPCs) on durations may sometimes require the issue of dynamic selection bias to be addressed);
- **proportions of time** spent in a particular state (i.e. in work, on Jobseeker's Allowance (JSA)).

The following sections give more detail on the different sources of data.

The IS History File

The research team believes that the eventual IA would have to make use of the IS History File. This dataset gives precise details of all of a claimant's changes in circumstances during an Income Support (IS) claim. This would be needed because the information in the National Benefits Database (NBD) is not sufficiently detailed for researchers to be able to calculate the date of birth of youngest child at all points in time during an IS claim, and this information would be needed to estimate when a lone parent would lose eligibility to IS.

Although it has been used in previous impact assessments (IAs) (most recently, Brewer *et al.*, 2009)), further investigation of the quality of these data would be useful before it is used in any eventual IA.

Employment data

It is suspected that the P6/P46 information of employment in the WPLS provides an inaccurate measure of employment. There seem to be two underlying reasons:

- employment end dates appear to be badly recorded, and so some employment spells appear to be ongoing when the actual job has ended;
- not all genuine jobs appear in the WPLS (this is partly because jobs which pay less than the personal allowance need not appear in the WPLS).

For example, the IA of the Lone Parent Pilots²⁶ shows that 30 per cent of lone parents who left IS and started an IWC claim do not have a matching job start in the WPLS.

The WPLS contains administrative data on tax credit receipt, which can, in theory, also be used to determine which lone parents are in work on any given day.²⁷ Before it is used in any eventual IA, further investigation of the quality of these data would be useful.

Earnings data

The P14 data in the WPLS contains information on individuals' annual taxable income for each tax year. But its usefulness is limited by the fact that:

- it is only available with a considerable lag from the end of the tax year to which it refers.
- the data is annual, and it is very difficult to calculate weekly or monthly earnings for an employment spell starting mid-way through the tax year.

An alternative source of data on earnings might be HMRC's administrative data on tax credits, which potentially has a record of pre-tax annual family earnings for each 'entitlement sub-period' (that is to say, a period when a family's circumstances in terms of structure and hours worked remains constant within a tax year). However, this data may also not be very timely, as tax credit recipients are not obliged to keep HMRC informed of changes in annual earnings. Again, further investigation into the quality of these data would be useful.

Income from benefits and tax credits

Amounts of DWP benefits paid to lone parents (including IS, JSA, Employment and Support Allowance (ESA) and IB) can be obtained from the National Benefits Database (NBD). Benefits that are administered by local authorities, principally Housing Benefit and Council Tax Benefit, are not included in this though, which would limit the ability of an IA to calculate a measure of family income. Tax credits are administered by HMRC and administrative data is available to researchers, although at present this does not include information on tax credit amounts paid or family income.

²⁶ Brewer *et al.*, 2009.

²⁷ The 'award element history' dataset contains information on when lone parents were entitled to Working Tax Credit and the full-time premium. These identify those who are working 16 or more and 30 or more hours a week, respectively. This information could also be obtained from the 'hours worked history' dataset, although this contains self-reported information which may not be reliable.

However, since this data does not include information on certain non-taxable income and direct tax payments, it would not be possible to construct a measure of income that was similar to that used by the Government when calculating official poverty statistics. Any assessment of the impact of the LPPC on child poverty, therefore, would have to use non-administrative data sources such as the Family Resources Survey (FRS). The limitations of this dataset are discussed in Table 4.1.

Appendix B

Comparison groups for each of the phases of the roll-out

Table B.1 Comparison groups for each of the phases of the roll-out

Phase	Lone parents in treatment group	Lone parents in comparison group A	Lone parents in comparison group B	Lone parents in comparison group C	Memo: all left IS by	Memo: cumulative number of lone parents directly affected
1/1a stock	DOB of youngest child 02/03/1993 to 01/03/1995. Lose IS entitlement in three months from 02/03/2009, youngest child aged 14-16 at this time	DOB of youngest child 02/03/2002 to 25/10/2002. Aged six at this time. And/or DOB of youngest child from 02/03/2004 onwards (how far depends on desired size of comparison group). Aged under five at this time	DOB of youngest child 02/03/2005 onwards (how far depends on desired size of comparison group). Aged under four at this time	DOB of youngest child 02/03/2006 onwards (how far depends on desired size of comparison group). Aged under three at this time	02/06/2009	35,750
1a flow	DOB of youngest child 02/03/1995 to 24/11/1995. Lose IS entitlement on child's 14th birthday, from 02/03/2009 to 24/11/2009	DOB of youngest child 02/03/2002 to 25/10/2002. Aged seven at this time. And/or DOB of youngest child 02/03/2004 to 24/11/2004. Aged five at this time	DOB of youngest child 02/03/2005 to 24/11/2005. Aged four at this time	DOB of youngest child 02/03/2006 to 24/11/2006. Aged three at this time	24/11/2009	46,550
1b stock	DOB of youngest child 25/11/1995 to 05/07/1997. Lose IS entitlement in six months from 06/07/2009, youngest child aged 12-14 at this time	DOB of youngest child 06/07/2004 onwards (how far depends on desired size of comparison group). Aged under five at this time	DOB of youngest child 06/07/2005 onwards (how far depends on desired size of comparison group). Aged under four at this time	DOB of youngest child 06/07/2006 onwards (how far depends on desired size of comparison group). Aged under three at this time	06/01/2010	82,250

Continued

Table B.1 Continued

Phase	Lone parents in treatment group	Lone parents in comparison group A	Lone parents in comparison group B	Lone parents in comparison group C	Memo: all left IS by	Memo: cumulative number of lone parents directly affected
1b flow	DOB of youngest child 06/07/1997 to 24/11/1997. Lose IS entitlement on child's 12th birthday, from 06/07/2009 to 24/11/2009	DOB of youngest child 06/07/2002 to 25/10/2002. Aged seven at this time. And/ or DOB of youngest child 06/07/2004 to 24/11/2004. Aged five at this time.	DOB of youngest child 06/07/2005 to 24/11/2005. Aged four at this time	DOB of youngest child 06/07/2006 to 24/11/2006. Aged three at this time	24/11/2009	91,750
2a stock	DOB of youngest child 25/11/1997 to 31/01/1999. Lose IS entitlement in three months from 01/02/2010, youngest child aged 11-12 at this time	DOB of youngest child from 01/02/2005 onwards (how far depends on desired size of comparison group). Aged under five at this time	DOB of youngest child from 01/02/2006 onwards (how far depends on desired size of comparison group). Aged under four at this time	DOB of youngest child from 01/02/2007 onwards (how far depends on desired size of comparison group). Aged under three at this time	01/05/2010	123,650
2a flow	DOB of youngest child 01/02/1999 to 26/10/1999. Lose IS entitlement on child's 11th birthday, from 01/02/2010 to 26/10/2010	DOB of youngest child 01/02/2005 to 26/10/2005. Aged five at this time	DOB of youngest child 01/02/2006 to 26/10/2006. Aged four at this time	DOB of youngest child 01/02/2007 to 26/10/2007. Aged three at this time	26/10/2010	143,860
2b stock	DOB of youngest child 27/10/1999 to 06/06/2000. Lose IS entitlement in three months from 07/06/2010, youngest child aged 10 at this time	DOB of youngest child from 07/06/2005 onwards (how far depends on desired size of comparison group). Aged under five at this time	DOB of youngest child from 07/06/2006 onwards (how far depends on desired size of comparison group). Aged under four at this time	DOB of youngest child from 07/06/2007 onwards (how far depends on desired size of comparison group). Aged under three at this time	07/09/2010	161,370

Continued

Table B.1 Continued

Phase	Lone parents in treatment group	Lone parents in comparison group A	Lone parents in comparison group B	Lone parents in comparison group C	Memo: all left IS by	Memo: cumulative number of lone parents directly affected
2b flow	DOB of youngest child 07/06/2000 to 26/10/2000. Lose IS entitlement on child's 10th birthday between 07/06/2010 and 26/10/2010	DOB of youngest child from 07/06/2005 to 26/10/2005. Aged five at this time	DOB of youngest child from 07/06/2006 to 26/10/2006. Aged four at this time	DOB of youngest child from 07/06/2007 to 26/10/2006. Aged three at this time	26/10/2010	172,580
3a stock	DOB of youngest child 27/10/2000 to 24/10/2001. Lose IS entitlement in three months from 25/10/2010, youngest child aged 9-10 at this time	DOB of youngest child from 27/10/2005 onwards (how far depends on desired size of comparison group). Aged under five at this time	DOB of youngest child from 27/10/2006 onwards (how far depends on desired size of comparison group). Aged under four at this time	DOB of youngest child from 27/10/2007 onwards (how far depends on desired size of comparison group). Aged under three at this time	25/01/2011	206,680
3a flow	DOB of youngest child 25/10/2001 to 25/10/2002. Lose IS entitlement on child's ninth birthday, from 25/10/2010 to 25/10/2011	DOB of youngest child 25/10/2005 to 25/10/2006. Aged five at this time	DOB of youngest child 25/10/2006 to 25/10/2007. Aged four at this time	DOB of youngest child 25/10/2007 to 25/10/2008. Aged three at this time	25/10/2011	240,100
3b stock	DOB of youngest child 26/10/2002 to 02/01/2004. Lose IS entitlement in three months from 03/01/2011, youngest child aged 7-8 at this time	DOB of youngest child from 03/01/2006 onwards (how far depends on desired size of comparison group). Aged under five at this time	DOB of youngest child from 03/01/2007 onwards (how far depends on desired size of comparison group). Aged under four at this time	DOB of youngest child from 03/01/2008 onwards (how far depends on desired size of comparison group). Aged under three at this time	03/04/2011	283,600

Continued

Table B.1 Continued

Phase	Lone parents in treatment group	Lone parents in comparison group A	Lone parents in comparison group B	Lone parents in comparison group C	Memo: all left IS by	Memo: cumulative number of lone parents directly affected
3b flow/ long term policy	DOB of youngest child after 03/01/2004. Lose IS entitlement on child's seventh birthday, from 03/01/2011 onwards	DOB of youngest child from 03/01/2006 onwards. Aged five at this time	DOB of youngest child from 03/01/2007 onwards. Aged four at this time	DOB of youngest child from 03/01/2008 onwards. Aged three at this time	n/a	n/a
<p>Comparison group A is for an analysis with a one-year anticipation effect and outcomes measured one year after the end of IS.</p> <p>Comparison group B is for an analysis with a one-year anticipation effect and outcomes measured two years after the end of IS, or with a two-year anticipation effects and outcomes one year after end of IS.</p> <p>Comparison group C is for an analysis with a one-year anticipation effect and outcomes measured three years after the end of IS, or with a two-year anticipation effects and outcomes two years after end of IS.</p>						

Appendix C

Likely sample sizes for different phases of the roll-out by date

Income Support eligibility is withdrawn

Table C.1 Likely sample sizes for different phases of the roll-out, by date Income Support eligibility is withdrawn

Phase	2009												2010											
	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec		
Phase 1 stock	1,600	1,600	1,650	1,600	1,650	1,490																		
Phase 1a stock	7,880	7,880	7,880	840	840	840																		
Phase 1a flow	1,770	1,770	1,830	1,770	1,830	1,830																		
Phase 1b stock					5,070	6,050	5,850	6,050	5,850	6,050	780													
Phase 1b flow					1,740	2,070	2,010	2,070	1,610															
Phase 2a stock																								
Phase 2a flow																								
Total	11,250	11,250	11,360	4,210	11,140	12,290	9,630	9,950	8,870	6,050	780	12,150	13,450	13,010	2,340	2,260	10,750	11,110	10,040	2,110	2,340			

Phase	2010												2011											
	June	July	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar		
Phase 2a flow	2,260	2,340	2,340	2,260	1,960																			
Phase 2b stock	4,720	6,100	6,100	590																				
Phase 2b flow	1,890	2,450	2,450	2,370	2,050																			
Phase 3a stock					2,680	11,490	11,880																	
Phase 3a flow					640	2,740	2,830																	
Phase 3b stock																								
Phase 3b flow																								
Total	8,870	10,880	10,880	5,220	7,330	14,230	14,710	27,990	19,080	21,120	6,270	5,970	5,780	5,970	5,970	5,780	5,780	3,140	3,040	2,940	14,170			

Source: DWP, from administrative data.

Appendix D

A formalisation of the regression discontinuity design

This appendix builds on the notation introduced in Chapter 3 to formalise the regression discontinuity (RD) design.

Starting from the potential outcomes framework outlined in Chapter 3:

$$Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i}$$

so that:

$$Y_i = Y_{0i} + \Delta D_i$$

In this instance, assume that treatment depends on a ‘forcing variable’ x_i ; c is a cut-off above which individuals are treated and below which they are not.

$$D_i = 1[x_i \geq c]$$

The idea then is to estimate the treatment effect at the cut-off:

$$\Delta_{SRD} = E[Y_{1i} - Y_{0i} | X_i = c]$$

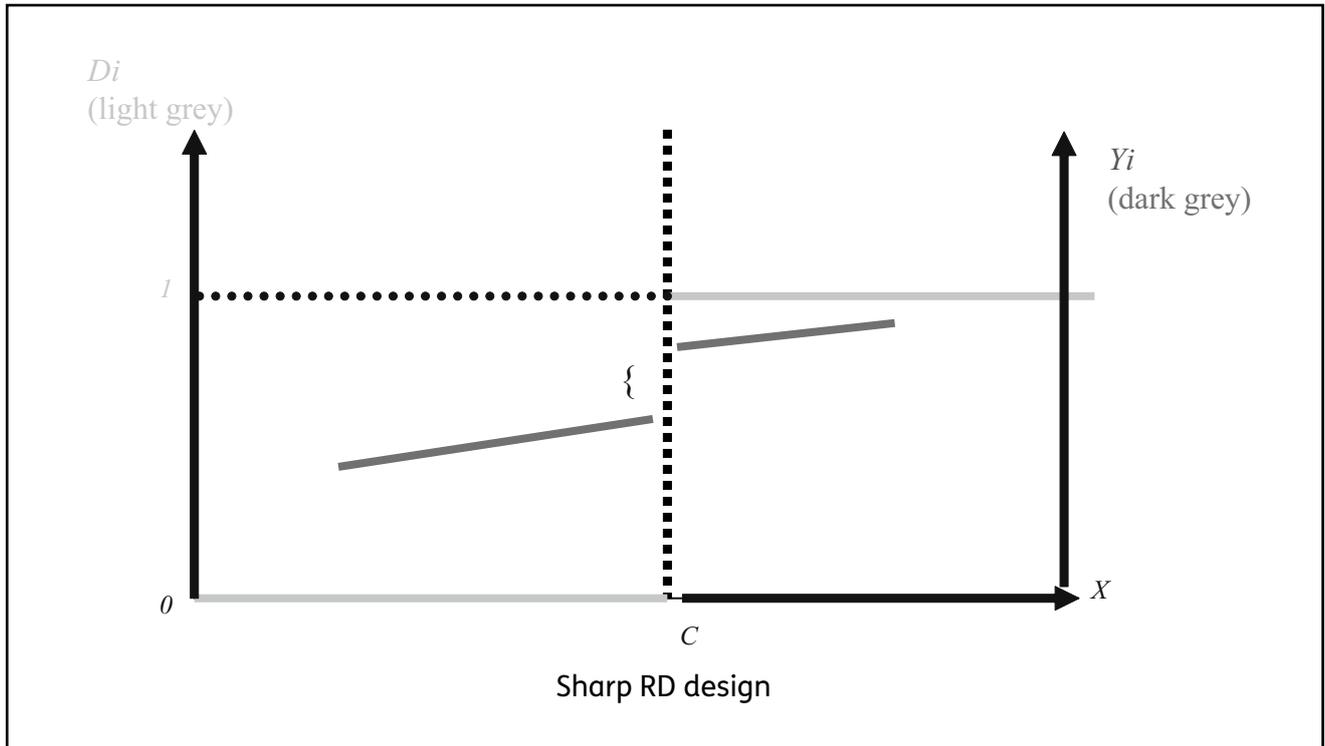
by:

$$\lim_{x \downarrow c} E[Y_i | X_i = x] - \lim_{x \uparrow c} E[Y_i | X_i = x]$$

Note that the conditional independence (unconfoundedness) assumption – which is necessary for regression and matching – will hold trivially (because given the covariates, there is no variation in treatment status). However, the ‘common support’ condition is fundamentally violated. Instead, the RD design works under a continuity assumption: that $E[Y_{1i} | X_i = x]$ and $E[Y_{0i} | X_i = x]$ are continuous in x (at least in the region of $x = c$).

Figure D.1 shows a hypothetical example of applying the RD design.

Figure D.1 An example of the regression discontinuity design



References

- Adam, S., Bozio, A. and Emmerson, C. (2009). *Can we estimate the impact of the Choices package in Pathways to Work?*. DWP Working Paper No. 60.
- Angrist, J. and Krueger, A. (1999). 'Empirical Strategies in Labor Economics'. *Handbook of Labor Economics*, Elsevier.
- Arellano, M. (1987). 'Computing Robust Standard Errors for Within-group Estimators'. *Oxford Bulletin of Economics and Statistics*, vol. 49 (4), pp. 431-434.
- Ashenfelter, O. (1978). 'Estimating the effect of training programs on earnings'. *Review of Economics and Statistics*, 60: 47-57.
- Bell, B., Blundell, R. and van Reenen, J. (1999). 'Getting the Unemployed Back to Work: An Evaluation of the New Deal Proposals'. *International Tax and Public Finance*, 6(3): 339-360.
- Bell, R.M. and McCaffrey, D.F. (2002). 'Bias Reduction in Standard Errors for Linear Regression with Multi-Stage Samples'. *Survey Methodology*, vol. 28 (2), pp. 169-181.
- Bertrand, M., Duflo, E. and Mullainathan, S. (2004). 'How much should we trust difference-in-difference estimates?'. *Quarterly Journal of Economics*, vol. 119 (1), pp. 249-275.
- Bewley, H., Dorsett, R. and Ratto, M. (2008). *Evidence on the effect of Pathways to Work on existing claimants*. DWP Research Report No. 488. Leeds: CDS.
- Blundell, R. and Costa Dias, M. (2000). 'Evaluation methods for non-experimental data'. *Fiscal Studies*, 21(4), pp. 427-468.
- Blundell, R. and Costa Dias, M. (2008). 'Alternative approaches to evaluation in empirical microeconomics', cemmap working paper CWP26/08. <http://www.cemmap.ac.uk/wps/cwp2608.pdf>
- Blundell, R. and Costa Dias, M. (2008). 'Alternative Approaches to Evaluation in Empirical Microeconomics', cemmap working paper CWP26/08.
- Blundell, R., Costa Dias, M., Meghir, C. and van Reenen, J. (2004). 'Evaluating the Employment Impact of a Mandatory Job Search Programme'. *Journal of the European Economic Association*, vol. 2(4), pp. 569-606.
- Brewer, M., Browne, J., Chowdry, H. and Crawford, C. (2009). *The lone parent pilots after 24-36 months: the final impact assessment of In-Work Credit, Work Search Premium, Extended Schools Childcare, Quarterly Work Focused Interviews and New Deal Plus for Lone Parents*. DWP Research Report No. 606. Leeds: CDS.
- Brewer, M., Browne, J., Crawford, C. and Knight, G. (2007). *The lone parent pilots after 12 to 24 months: an impact assessment of In-Work Credit, Work Search Premium, Extended Schools Childcare, Quarterly Work Focused Interviews and New Deal Plus for Lone Parents*. DWP Research Report No. 415. Leeds: CDS.
- Brewer, M., Muriel, A., Phillips, D. and Sibieta, L. (2009). 'Poverty and Inequality in the UK: 2009'. IFS Commentary 109. London: IFS.
- Bloom, H. (1984). 'Accounting for no-shows'. *Evaluation Review*, 8(2):224-246.

- Cameron, C., Gelbach, J. and Miller, D.L. (2008). 'Bootstrap-Based Improvements for Inference with Clustered Errors'. *The Review of Economics and Statistics*, vol. 90, pp. 414-427.
- Cameron, C., Gelbach, J. and Miller, D.L. (2010). 'Robust Inference with Multi-Way Clustering', forthcoming in JBES. Previous version available at <http://www.nber.org/papers/T0327>
- Card, D. and Hyslop, D. (2005). 'Estimating the effects of a Time-Limited Earnings Subsidy for Welfare Leavers'. *Econometrica*, 73(6): 1723-1770.
- Chay, K., McEwan, P. and Urquiola, M. (2005). 'The Central Role of Noise in Evaluating Interventions That Use Test Scores to Rank Schools'. *The American Economic Review*, 95(4): 1237-1258.
- Cobb-Clark, D. and Crossley, T.F. (2003). 'Econometrics for Evaluations: An Introduction to Recent Developments'. *Economic Record*, 79(247):493-513. Sections I and II.
- De Giorgi, G. (2005). 'The New Deal for Young People: Five Years On'. *Fiscal Studies*, vol. 26, no. 3, pp. 371-383.
- Dolton, P. and O'Neill, D. (xxxx). 'Unemployment duration and the Restart effect: some experimental evidence'. *The Economic Journal*, 106 (435): 387-400.
- Eissa, N. and Liebman, J.B. (1996). 'Labor Supply Response to the Earned Income Tax Credit'. *The Quarterly Journal of Economics*, 111(2): 605-637.
- Finn, D. and Gloster, R. (2010). *Lone Parent Obligations: A review of recent evidence on the work-related requirements within the benefit systems of different countries*. DWP Research Report No. 632. Leeds: CDS
- Ham, J. and Lalonde, R. (1996). 'The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training.' *Econometrica*.
- Holland, P.W. (1986). 'Statistics and causal inference'. *Journal of the American Statistical Association*, 8(396): 945-960.
- Imbens, G. and Lemieux, T. (2007). 'Regression Discontinuity Designs: A guide to Practice'. Forthcoming, *Journal of Econometrics* (see also NBER TWP 337. www.nber.org/papers/t0337).
- Mulheirn, I. and Pisani, M. (2008). 'Working Tax Credit and labour supply'. HM Treasury Economic Working Paper No. 3.
- Moulton, B. (1986). 'Random Group Effects and the Precision of Regression Estimates'. *Journal of Econometrics*, vol. 32, pp. 385-397.
- Pronzato, C. and Mogstat, M. (2008). 'Are Lone Mothers Responsive to Policy Changes?'. ISER Working Paper Series No. 2008-15.
- Rubin, D.B. (1974). 'Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies'. *Journal of Educational Psychology*, 66: 688-701.
- Thistlethwaite, D. and Campbell, D. (1960). 'Regression-Discontinuity Analysis: An alternative to the ex post facto experiment'. *Journal of Educational Psychology*, 51, 309-317.
- Wooldridge, J. (2007). 'Difference-in-difference estimation', http://www.nber.org/WNE/lect_10_diffindiffs.pdf

Lone Parent Obligations (LPO) were introduced in November 2008. Since then, most lone parents with a youngest child aged 12 or over lost entitlement to Income Support solely on the grounds of being a lone parent. The age of the youngest child was lowered to ten in October 2009 and to seven from October 2010.

This report presents findings from a feasibility study for an impact assessment of LPO. It provides evidence on the suitability of a difference-in-difference methodology and the use of lone parents with younger children as a comparison group. The report also presents the results of analysis carried out to test whether the 'common trends' assumption, needed for any difference-in-difference estimate to be valid, held between the period from 2001 and 2007.

An impact assessment would form part of a programme of evaluation research, using a mixed methods approach, to assess the effects of LPO. The evaluation includes in-depth interviews with customers and staff, a large-scale survey of customers, as well as analysis of in-house and other data sources. It is being carried out by a consortium of independent research organisations, led by the Centre for Economic and Social Inclusion.

If you would like to know more about DWP research, please contact:
Paul Noakes, Commercial Support and Knowledge Management Team,
Work and Welfare Central Analysis Division, 3rd Floor, Caxton House,
Tothill Street, London SW1H 9NA.
<http://research.dwp.gov.uk/asd/asd5/rrs-index.asp>

DWP Department for
Work and Pensions

Published by the
Department for Work and Pensions
November 2010
www.dwp.gov.uk
Working paper no. 92
ISBN 978-1-84712-891-1