

In-work poverty: a systematic review

by Janice Tripney, Mark Newman, Mukdarut Bangpan,
Amelia Hempel-Jorgensen, Marian Mackintosh,
Helen Tucker and Jennifer Sinclair

Department for Work and Pensions

Research Report No 549

In-work poverty: a systematic review

**Janice Tripney, Mark Newman, Mukdarut Bangpan,
Amelia Hempel-Jorgensen, Marian Mackintosh, Helen Tucker and
Jennifer Sinclair**

A report of research carried out by the Evidence for Policy and Practice
Information and Co-ordinating Centre on behalf of the Department for Work
and Pensions

© Crown Copyright 2009. Published for the Department for Work and Pensions under licence from the Controller of Her Majesty's Stationery Office.

Application for reproduction should be made in writing to The Copyright Unit, Her Majesty's Stationery Office, St Clements House, 2-16 Colegate, Norwich NR3 1BQ.

First Published 2009.

ISBN 978 1 84712 466 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..... | vii |
| Summary | 1 |
| 1 Background | 7 |
| 1.1 Introduction..... | 7 |
| 1.2 The contribution of systematic reviews | 9 |
| 1.3 Type of review | 10 |
| 1.4 User involvement..... | 10 |
| 1.5 Aims and review questions | 10 |
| 1.6 Barriers and facilitators: a conceptual framework..... | 11 |
| 2 Mapping exercise: methodology | 13 |
| 2.1 Overview | 13 |
| 2.2 Defining relevant studies: inclusion and exclusion criteria | 13 |
| 2.3 Identification of potential studies: search strategy | 16 |
| 2.4 Screening studies: applying inclusion and exclusion criteria | 16 |
| 2.5 Characterising included studies..... | 17 |
| 2.6 Identifying and describing studies: quality-assurance process..... | 17 |
| 3 Systematic map results..... | 19 |
| 3.1 Overview | 19 |
| 3.2 Identifying studies | 19 |

| | | |
|-------|---|----|
| 3.3 | Summary of findings from the systematic map | 22 |
| 3.3.1 | <i>Characteristics of factors studies (201 studies)</i> | 22 |
| 3.3.2 | <i>Characteristics of intervention studies (285 studies)</i> | 22 |
| 4 | Synthesis methodology | 27 |
| 4.1 | Overview | 27 |
| 4.2 | Moving from broad characterisation (mapping) to synthesis | 27 |
| 4.3 | Detailed description of studies in the synthesis | 28 |
| 4.4 | Assessing quality of studies and weight of evidence for the review question | 29 |
| 4.5 | Synthesis: quality-assurance process | 31 |
| 4.6 | Synthesis of evidence | 31 |
| 4.7 | Deriving conclusions and implications | 32 |
| 5 | Synthesis results | 33 |
| 5.1 | Overview | 33 |
| 5.2 | Selecting studies for synthesis | 33 |
| 5.3 | Characteristics of studies included in the synthesis | 34 |
| 5.3.1 | <i>Population</i> | 34 |
| 5.3.2 | <i>Interventions</i> | 36 |
| 5.3.3 | <i>Outcomes</i> | 39 |
| 5.3.4 | <i>Evaluation design</i> | 40 |
| 5.3.5 | <i>Summaries of the characteristics of each study</i> | 41 |
| 5.4 | Weight of Evidence | 52 |
| 5.5 | Synthesis of evidence | 53 |
| 5.5.1 | <i>Introduction</i> | 53 |
| 5.5.2 | <i>Synthesis of results: wages and household income</i> | 53 |
| 5.5.3 | <i>Synthesis of results: employment participation</i> | 55 |
| 5.5.4 | <i>Synthesis of results: working hours</i> | 57 |
| 5.6 | Summary of synthesis results | 59 |

| | | |
|-------|--|----|
| 6 | Conclusions and implications | 61 |
| 6.1 | Discussion..... | 61 |
| 6.1.1 | <i>Strengths and limitations of this systematic review</i> | 65 |
| 6.2 | Implications | 67 |
| 6.2.1 | <i>Policy</i> | 67 |
| 6.2.2 | <i>Research</i> | 68 |
| | References | 71 |

List of tables

| | | |
|-----------|---|----|
| Table 1.1 | Causes and solutions to in-work poverty..... | 8 |
| Table 2.1 | Inclusion/exclusion criteria (Stage 1: Mapping) | 14 |
| Table 3.1 | Type of intervention (305 interventions* in 285 studies)..... | 25 |
| Table 5.1 | Outcomes measured (18 studies) | 40 |
| Table 5.2 | WoE judgements | 52 |
| Table 5.3 | Summary direction of effect: changes in household income | 54 |
| Table 5.4 | Summary directions of effect: changes in wages | 54 |
| Table 5.5 | Summary directions of effect: changes in employment participation | 56 |
| Table 5.6 | Summary directions of effect: changes in working hours | 58 |

List of figures

| | | |
|------------|--|----|
| Figure 1.1 | The contribution of systematic reviews..... | 9 |
| Figure 1.2 | Conceptual diagram of map coverage | 11 |
| Figure 3.1 | Selection of studies for the systematic map..... | 21 |
| Figure 3.2 | Evaluations by study country (285 studies) | 23 |
| Figure 3.3 | Outcomes measured (285 studies) | 26 |
| Figure 5.1 | Selection of studies for the synthesis | 34 |

Acknowledgements

This review was commissioned by the Department for Work and Pensions (DWP).

The Review Group would like to thank members of the Advisory Group for their contributions, including Jane Francis, Claire Murphy, Karen Elsmore, Harriet Cameron, and Richard Given (DWP) and Kevin Dodds (Her Majesty's Treasury).

Special thanks are also due to Jane Francis the project manager at DWP (Strategy Unit) for her support throughout the project.

Catherine Aicken provided support with document retrieval.

The opinions expressed in this publication are not necessarily those of the Evidence for Policy Practice Information and Co-ordinating Centre (EPPI-Centre) or the funders. Responsibility for the views expressed remains solely with the authors.

The Authors

Janice Tripney is a Research Officer at the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre), Social Science Research Unit, Institute of Education, University of London.

Dr. Mark Newman is a Senior Research Officer and Associate Director of the EPPI-Centre, Social Science Research Unit, Institute of Education, University of London, where he manages the EPPI-Centre's programme of reviews in education and social policy.

Mukdarut Bangpan is a Research Officer at the EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Amelia Hempel-Jorgensen is a Research Officer at the EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Marian Mackintosh is a Research Associate at the EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Helen Tucker is a Research Associate at the EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Dr. Jennifer Sinclair is a Research Officer at the EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Abbreviations

| | |
|-------------|---|
| AHC | After Housing Costs |
| BHC | Before Housing Costs |
| BHPS | British Household Panel Survey |
| CTC | Child Tax Credit |
| CPS | Current Population Survey |
| DWP | Department for Work and Pensions |
| EITC | Earned Income Tax Credit |
| EPPI-Centre | Evidence for Policy and Practice Information and Co-ordinating Centre |
| FACS | Family and Children Study |
| FC | Family Credit |
| FES | Family Expenditure Survey |
| FRS | Family Resources Survey |
| HEK | Swedish Income Distribution Survey |
| HMT | Her Majesty's Treasury |
| HBAI | Household Below Average Income |
| LFS | Labour Force Survey |
| MDRC | Manpower Demonstration Research Corporation |
| MITTS | Melbourne Institute Tax and Transfer Simulator |

| | |
|--------|---|
| NVQ | National Vocational Qualification |
| OBRA93 | Omnibus Budget Reconciliation Act of 1993 |
| RCT | Randomised Controlled Trial |
| SIHC | Survey of Income and Housing Costs |
| SOEP | German Socio-Economic Panel |
| UK | United Kingdom |
| USA | United States of America |
| WoE | Weight of evidence |
| WFTC | Working Families' Tax Credit |
| WTC | Working Tax Credit |

Summary

Introduction

Government strategy is committed to halving child poverty in the United Kingdom (UK) by 2010 and eradicating it by 2020. A range of policy initiatives have made encouraging progress, with 600,000 children lifted out of poverty in the last ten years. Yet, the total number of poor children living in **working** households has stayed the same at 1.4 million (nearly half of all poor children). If the ambitious targets are to be met, it will be important that the next phase of the strategy for tackling child poverty prioritises these families.

The systematic review detailed in this report was commissioned by the Department for Work and Pensions (DWP). Systematic reviewing is a specialist technique which employs standardised and explicit methods to minimise the risk of drawing the wrong or misleading conclusions from a body of evidence. Explicit reporting of how the review was conducted allows others to assess the validity of its findings.

User involvement

An Advisory Group was set up to inform the development of the review, thereby increasing its relevance to policy.

Aims and review questions

The overall aims of the research were to produce findings to inform policy in relation to the reduction of in-work poverty among families with dependent children and highlight gaps in the research in this area that might be filled by future research.

There were two stages to the systematic review, and therefore two research questions:

- 1 What is the nature and extent of the research that has been undertaken on the barriers to, and facilitators of, reducing in-work poverty in families with dependent children?

- 2 What is the effectiveness of interventions with the potential to reduce in-work poverty in couple families with dependent children?

Review methodology

The systematic review followed standardised systematic review processes developed by the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre). Quality-assurance mechanisms were used throughout the review process, to ensure rigour and consistency between members of the review team.

The systematic review was conducted in two stages: mapping and synthesis.

At stage one of the review (mapping), the potentially relevant studies identified by the searching process were screened against ten criteria designed to facilitate the systematic selection of only the relevant studies. For example, the criteria for inclusion in the map were set to identify studies that considered poor families (one- or two-parent) with dependent children. Our definition of poor families included those described as:

- low income;
- low paid;
- low skilled;
- lone parents;
- recipients of social welfare benefits or in-work financial support and/or part-time employees.

Included studies were coded for contextual and methodological information, using standardised EPPI-Centre frameworks and coding questions developed specifically for the review. The map results are presented in the form of a descriptive analysis of the available research literature.

The information contained in the map provided the basis for developing the more narrowly focused policy question to be answered by the second stage of the review (synthesis). A second set of selection criteria was applied to the map studies, and the references of those that were capable of answering the synthesis question were scanned, to identify further relevant research. Included studies were analysed in depth and their quality appraised. Three outcome categories provided the framework for comparing studies in the synthesis. The synthesis focused on searching for patterns of similarity or difference in the direction of effects (positive or negative) in each outcome category.

Identifying studies (mapping)

Altogether, the literature searches produced 9,144 records. These were narrowed down to a total of 594 reports of 439 separate studies which were included in the map. Of these, 201 studies provided non-evaluative evidence on factors associated with in-work poverty in families with dependent children (factors studies). A total of 285 studies provided evaluative evidence about the design and impact of interventions to reduce in-work poverty among such families (intervention studies). These groups are not mutually exclusive.

Key findings (mapping)

Factors studies: Reflecting our inclusion criteria, all 201 factors studies were conducted in the UK. Of those identified, 12 were cross-national comparison studies. Studies used a range of study methods and designs (including 67 cross-sectional and 75 'views' studies).

Seventeen of the 201 factors studies sampled coupled parents only. A further 85 studies included both lone and two-parent families. Only one of the 63 studies that purposively sampled lone parents was focused solely on lone fathers. A majority of the 201 studies included both individuals who were in paid employment, and those who were not (68 per cent). No studies included full-time workers only.

Intervention studies: A total of 285 studies evaluated an intervention. Nearly one-third of these (31 per cent) were conducted in the UK (88 studies). In total, there were 45 Randomised Controlled Trials (RCTs), only one of which was conducted in the UK.

The study population in 13 of the intervention studies that we identified contained two-parent families only. A further 115 studies included both lone and coupled parents in their samples. In 162 studies (57 per cent) the study population consisted of some participants who were in-work and others who were not. No studies included full-time workers only.

Studies evaluated a broad range of intervention types. A total of 305 interventions, or components of interventions, were evaluated in the 285 studies:

- Employment support (N = 118).
- Education (N = 18).
- Financial (N = 144).
- Social/medical (N = 7).
- Resource provision (N = 17).
- Family-friendly initiatives (N = 1).

The financial interventions were further categorised, based on whether they 'push' (N=36) or 'pull' (N=113) people not doing paid work into the labour market. A few interventions combined both these strategies within one initiative.

Reflecting the multi-component nature of the interventions, many sought to influence behaviour by addressing more than one barrier to employment: for example, employment support services that offered job searching advice and a 'back-to-work' financial bonus on entering employment.

There is a relatively large body of research providing evidence on post-employment initiatives (142 studies). A total of 65 studies evaluated interventions that provide a continuum of support, addressing both pre-employment and in-work support needs.

Studies measured a range of outcomes relevant to this review, including employment participation (194 studies), wages/earnings/income (136 studies), and hours worked (54 studies).

Identifying studies (synthesis)

After application of the second set of selection criteria to the 439 map studies, 12 studies remained. Scanning their references lists led to a further five relevant studies. An additional study came to our attention through personal communication. A total of 18 evaluation studies were identified for inclusion in the synthesis.

Key findings (synthesis)

Of the 18 included studies, 12 were conducted in the UK, three in the United States of America (USA), and one each in Germany, Sweden and Australia. All were retrospective quantitative evaluations, with all but one utilising statistical modelling of varying degrees of complexity.

Four studies were judged to provide an overall medium weight of evidence (WoE) and the remaining 14 studies were graded low/medium overall.

Reflecting our inclusion criteria, the study populations of all 18 studies contained working two-parent families. Three studies focused solely on couple families with dependent children. Most study samples (89 per cent) also contained workless families.

All 18 studies evaluated financial interventions available only to those in employment. No evaluations of other kinds of interventions for in-work poor couple families with dependent children were identified.

All studies evaluated the impact of changes (real or hypothetical) to the tax and/or benefit system. Despite differences in specific mechanisms, the main focus of the

interventions was on providing greater financial return on paid employment, than the system to which they were compared.

Fifteen studies evaluated tax credits schemes, in either the UK or the USA (83 per cent overall). Two of the three remaining studies also evaluated initiatives that were targeted at the working poor. One study evaluated a reform available to all workers.

Although we set our selection criteria to identify studies that considered transitions into and out of income poverty, none of the studies included for synthesis measured changes in income using thresholds derived from percentages of median income for the whole population (i.e. poverty lines).

All 18 studies measured alternative (proxy) outcomes which, in our view, have the potential to reduce in-work poverty:

- wages/household income;
- employment participation;
- working hours.

This feature of the review reflects the nature of the studies themselves, rather than a deliberate decision to focus on these particular outcomes. Similarly, the inclusion of only financial interventions in the synthesis was also determined by the nature of the evidence base.

One study measured the impact on household income following behavioural responses to a financial intervention. An increase in income was found. Two studies measured the impact on wage growth and reported different findings.

Eight studies informed us about the effect of financial interventions on the overall employment participation of second earners (i.e. taking into consideration both male and female second earners). Six studies found a net reduction in participation and two studies found a net increase.

Eight studies informed us about the effect of financial interventions on the overall working hours of couple families. Six studies found a net decrease in working hours and two studies found a net positive effect.

Conclusions and implications

Since this review relies on the use of proxy outcomes for poverty, clearly there are limitations to how far the selected studies can illuminate us about effective solutions to in-work poverty.

Our interpretation of the results of the synthesis is that the financial interventions that have been evaluated do not appear, on average, to have resulted either in attracting more potential second earners into work or encouraging two-parent

families to work more hours. Given the overall quality of the evidence, however, and the lack of complete consistency between the primary study findings, this interpretation is to some degree speculative and should be treated with caution.

What we can reliably say on the basis of these studies is that they do not provide a conclusive answer to the policy question about effective solutions to the problem of in-work poverty. This is an important finding in itself. Therefore, our recommendations for policy are concentrated on the urgent need for further research that can answer this question.

Although there is a substantial body of empirical research evidence on factors appertaining to in-work poverty and interventions to reduce it, a relatively limited amount of evaluative research has focused on working two-parent families. This needs to be addressed.

The absence of evaluations of anything other than financial interventions would suggest that, if non-financial interventions are in place, there is a need to commission rigorous evaluations of them. This also applies to other types of financial interventions that were not identified.

Overall, we would strongly recommend that resources are devoted to further quantitative work on the relationship between financial interventions and changes to household income that take full account of behavioural responses to reforms and the interaction between different means-tested benefits. This should be supplemented by rigorous qualitative analysis to unpack some of the more subtle relationships that quantitative analysis cannot detect.

Prospective randomised experiments are widely viewed as the 'gold standard' for proving the efficacy of interventions. While acknowledging that there are numerous practical, methodological and ethical issues that need to be addressed for such study designs to be used in this field, it is important that policymakers consider evaluation before rolling out widespread changes to systems.

Researchers need to engage in finding solutions to the many difficulties posed by evaluation practice in this field.

1 Background

1.1 Introduction

The Government has set ambitious targets to halve the number of children in child poverty by 2010 and then eradicate child poverty altogether by 2020. A range of new policy initiatives, including both employment measures and cash transfers, have made encouraging progress. The latest figures, published in June 2008 as part of the annual Households Below Average Income (HBAI) series, indicate that, from 1998/99 to 2006/07, the number of children in relative poverty fell by 600,000 'before housing costs' (BHC) and 500,000 'after housing costs' (AHC) (DWP, 2008). However, the Government is still considerably short of its targets, having missed the 2004/05 milestone by a significant margin. Also of concern, the HBAI figures show that, between 2005/06 and 2006/07, the number of children in relative poverty rose by 100,000, both before and after housing costs are taken into account, to 2.9 million (BHC) and 3.9 million (AHC) respectively. Particularly in light of this recent increase, meeting the 2010 and 2020 targets will be challenging.

Working poor families are of particular interest in the current policy climate. A central feature of the Government's child poverty strategy has been the message that presents employment as the most effective and sustainable route out of poverty. For many families, however, work does not provide an income sufficient to lift them out of poverty. The 2008 HBAI statistics indicate that there are 1.4 million poor children living in working households, the same number of children as in 1997. Furthermore, poor children are increasingly likely to come from a working family. Today, 50 per cent of all poor children live in families where at least one parent works. This compares with 40 per cent ten years ago. Almost 80 per cent of all working poor families with children are two-parent families (DWP, 2008). From the point of view of meeting the Government's child poverty targets, working poverty, particularly among couple families with children, is clearly a substantial problem.

In June 2006, the DWP commissioned independent policy advisor Lisa Harker to carry out an independent review of their child poverty strategy. *'Delivering on*

Child Poverty: What would it take? was published later that year (Harker, 2006). The research identified three broad causes of in-work poverty (reproduced verbatim in Table 1.1).

Table 1.1 Causes and solutions to in-work poverty

| | |
|--|--|
| Low pay | Requires measures to improve wage levels – via the minimum wage, sector pay agreements or a voluntary approach. Better support for parents to advance in-work, so that low-paid workers do not remain trapped on low pay. Working Tax Credit is sufficient to lift some in-work couple families out of poverty – more help is required via the tax credits system. |
| Families relying on one earner | There is a financial disincentive for some second earners to enter work. Second earners need help with preparing for and moving into work. |
| Single/dual earners not working enough hours | Single/dual earners need support to increase their hours and/or progress in-work. |

A key message of this review was that finding the right balance of employment support and tax credit/benefit support will maximise the chances of meeting the 2010 and 2020 targets. The argument was made that continuing with current policy as it stands will not achieve the necessary reduction in poverty, and further reforms are required if significant progress is to be made. The tentative solutions proposed in the review appear to be largely based on existing initiatives. For example, the New Deal programme for lone parents has, it was argued, had a positive impact on getting this group into work, and thus its principles could also be extended to all parents, in the form of a New Deal for Parents. Loosely outlined reforms of various forms of financial in-work support, including those that lie outside DWP responsibility, were also discussed. In brief, the review suggested more personalised support, tailored to the needs of parents, with more flexibility between programmes. To reach working poor families not in contact with existing initiatives, the review advocated the use of largely unused channels, such as Children's Centres. Many of the recommendations in the review focused on the need for more evidence upon which to base the design of future measures to reduce in-work poverty in families with children.

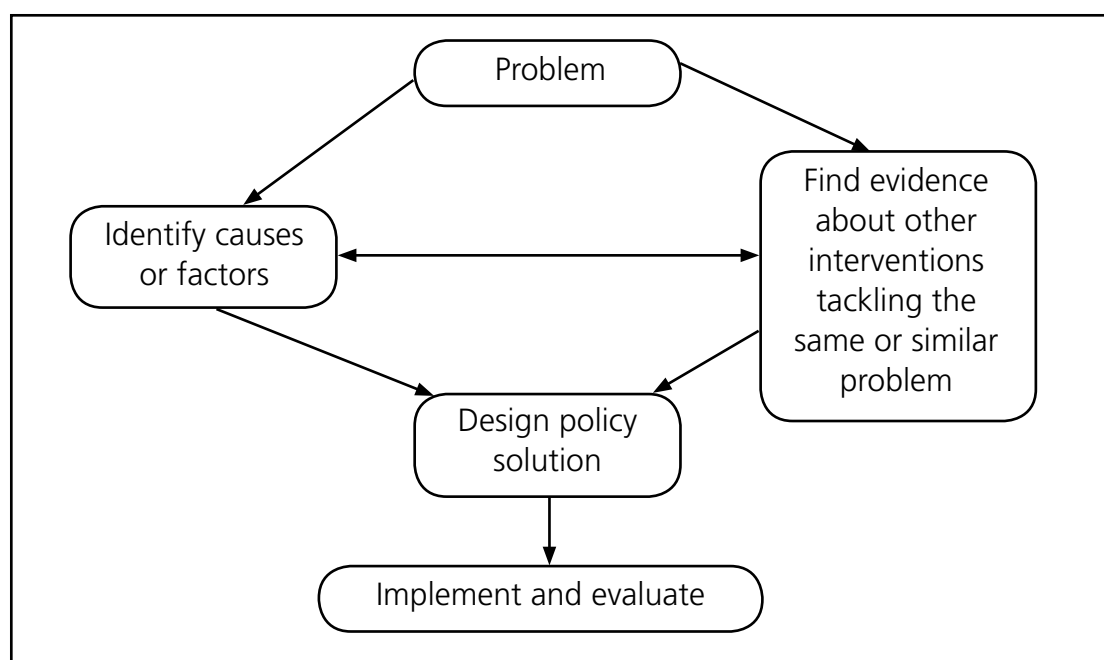
The DWP is keen to ensure that, where possible, the evolution and development of further initiatives should continue to be informed by the results of high quality research. The research evidence base on poverty has grown steadily since its 'rediscovery' in the mid-1960s and the re-opening of the poverty debate (Abel-Smith and Townsend, 1965). There are a number of recent surveys of literature in this broad area, including Gradus and Julsing (2001), Millar and Ridge (2001), Brewer and Shephard (2004), and Kemp *et al.* (2004). To date, in-work poverty has not been the topic of a systematic review.

The EPPI-Centre was commissioned by DWP to undertake the systematic review of research on in-work poverty among families with dependent children detailed in this report. This review will support their own work in this area, to identify what needs to be addressed and, to some extent, how they might go about it.

1.2 The contribution of systematic reviews

Research evidence should play a role in informing policy processes. The general simplified conceptual model of the policy-making process to which the systematic review can contribute evidence is illustrated in Figure 1.1. In essence, this is a problem-solving framework in which policy is viewed as providing a solution in the form of an intervention. A systematic review can contribute evidence to help identify the causes or factors which create or affect a particular problem, and thus which will need to be addressed in the design of any policy solution; and/or it can contribute evidence about the effectiveness, design and implementation of interventions which have been developed to address the same or similar problems.

Figure 1.1 The contribution of systematic reviews



Systematic reviewing is a specialist technique which uses standardised and explicit methods. These methods are employed in order to minimise the risk of drawing the wrong or misleading conclusion from a body of evidence. Explicit reporting of how the review was conducted allows others to assess potential sources of bias in the review and thus the validity of its findings. Methods of communicating review findings emphasise careful presentation, in order to avoid misrepresentation, and measures to increase accessibility, particularly for those who have to make policy decisions.

1.3 Type of review

The systematic review described in this report was carried out in two stages: Stage 1, a mapping exercise, followed by Stage 2, an in-depth review (or synthesis) of a subset of studies.

At the mapping stage, the initial question was broad, searching was extensive, and contextual and methodological information was collected about the studies in the map. The map results are presented in the form of a descriptive analysis of the available research literature on in-work poverty among families with dependent children (within certain pre-specified parameters). By identifying and describing this research, including where there are gaps in research activity, the mapping exercise has produced a useful product in its own right. A database of studies in the field has been created, facilitating sustainability and the potential for development of a cumulative knowledge base.

The information presented in the systematic map also provided the basis for identifying a more focused policy question to be answered by the second stage of the review (synthesis). A key benefit of a two-stage commissioning process is its provision for users of the review to be closely involved in this process, thereby helping to ensure that systematic reviews are relevant and useful.

At the second stage of the review, the synthesis focused on a discrete sub-group of studies from the systematic map which were capable of addressing the second review question. Employing strategies designed to minimise bias and error, the synthesis brings together the relevant findings of the individual primary research studies. In common with other systematic reviews, the methods of synthesis reflect the types of studies included in the review, the detail and quality of reporting, and their heterogeneity. The new findings generated by the synthesis are used to draw conclusions and identify implications for policy and research.

1.4 User involvement

An Advisory Group was set up to inform the scope and development of the review, and to increase its relevance to policy. Group membership comprised a variety of researchers and policy specialists representing a number of streams within the DWP and Her Majesty's Treasury (HMT) (listed in the acknowledgements section of the report).

1.5 Aims and review questions

The overall purpose of the research is to produce findings that will inform policy in relation to the reduction of in-work poverty among families with dependent children and highlight gaps in the research in this topic area.

Each stage of the review had a key research question:

- 1 **Mapping stage:** What is the nature and extent of the research that has been undertaken on the barriers to, and facilitators of, reducing in-work poverty in families with dependent children?
- 2 **Synthesis:** What is the effectiveness of interventions with the potential to reduce in-work poverty in couple families with dependent children?

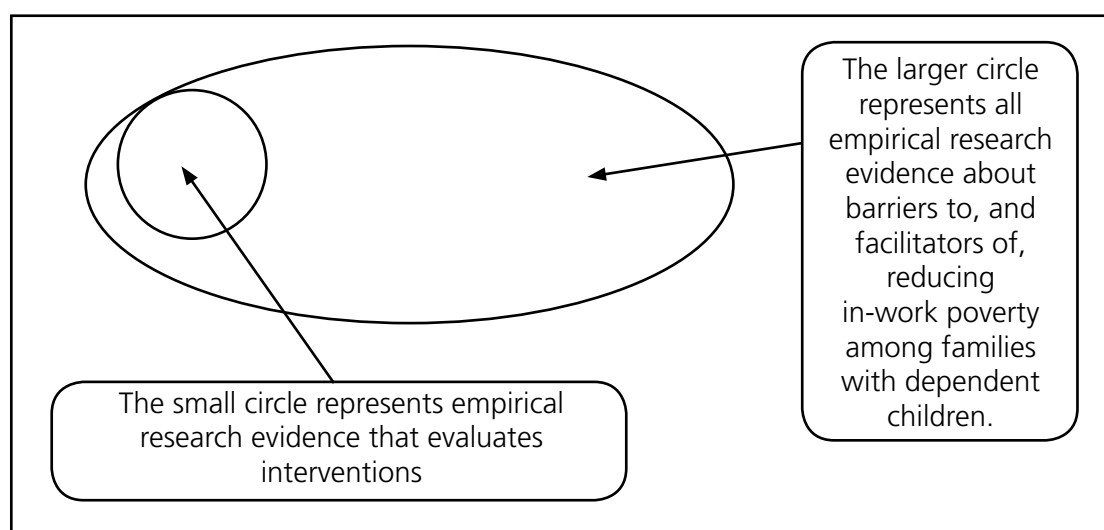
These two questions, both produced in consultation between the DWP and the EPPI-Centre team, are answered separately by the two stages of the review (mapping and synthesis).

The mapping stage of the review aims to identify, collate and describe the nature and extent of the available literature that is capable of answering the broad first review question, thereby providing the basis for informed discussion and decision-making about the direction of the synthesis. The main aim of the synthesis is to help policymakers identify interventions that reduce in-work poverty in couple families with dependent children, and on which the design of future initiatives can be based.

1.6 Barriers and facilitators: a conceptual framework

The first review question was necessarily broad. It drove the initial mapping stage of the review and provided the conceptual basis for the systematic map of research in the topic area. For the purposes of the map, 'barriers and facilitators' have been conceptualised as encompassing interventions. The conceptual content of the map is illustrated in Figure 1.2.

Figure 1.2 Conceptual diagram of map coverage



The literature answering the second review question forms a subset of the smaller circle (not represented in the diagram).

2 Mapping exercise: methodology

2.1 Overview

This chapter describes the methods used in the first stage of the review: the mapping of research relevant to the barriers to, and facilitators of, reducing in-work poverty among families with dependent children. The mapping exercise was conducted in three stages:

- defining relevant studies, involving the development of inclusion and exclusion criteria (see Section 2.2);
- identification of relevant studies, including the searching and screening processes (see Sections 2.3 and 2.4); and
- describing, or characterising, these studies (see Section 2.5).

The quality assurance process is outlined in Section 2.6.

2.2 Defining relevant studies: inclusion and exclusion criteria

The aim of the literature search was to locate a wide variety of research that was relevant for answering the broad review question (research question no.1):

What is the empirical research evidence about the barriers to, and facilitators of, reducing in-work poverty among families with dependent children?

This research question was designed to be broad. For the purpose of conducting the map, it was broken down into two inter-related sub-questions which helped further define the field of enquiry. These are based on the conceptualisation of map coverage demonstrated in Figure 1.2 in Chapter 1):

- *What is the nature and extent of the non-evaluative research that has been undertaken on the barriers to, and facilitators of, reducing in-work poverty among families with dependent children?*

- *What is the nature and extent of the evaluative research that has been undertaken on the design and impact of interventions with the potential to reduce in-work poverty among families with dependent children?*

To identify relevant studies for inclusion in the map, the following inclusion and exclusion criteria were developed and agreed through lengthy discussions, including consultations with members of the Advisory Group. These criteria defined the boundaries of the map.

Table 2.1 Inclusion/exclusion criteria (Stage 1: Mapping)

| Number | Inclusion criteria | Exclusion criteria |
|--------|--|---|
| 1 | Study citations provide sufficient information to be immediately screened or for retrieval for full-text screening. | An incomplete reference. |
| 2 | Studies published in English. | Not published in English. |
| 3 | Studies published from 1992 onwards. | Published before 1992. |
| 4 | Study population (sample) includes parents (lone or partnered) of dependent children. Dependent children are defined by age or circumstances (e.g. in education, disabled adults dependent on their parents). Studies that include other stakeholders are within the scope of the map, conditional on references to in-work poverty in the context of families with dependent children. | Study sample does not include lone parent or couple families that contain one or more dependent children. |
| 5 | Study population includes parents who are 'poor'. Our definition of 'poor' includes, but is not limited to, those described as: low paid, low skilled, lone parents, recipients of social welfare benefits or in-work financial support, and/or part-time employees. Include studies where intervention is targeted at these groups (even if the sample contains families from across the entire income spectrum). | Families are not in poverty. |

Continued

Table 2.1 Continued

| Number | Inclusion criteria | Exclusion criteria |
|--------|---|---|
| 6 | Study evaluates an intervention. | Is an evaluative study of (i) professional skills training programmes; or (ii) childhood nutritional or prenatal programmes (for example, home visiting) targeting premature, or other at-risk children; or (iii) longitudinal initiatives designed to improve child outcomes and with no explicit provision for parents (e.g. Perry Pre-school Programme); or (iv) based on micro-simulation modelling that makes no allowance for the impact of tax/benefit changes on individual behaviour (i.e. estimates the immediate impact of a policy change by itself). |
| 7 | Is an evaluative study that measures changes to in-work poverty or relevant proxy outcomes. Relevant proxy outcomes include: (i) changes to the number of working members within the family; (ii) changes to the number of hours worked by one, or more, family members; (iii) changes to wages, earnings or income. Studies measuring 'softer' outcomes that can indicate willingness to enter employment (such as job applications, job searches, and attitudes to work) are within the scope of the map. Studies reporting on a broad range of softer outcomes related to changes in wages/earnings/income, reflecting low pay as a cause of in-work poverty, are also relevant. These outcomes include, but are not limited to, career advancement, job promotion, acquisition of educational qualifications, and participation in education or training. | Is an evaluative study of an intervention that has the potential to reduce in-work poverty, but which reports only outcomes that are not relevant to this review (e.g. marriage, health). Welfare receipt (caseloads reductions etc.) is not considered a sufficient basis on which to evaluate the interventions, as such data does not indicate why the welfare spell ended (i.e. may have been unrelated to employment). |
| 8 | Is non-evaluative research on the barriers to, or facilitators of, reducing in-work poverty among families with dependent children. | Is non-evaluative research that is not about the barriers to, or facilitators of, reducing in-work poverty among families with dependent children. Non-evaluative studies that investigate the strategies families use to cope with the financial constraints of in-work poverty (e.g. borrowing from neighbours) are not within the scope of the map. |

Continued

Table 2.1 Continued

| Number | Inclusion criteria | Exclusion criteria |
|---------------|---|---|
| 9 | Is non-evaluative research on the barriers to, and facilitators of, reducing in-work poverty among families with dependent children studies (factors studies) conducted in the UK. | Is non-evaluative research on the barriers to, and facilitators of, reducing in-work poverty among families with dependent children (factors studies) conducted outside the UK. |
| 10 | Report describes primary empirical research: that is, it is a study presenting data obtained by systematic processes. The data is inferential, rather than descriptive, in their application. | Is not primary empirical research: for example, editorials, commentaries or book reviews; policy documents; non-systematic reviews; systematic reviews; resources; bibliographies; theoretical or methodological studies. |

2.3 Identification of potential studies: search strategy

A highly sensitive search strategy was developed using the review questions, the three key concepts (families with dependent children, employment, poverty), and the inclusion/exclusion criteria. Full details of the search strategy are given in Appendix A of the Technical Report.

The following range of sources was used:

- 19 electronic bibliographic databases;
- websites of 15 organisations known to have an interest in the topic area of the review;
- search engines (Google and Google Scholar).

The search terms (or key words) were developed iteratively using a combination of techniques.

- Free-text terms and relevant index terms were identified (both synonyms and antonyms) which could be used to describe the important concepts (families with dependent children, employment, poverty).
- Pilot searches were undertaken to test the terms identified, which were then refined and used to search the bibliographic databases.

Searches were conducted during June 2007. All searches covered the period 1992-2007. All citations identified in the above searches were imported into the EPPI-Centre's custom-designed in-house reviewing software, EPPI-Reviewer, and scanned for relevance against the criteria (Thomas and Brunton, 2006).

2.4 Screening studies: applying inclusion and exclusion criteria

The screening process involved the hierarchical application of the exclusion criteria (see Table 2.1) to titles and abstracts. Full reports were obtained for those

studies that appeared to meet our inclusion criteria, or where we had insufficient information to be sure. The exclusion criteria were re-applied to the full reports, and those that did not meet the criteria were excluded.

2.5 Characterising included studies

The coding of studies included in the map provided data for the purposes of describing, or mapping, the overall field of research on the topic area. Included studies were coded for contextual information (e.g. characteristics of the population and intervention, where applicable) and methodological information, using the standardised EPPI-Centre coding framework and coding questions developed specifically for this review. For full details of the coding tool, see Appendix B of the Technical Report.

The coding included the following:

- where research was conducted (i.e. country and other contextual features);
- research design and data collection methods;
- who was studied;
- what was researched (e.g. type of intervention);
- what kinds of outcomes were investigated.

In order to understand and group similar interventions more effectively, reviewers attempted to identify the theory of change underpinning each intervention; in other words, the mechanism(s) through which the intervention was trying to bring about the outcomes it intended to promote or achieve. Where possible, interventions were also categorised into pre-employment, transition, and/or post-employment measures.

2.6 Identifying and describing studies: quality-assurance process

Researchers involved in screening and coding took part in moderation exercises where samples of papers were screened or coded (depending on the stage of the review) by all members involved. Results were compared and discussed. When reviewers were unable to reach a decision, discussions were held until a consensus was reached. The team leader also carried out independent audits of each team member's screening decisions and coding, on random samples of papers. These procedures aimed to ensure consistency in interpretations of the selection criteria and coding tool for this review.

3 Systematic map results

3.1 Overview

This chapter presents a summary of the results of the systematic mapping of studies. Section 3.2 describes the flow of literature through the map, and Section 3.3 reports some of the main characteristics of the studies included in the map. The full results of the mapping stage are presented in Appendix D of the Technical Report, while the reference list for these studies can be found in Appendix H.

3.2 Identifying studies

The searches identified a total of 12,322 records. Comprehensive searches of electronic bibliographic databases identified 12,131 citations; searches of Google, Google Scholar and websites known to have an interest in this area led to the identification of a further 191 citations. Figure 3.1 describes the flow of this literature through each stage of the mapping process.

After removing 3,178 duplicate records, the titles and abstracts of 9,144 records were screened. Of these, most did not meet the inclusion criteria and so were excluded from the map (93 per cent). A high proportion of these were excluded because the study population did not include parents with dependent children (38 per cent). Nearly one-fifth of studies were excluded because the study population did not meet our definition of being 'in poverty' (19 per cent). Thirteen per cent of studies were excluded on the grounds that, although they provided non-evaluative research on barriers/facilitators to reducing in-work poverty among families with dependent children, they were carried out in settings outside the UK.

A total of 627 reports were identified which answered the two broad review sub-questions. Thirty-three items were either unavailable, or arrived too late for the map. A total of 439 studies (reported in 594 'papers') were included in the map. Forty-seven studies answered both sub-questions.

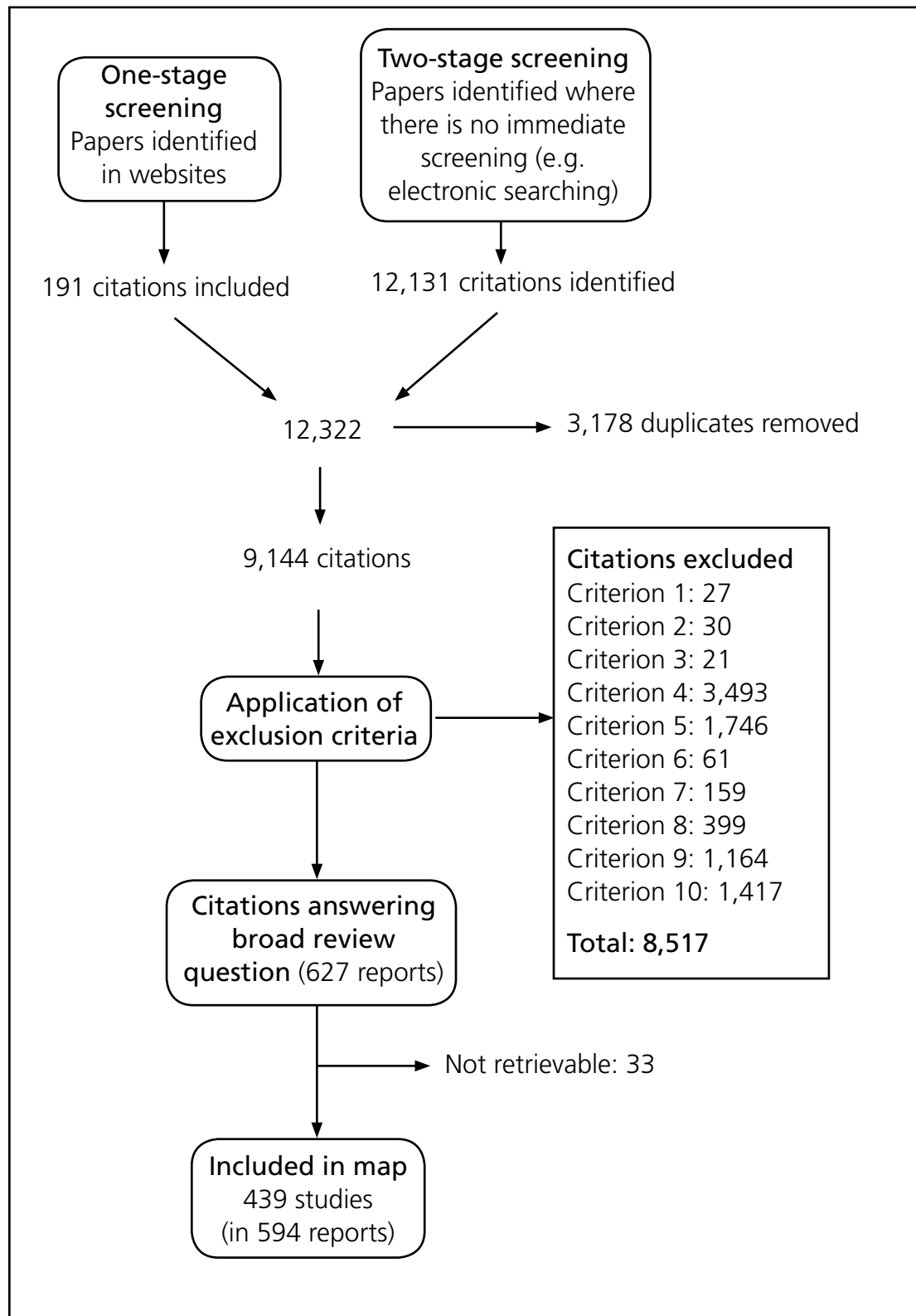
- 1 *What is the nature and extent of the evaluative research that has been undertaken on the design and impact of interventions with the potential to reduce in-work poverty among families with dependent children?*

(285 studies)

- 2 *What is the nature and extent of the non-evaluative research has there been undertaken on the barriers to, and facilitators of, reducing in-work poverty among families with dependent children?*

(201 studies)

Around one-quarter of the total number of studies included in the map were identified through website searches, with the bulk of these originating from two organisations: the DWP in the UK, and the Manpower Demonstration Research Corporation (MDRC) in the USA.

Figure 3.1 Selection of studies for the systematic map

3.3 Summary of findings from the systematic map

This section reports a summary of findings from the systematic map. The full results of the mapping stage are presented in Appendix D of the Technical Report.

A total of 439 empirical studies were included in the map. Of these, 285 studies provided evaluative evidence about the design and impact of interventions with the potential to help families with dependent children transition out of in-work poverty (**intervention studies**). A total of 201 studies provided non-evaluative research evidence on the barriers to, and/or facilitators of, reducing in-work poverty among these families (**factors studies**). The groups are not mutually exclusive.

3.3.1 Characteristics of factors studies (201 studies)

A range of data was collected, principally details of the study itself and characteristics of the study population.

Countries in which studies were conducted

Reflecting the selection criteria, all 201 factors studies were conducted in the UK.

Study design

There was a range of study methods and designs, including 67 cross-sectional studies where data has typically been collected in national surveys, and 75 'views' studies. In this latter group of studies, researchers have tried to understand phenomenon from the 'worldview' of a particular group, culture or society, placing attention on subjective meaning, perspectives and experience.

Population

Seventeen of the 201 factors studies sampled only couple families with dependent children. A further 85 studies included both lone parents and members from couple families. Relatively few studies appear to have included both parents from the same family within the sample (nine per cent); even fewer conducted paired in-depth interviews. One of the 63 studies that purposively sampled lone parents was focused solely on lone fathers. Forty-one studies included employers and other stakeholders in their sample.

A majority of the 201 factors studies included both those who were in paid employment, and those who were not (68 per cent). No studies included full-time workers only.

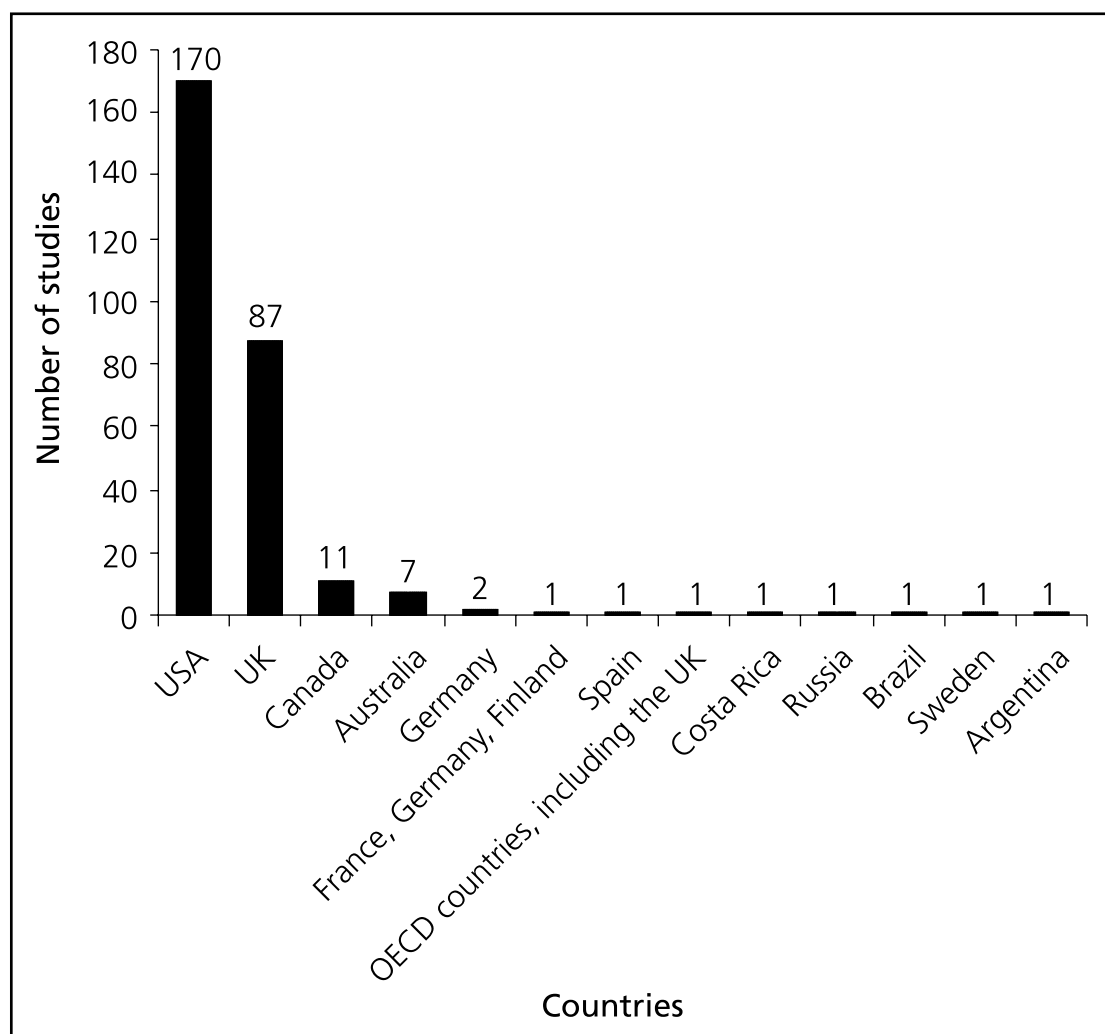
3.3.2 Characteristics of intervention studies (285 studies)

In these types of studies, in addition to details of the study itself and characteristics of the study population, information was collected about the interventions and outcomes measured.

Countries in which studies were conducted

Evaluative studies of interventions which have the potential to reduce in-work poverty among families with dependent children were found to be more common in the USA than in any other country (170 studies; 60 per cent). Studies conducted in the UK accounted for 31 per cent of the total number of intervention studies (88 UK studies in total, including one conducted in both the UK and other OECD countries). Figure 3.2 details the range of countries in which studies were conducted.

Figure 3.2 Evaluations by study country (285 studies)



Evaluation design

A total of 45 intervention studies (16 per cent) were RCTs. Only one RCT was conducted in the UK.

Population

Thirteen intervention studies included only couple families with dependent children in their samples. A further 115 studies included both lone and two-parent families.

In 162 studies (57 per cent), the study population consisted of some participants who were in-work and others who were not. No studies included full-time workers only.

Of the 13 studies that focused solely on two-parent families, four were based on UK populations; six studies were conducted in the USA, and one each in Spain, Sweden and Australia. Around half of these studies included at least some participants who were in paid employment, although, typically, the interventions were targeted at participants receiving social welfare payments.

Type of intervention

Different types of interventions with the potential to reduce in-work poverty in working families with children have been the focus of evaluation (see Table 3.1). A total of 305 interventions, or individual components of interventions, were evaluated in the 285 studies that we identified (i.e. some studies separately evaluated more than one).

Most of the reports described studies that evaluated financial initiatives (47 per cent) or employment support services (39 per cent). Very few studies evaluated interventions that were categorised as education initiatives. Both UK and non-UK studies displayed similar patterns in this respect. All ten interventions grouped as resource provision that were evaluated in the UK were initiatives that expanded childcare provision in some way, such as extended or more flexible opening hours. In contrast, only one of the seven non-UK resource provision interventions was concerned primarily with improving the availability and/or accessibility of childcare services to help parents move out of in-work poverty. Of the remaining six non-UK interventions in the resources category, two studies evaluated transitional housing programmes (one of them for teenage parents) and four studies assessed the effectiveness of residential relocation schemes. No evaluations of comparable schemes in the UK were identified. All seven evaluative studies of medical/social services interventions included in the map were conducted outside the UK. Only one study of family-friendly initiatives, such as flexible working policies, was located. The evaluation was conducted in the UK.

Table 3.1 Type of intervention (305 interventions* in 285 studies)

| Type of intervention | Number of UK studies | Number of non-UK studies |
|-----------------------------|----------------------|--------------------------|
| Employment support | 35 | 83 |
| Education | 5 | 13 |
| Financial | 43 | 101 |
| Social/medical | 0 | 7 |
| Resource provision | 10 | 7 |
| Family-friendly initiatives | 1 | 0 |

*or individual components of interventions.

Interventions in the financial group were further categorised, based on whether they pushed (e.g. via benefit sanctions) or pulled (e.g. via earnings supplements) people not doing paid work into the labour market. In both the UK and non-UK studies, most of the financial interventions were of the latter type. The majority of the studies evaluating financial 'push' interventions were conducted outside the UK.

A further aspect of the interventions included in the map concerns the timing of provision. The most commonly evaluated interventions were those targeted at the pre-employment phase (199 studies); post-employment support was evaluated in 142 studies and transitional measures in 26 studies. A relatively large number of studies evaluated interventions that offered pre- and post-employment support (65 studies). In 11 studies, the interventions offered a continuum of support across all three time periods (pre-employment, transition, post-employment).

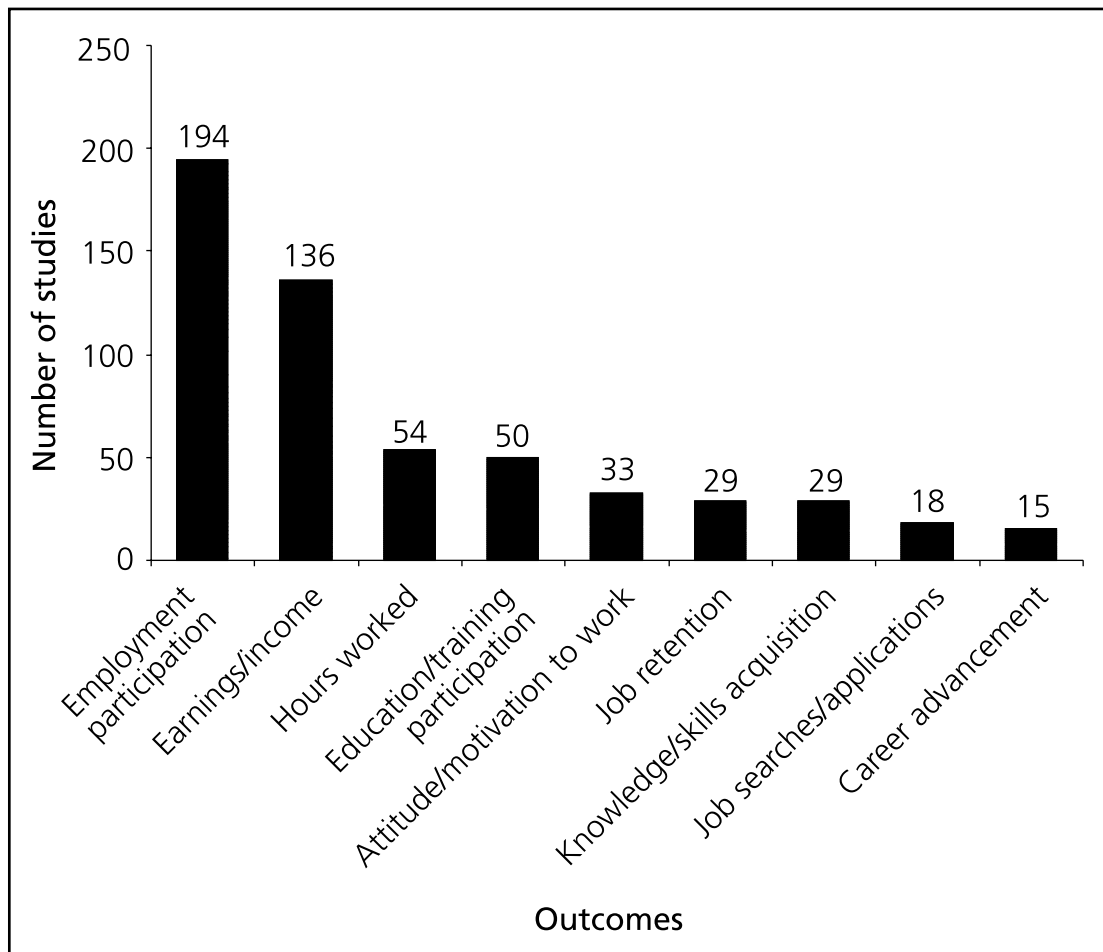
Reflecting the multi-component nature of the interventions, many sought to influence behaviour by addressing more than one barrier to employment. For example, in addition to enhancing participants' employment-related knowledge and skills, such as job searching or interview techniques, employment support programmes were often designed to change participants' qualifications and/or finances. Both UK and non-UK interventions displayed similar patterns in this respect. However, there were differences between UK and non-UK studies in terms of some of the other mechanisms of change adopted by the interventions: for example, there were no evaluations of UK interventions incorporating services that addressed participants' health difficulties or behaviours, such as substance-abuse.

Outcomes

A broad range of outcomes was measured by the evaluative studies included in the map. Three of the most common outcomes were changes in employment participation (194 studies), wages/earnings/income (136 studies) and hours worked (50 studies). In a relatively large number of studies there was a focus on detailing participants' perceptions of impact: for example, whether they thought the intervention had contributed to a change in employment status,

attitude to education, and so on (75 studies overall). Many studies also measured non-employment-related outcomes, including changes to fertility, marriage patterns, and/or health and wellbeing. Some of the most common employment-related outcomes are presented in Figure 3.3.

Figure 3.3 Outcomes measured (285 studies)



4 Synthesis methodology

4.1 Overview

This chapter describes the methods used in the second stage of the review: the synthesis. This stage of the review involved identifying the synthesis question and the relevant subset of studies capable of answering it, including the development of a second set of exclusion criteria (see Section 4.2); extracting data for in-depth analysis of the studies (see Section 4.3); and, appraisal of study quality (see Section 4.4). The quality-assurance process is outlined in Section 4.5, the methods of synthesis in Section 4.6, and deriving conclusions and implications in Section 4.7.

4.2 Moving from broad characterisation (mapping) to synthesis

As indicated in Chapter 2, this review involved the use of two stages: mapping (stage 1) and synthesis (stage 2). The map describes the nature and extent of the empirical literature in the broad field investigated by this review, identifying where research activity has been concentrated and where there are gaps. This information was used to develop a narrower, more focused, question to be answered by the synthesis. The Advisory Group met in February 2008 to discuss the results of the mapping exercise. The result of the meeting was the following question to be answered by the synthesis:

What is the effectiveness of interventions with the potential to reduce in-work poverty in couple families with dependent children?

To be included in the descriptive mapping, studies had to pass a total of ten exclusion criteria which were applied hierarchically to the titles and abstracts, and then full texts (see Table 2.1 in Chapter 2). To identify relevant studies for inclusion in the synthesis, a second set of selection criteria was developed and applied to the 439 studies included in the map. The stage 2 exclusion criteria, numbered 11 to 20, follow. Again, these were developed in consultation with members of the Advisory Group.

Exclusion criteria (Stage 2: synthesis)

- 11 Study does not evaluate an intervention.
- 12 Study was published before 1998.
- 13 Study population does not include couple families with dependent children.
- 14 Study does not estimate the effect of the intervention separately for parents: in other words, the intervention is available to both parents and childless adults but findings have only been provided for the whole group.
- 15 Study does not estimate the effect of the intervention separately for two-parent families, either for the couple as a whole, or for individuals in couples: for example, the study sample contains both lone and partnered parents but findings have only been provided for the whole group.
- 16 Study does not estimate the effect of the intervention separately for poor couple families with dependent children (others in the sample may be poor).
- 17 Focus of the study (either in terms of the sample, or the design of the intervention) was on workless couple families with dependent children (broadly defined as those receiving social welfare payments).
- 18 Study does not estimate the effect of the intervention separately for two-parent families containing at least one working member.
- 19 Study estimates the impact of a UK intervention on a non-UK population.
- 20 Study measures participants' perceptions of the intervention's impact only.

At the end of this stage of further selection, the reference lists of those studies to be included in the synthesis were scanned for further studies that were capable of answering the synthesis question. Any potentially relevant study was screened against the selection criteria. Each new relevant item that was identified was also searched for relevant references, until this process was exhausted. A number of additional relevant 'linked' reports were also identified at this stage (i.e. reports of the same study in different publications). In addition, the process of checking reference lists identified further studies that were potentially relevant for the map, especially studies focusing on lone parents. These studies were not added to the map.

4.3 Detailed description of studies in the synthesis

Studies identified as meeting the selection criteria were analysed in depth, using the EPPI-Centre's detailed data-extraction software, EPPI-Reviewer. Again, the standard EPPI-Centre guidelines were adapted. Additional review-specific questions were identified and added to the coding guidelines that had been produced for the task of characterising the studies included in the systematic map. Data was collected on the study aims; research questions and focus; research methods, including design and sampling; data collection; data analysis; and results and conclusions

(see Appendix B of the Technical Report). Where relevant, authors were contacted for clarification.

Each study was data-extracted by the team leader, with a second member of the team acting as second reviewer (four team members in total). The results and conclusions for each study, comprising both numerical data and authors' narrative reports, were extracted and considered by reviewers. Only those study findings and conclusions considered to be relevant to the review were extracted, for potential inclusion in the synthesis. The results as reported here represent a critical interpretation of those reported in the original studies.

4.4 Assessing quality of studies and weight of evidence for the review question

The policy level question addressed by this review is whether systematic intervention to change household income leads to families transitioning out of in-work poverty. Will it lead to additional 'benefit' (and/or prevent 'harm')? The review question is therefore about impact; i.e. does any difference in the independent variable (income) 'cause' systematic differences in the dependent outcome variable(s).

An important issue for interpretation and synthesis is to what extent the results obtained in an individual study that attempts to answer a 'what works' question can be said to be an estimate of the 'true' state, rather than an artefact of the study design and/or method of analysis (i.e. bias). A study that aimed to investigate this question would attempt to minimise the various threats to validity that can produce bias in the study results (see Appendix C of the Technical Report). Studies that do not control for all potential biases yield estimates that may deviate from the true underlying relationship beyond the play of chance, due to the effects of confounding factors, biases or both. The main problem is not the lack of precision, but the production of findings that are seriously biased or confounded. Empirical evidence has demonstrated that the optimal research design for questions of impact is prospective randomised experiment. This design is widely regarded as optimal, both because it establishes the direction of causality and minimises the effects of bias (Boruch and Wortman, 1979; Tate, 1982; Torgerson and Torgerson, 2001). While none of the studies in the synthesis are of this type, they were all designed with the intention of answering an impact question.

When considering the interpretation and synthesis of the results, it is therefore appropriate to estimate the extent to which each of the individual studies 'controlled' for the various threats to validity. In this synthesis, all the studies were retrospective evaluations, with all but one employing statistical modelling of varying degrees of complexity. Several were econometric evaluations. Methods for reviewing this type of research are under-developed, and a literature search was unable to identify any fully developed and tested quality appraisal tools (e.g. checklists). In the absence of this, we took a pragmatic approach. This was based on an adaptation of the standard EPPI-Centre WoE framework, the criteria

for which are based on the evidence of the degree to which a particular study controlled for threats to internal validity (Gough, 2007), with supplementary criteria specific to model validation incorporated as appropriate. Here we were guided by methodological developments in the field of health economics (Soto, 2002; Weinstein *et al.* 2003). Further details can be found in Appendices B and C of the Technical Report.

The WoE framework consists of the following three elements:

WoE A: Execution of the study. Was the study carried out well? This took into consideration sampling and allocation (where appropriate), data collection, data analysis and quality of the reporting. WoE A was based on reviewers' judgments, taking into account questions in sections H, I, J and K from the coding tool. Each reviewer coded WoE A individually and any discrepancies were discussed in depth and resolved.

High = met most criteria specified in sections H, I, J and K of the coding tool.

Medium = met some criteria specified in sections H, I, J and K of the coding tool.

Low = met fewer than half the criteria specified in sections H, I, J and K of the coding tool.

WoE B: Appropriateness of the research design and analysis used for answering the synthesis question. The review question, in asking about the impact of interventions, implies causality.

High = prospective experiment with random allocation to groups.

Medium = prospective experiment with statistically equivalent groups.

Low = all other designs.

WoE C: Relevance of the focus of the study (including conceptual focus, context, sample, measures, scenario, or other indicators of the focus of the study) for addressing the synthesis question.

High = reported the employment status of both partners and was conducted in the UK.

Medium = reported the employment status of one partner only and/or was conducted outside the UK.

No studies could be judged as low.

WoE D: An overall weight of evidence, taking into account A, B and C.

WoE D = the average of A, B and C. However, D could not be higher than the average of A and B.

4.5 Synthesis: quality-assurance process

For each of the studies selected for synthesis, the data extraction and assessment of the weight of evidence were conducted independently by two reviewers. Comparative reports were then discussed until any discrepancies, in either the application of the data-extraction tool or the apportioning of different weights to the findings and conclusions of the studies, were resolved.

4.6 Synthesis of evidence

The synthesis combines the results from individual studies to produce an overall result for answering the more narrowly specified second review question: *What is the effectiveness of interventions with the potential to reduce in-work poverty in couple families with dependent children?*

The synthesis of study results draws upon both the authors' interpretation of their study results and our (the review team's) interpretation. In the primary studies, effects are typically presented in the form of percentage point changes: where appropriate, actual figures have been reported in the review. The real sizes of the changes reported in each study are difficult to compare and synthesise. For example, technically it is not possible to create 'an average' using meta-analysis. What is comparable across the study results is the 'direction of effect'. Whether a study result is presented in terms of a percentage change or a regression coefficient, the result can be interpreted as either 'positive' (i.e. the outcome favours the intervention) or 'negative' (i.e. the outcome does not favour the intervention). A 'vote-counting' approach has been used to summarise the effects across studies, i.e. to compare the number of positive studies with the number of negative studies (Higgins and Green, 2008). In this review, this vote counting approach is based on direction of effect only, not any 'statistical significance' reported in the studies.¹

The next stage of the synthesis focused on searching for patterns of similarity or difference in the direction of effects in each outcome category. Potential explanations for any identified variation in direction of effect were then explored, based on any differences between studies in:

- study quality;
- study methods;

¹ Statistical significance is an estimate of the likelihood of a particular result occurring by chance. Statistical significance is linked to sample size and the probability of a study being able to detect an effect of a given magnitude. The absence of a statistical significant effect does not necessarily mean that there is no effect, but may simply be the result of the sample being too small to detect an effect at the given significance level. None of the studies in the review supply information about prior sample size calculations.

- sample/context characteristics;
- specific characteristics of the intervention.

Finally, we looked to see if any variation in direction of effect could be explained by reference to systematic differences between studies. If we were able to find such differences, we could then be more confident that any remaining consistent patterns in direction of effect provided valid summary estimates of the impact of this kind of intervention.

4.7 Deriving conclusions and implications

The synthesis was used to draw conclusions and identify implications for policy and research. This was based on discussions between reviewers.

5 Synthesis results

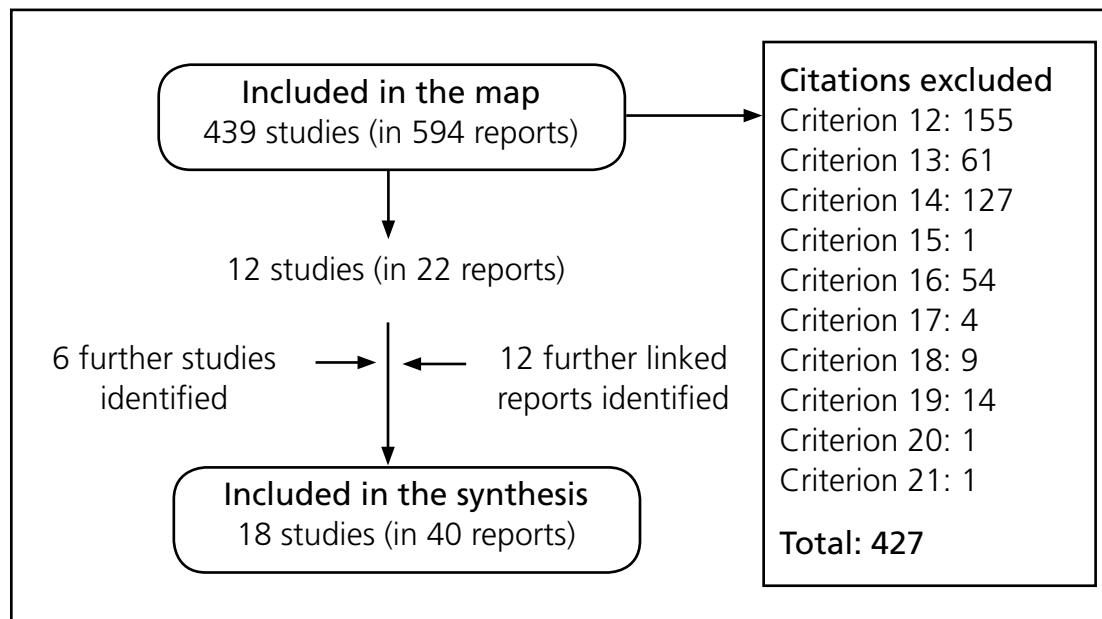
5.1 Overview

This chapter describes the selection of studies for synthesis (see Section 5.2) and the characteristics of the studies in the synthesis (see Section 5.3). It also outlines the quality judgements (weights of evidence) assigned to each of the studies (see Section 5.4), before presenting the synthesis of evidence (see Section 5.5) and summary of findings (Section 5.6).

5.2 Selecting studies for synthesis

The process of selecting studies for synthesis has already been outlined in Chapter 4. In summary, this involved developing and applying an additional set of inclusion/exclusion criteria, and scanning the reference lists of included studies.

Each of the 439 studies included in the map was screened against the second-stage selection criteria. The most common reasons for exclusion were that the study did not evaluate an intervention (155 studies), or there were no couple families with dependent children in the sample (127 studies). In addition, 61 studies were excluded because they were published before 1998. A further 54 studies were excluded on the grounds that, although the study population contained two-parent families, there was no sub-group analysis for these parents (i.e. the study did not conduct separate analyses for different types of families). Likewise, 14 studies were excluded because they did not estimate separate effects for the couple families in the sample who were working at the beginning of the study. In nine studies, the focus was on workless families. Four other exclusion criteria resulted in a further seven studies being excluded. The remaining 12 studies met the inclusion criteria for the synthesis. The reference lists of these studies were scanned, leading to a further five relevant studies. A further one study was brought to our attention through personal communication. Eighteen studies were included in the synthesis.

Figure 5.1 Selection of studies for the synthesis

5.3 Characteristics of studies included in the synthesis

The majority of studies in the synthesis were conducted on UK populations, with three studies from the USA, and one each carried out in Germany, Sweden and Australia. Proportionally, the number of studies from the UK is greater in the synthesis than in the systematic map (67 per cent compared to 31 per cent). Reflecting the question to be answered by the synthesis, all studies evaluated an intervention. Further characteristics of the studies are presented in Sections 5.3.1 (population), 5.3.2 (interventions), 5.3.3 (outcomes) and 5.3.4 (evaluation design).

5.3.1 Population

All studies evaluated the effectiveness of interventions on parents of dependent children from couple families. Fifteen studies included in their samples both lone parents and parents from couple families, with the results separately estimated for the two family types (this reflecting the second-stage selection criterion explicitly excluding studies that either sampled lone parents only or did not report a separate effect on two-parent families). Three studies included only couple families (Eissa and Hoynes, 2006; Francesconi *et al.*, 2007; Bonin *et al.*, 2003). In this review, no distinction is made between couple families based on marriage and de facto married (i.e. cohabiting) couples, although this distinction is made in some of the individual primary studies.

The main samples in three studies contained adults generally, and the effects on different sub-samples, including parents with dependent children, were reported. The number of parents in the adult samples was not indicated. In two studies, the intervention of interest was available only to parents (Gregg *et al.*, 1999; Bonin *et*

al., 2003). The third study evaluated an intervention that was available to working adults (Creedy *et al.*, 2003).

The majority of studies (15 in total) restricted their main samples to parents of dependent children, (although several of these studies used childless adults as the comparison group). The main samples in 12 of these 15 studies broadly represented the entire income distribution.² However, four of these 12 studies also subdivided their main sample by income level, and presented additional findings for sub-groups of poor and/or low qualified families with dependent children. Of these, three studies reported sub-group analysis findings that were relevant to this review (Ellwood, 2000; Flood *et al.*, 2004; Lydon and Walker, 2005). The remaining three studies (out of the total 15) purposively sampled only those parents who were most likely to be affected by the reform in question (i.e. lower income parents). Of these, two studies used low education (less than 12 years of schooling) as a proxy for low income (Eissa and Hoynes, 2006; Heim, 2006). The third study excluded from its main sample couples where the man worked more than 16 hours and earned in the top quartile of the earnings distribution (Francesconi *et al.*, 2007).

The majority of studies had samples containing both working and non-working couple families with dependent children. This limited the number of relevant findings that could be extracted for synthesis in the review. Only two studies sampled members of working couple families only, either single- or dual-earning. The study by Eissa and Hoynes (2006) excluded those couples who reported zero hours of work. The study by Ellwood (2000) focused solely on women; our understanding is that all married mothers in the sample had employed husbands.

Sixteen studies drew their samples from datasets generated by national, longitudinal surveys. National surveys covering Britain/UK included the Labour Force Survey (LFS), the Family Resources Survey (FRS), the British Panel Household Survey (BPHS) and the Family Expenditure Survey (FES). Non-UK national surveys included the USA's Current Population Survey (CPS), the German Socio-Economic Panel (SOEP), the Australian Survey of Income and Housing Costs (SIHC) and the Swedish Income Distribution Survey (HEK). Two UK studies used a survey that was specifically designed for policy evaluation (the Family and Children Study (FACS)).

As supplementary analyses, several studies measured impacts on sub-samples of the main population, grouped by the employment status of partners, age of parents, age of youngest child, and/or number of children in the family.

² In a number of studies parents facing constraints on their ability to participate in the labour market were excluded during the sample selection process (for example, those involved in education or on disability benefits).

5.3.2 Interventions

All 18 studies included for synthesis evaluated financial interventions. This suggests an absence of other types of interventions for working couple families with dependent children, and/or an absence of empirical research evaluating those that do exist.

All the interventions were post-employment initiatives available only to those who were engaged in paid employment. Seventeen studies evaluated interventions targeted primarily at the working poor: that is, available only to those on earnings/incomes below a certain limit. The remaining study evaluated an intervention that had two components, one of which was available to all workers and the other affecting only those on low incomes (Flood *et al.*, 2004). None of the interventions were targeted solely at couple families, but were available to all types of families with dependent children.

Although many of them differ in the detail of their design and/or operation, the reviewed financial interventions have elements in common. They all involve the transfer of income through the tax and/or the benefit system to qualifying families, either directly through cash payments, indirectly through reductions in tax or social insurance liabilities, or through a combination of these approaches.

Generally, the reviewed interventions provide maximum credits/benefits to those on the lowest earnings/incomes, and then reduce awards as the family's income rises; thereby, according to economic theory, acting as a work disincentive (the 'income effect').

One of the 18 studies is slightly different to the others, in that it evaluates an intervention that combines an approach that more generously transfers income to working families (financial 'pull') with an alternative strategy reducing social/housing assistance payments that may actually reduce the income of some families, unless they respond by increasing participation/hours (financial 'push'). These opposing strategies operate simultaneously, although only the poorest families claiming benefits are likely to be affected by the financial 'push' element (Flood *et al.*, 2004).

Sixteen studies evaluated national policies that had already been implemented. Of these, one study considered the effect of introducing a completely new policy. The comparison, therefore, was with 'no intervention' (strictly speaking, the tax/transfer system that prevailed prior to the intervention). The other 15 evaluations of up-and-running interventions measured the effect of replacing one version of a programme with another one (sometimes involving a change of name). In each case, the modifications involved increasing the financial award payable to eligible parents. (Although the modifications usually also addressed other problems, including fraud and administration, a major focus was always making payments more generous). The remaining two studies evaluated hypothetical changes to existing programmes (Flood *et al.*, 2004; Creedy *et al.*, 2003). Here, the estimated effects were relative to the situation that would have prevailed under the actual tax/benefit system.

Despite differences in what is being compared, a way of summarising the interventions might be to say that they all (in slightly different ways) allow workers to keep more of their salaries (and/or income, in the case of means-tested benefit changes) than in the scheme with which they are being compared.

Some authors reported that they had estimated the effects not only of the main intervention of interest, but also related tax/benefit and/or welfare reforms (including non-financial initiatives) that were introduced at the same time (Blundell *et al.*, 2004; Blundell *et al.*, 2005; Ellwood, 2000).³ This information is detailed in the individual study summaries in Section 5.3.5. However, in the following description of the interventions that were evaluated we have focused on the main interventions of interest.

UK interventions (12 studies)

- Family Credit (FC).
- Working Families' Tax Credit (WFTC).
- Working Tax Credit (WTC)/Child Tax Credit (CTC).

Non-UK interventions (six studies)

- USA: Earned Income Tax Credit (EITC).
- Germany: social insurance subsidies (more specifically, the variant of the Mainzer model that offered an additional child support subsidy).
- Sweden: hypothetical reform of the tax and benefit system (reduction in income tax liability operating alongside a 25 per cent cut in maximum benefit levels for housing/social allowances).
- Australia: hypothetical reform of the tax and benefit system (reduction in the withdrawal rate on means-tested benefits to 30 per cent).

The majority of studies (83 per cent) evaluated the impact of modifying tax credit programmes, either in the UK or the USA. Although there are key differences, the schemes are broadly similar.⁴

UK tax credits: FC was introduced in the UK in 1988 as a replacement for Family Income Supplement (FIS), partly in an attempt to tackle some of the problems associated with FIS, such as low take-up. A little over a decade later, FC was itself phased out and replaced by WFTC. Introduced in October 1999, WFTC had an enhanced credit for younger children, was more generous than FC in its higher earnings threshold (£90 compared with £79), was withdrawn at a gentler rate when income rose (55 per cent compared with 70 per cent) and incorporated

³ This was largely because the methodology used was unable to determine effects separately.

⁴ For a more detailed history and comparison of tax credit policies, see a number of the individual reviewed studies.

a new childcare credit that provided more generous help with childcare costs (up to 70 per cent of costs).

Less than four years after the introduction of WFTC, the UK's tax credit programme underwent further changes when WTC and CTC were introduced from April 2003. There are a number of differences between the current scheme and its predecessor, WFTC. For example, WTC can be paid to certain groups of childless individuals. CTC is designed to integrate all income-related support for children and therefore payment does not depend on participation in employment. Awards for childcare costs are also more generous: as of April 2006, they have met up to 80 per cent of eligible costs, up to a maximum of £175 for one child, and up to £300 for two or more children.

None of the UK tax credit schemes featured in this review (FC, WFTC or WTC/CTC) have involved the use of phase-in periods. In other words, families fulfilling the work and other conditions are immediately eligible for the maximum credit. Phase-out periods have been a feature of each of the schemes. As soon as income rises above a certain threshold, entitlement to the maximum credit is reduced, being tapered away until entitlement is zero. For those tax credits that depend on participation in employment (all except CTC), eligibility has also been conditional on number of hours worked and levels of earnings. As awards are based on net, rather than gross, earnings, interactions with other means-tested benefits are likely for many families (dampening any increase in generosity of tax credit awards).

None of the UK tax credit programmes have been targeted solely at couple families, but two of them incorporate initiatives that have been specifically designed to encourage greater employment participation in two-parent families, particularly among second earners. Both WFTC and WTC provide additional financial support (in the form of awards for the purchase of childcare services) that is paid in addition to the main tax credit award. Eligibility is conditional on both parents working at least 16 hours per week.

USA tax credits: The EITC was first introduced as a federal scheme in 1975. Since then, expansions to the scheme have taken place at regular intervals (approximately every three years). One of the most important expansions to the federal EITC followed, the Omnibus Budget Reconciliation Act of 1993 (OBRA93). The Act, substantially increased the generosity of payments: for example, by raising the subsidy rate, the maximum credit payable, and, for families with two or more children, the income eligibility range. State-level schemes have also been adopted in at least 17 states, making overall payments even more generous.

The EITC begins with a phase-in range, where the credit rises gradually as the family's earnings from employment increase. Then, when earnings reach a certain threshold, the credit stops rising and remains on a fixed rate; this is the maximum that can be received. This is known as the 'flat' or 'plateau' range. If the family's income rises above the flat range, the EITC moves into a phase-out range where credit falls gradually; at this point, both earnings and other sources of income,

such as investments and capital, are taken into account. For workers in the phase-in range of the EITC schedule, the subsidy acts exactly like an increase in their wage rate. Unlike the various UK tax credits programmes, the EITC is not conditional on the number of hours worked. It is also based on gross income, avoiding interactions with other benefits. For the versions of EITC evaluated by the studies in this review, there were no additional incentives specifically for two-parent families.⁵

5.3.3 Outcomes

The 18 studies included for synthesis reported a narrow range of outcomes which are viewed as playing some role in effecting reductions in poverty. On these grounds, they have been considered relevant proxy outcomes. The studies have been grouped into three categories, explanations of which follow. Table 5.1 details the outcomes measured in each study.

Wages and household income refers to studies measuring changes to either (a) wage levels of members of couple families with dependent children (either hourly rate or weekly wages), or (b) household income in such families.

Employment participation refers to studies measuring movements into (and/or out of) work among second earners in couple families with dependent children.

Working hours refers to studies measuring changes in the number of hours worked by one or both working members of couple families with dependent children.

Of the 16 studies measuring employment participation and/or the working hours, 15 chose to measure behavioural outcomes only. In the majority of these 15 studies, it appears to have been simply taken as given that the interventions improved families' financial situation. Interactions with other means-tested benefits aside, this is not an entirely unreasonable assumption since they all involved the direct or indirect transfer of cash to eligible families. The remaining study stands out, in that it measured not only the impact on working hours but also followed this up with an analysis of the effect on household income (i.e. having allowed for behaviour change associated with the policy). As noted earlier, the intervention evaluated in this study was slightly different from others in the review. The income tax cut – the financial 'pull' designed to make work pay – was combined with a strategy designed to 'push' people into entering work or increasing hours. The proposed income tax cut, by itself, would not necessarily financially compensate all those families who would experience the benefits cuts. This seems the most reasonable explanation for the study's focus on measuring the income effect after families had adjusted their work patterns in response to the policy.

⁵ Following more recent changes to the programmes, the phase-out stage now begins at a higher income for married couples than for single parents.

Table 5.1 Outcomes measured (18 studies)

| Study | Wages or household income | Participation | Working hours |
|----------------------------------|----------------------------------|----------------------|----------------------|
| Azmat (2006) | x | | |
| Blundell <i>et al.</i> (2000) | | x | x |
| Blundell <i>et al.</i> (2004) | | x | x |
| Blundell <i>et al.</i> (2005) | | x | |
| Bonin <i>et al.</i> (2003) | | x | x |
| Brewer <i>et al.</i> (2006) | | x | x |
| Chzhen and Middleton (2007) | | | x |
| Creedy <i>et al.</i> (2003) | | | x |
| Eissa and Hoynes (2006) | | x | x |
| Ellwood (2000) | | x | |
| Flood <i>et al.</i> (2004) | x | | x |
| Francesconi <i>et al.</i> (2007) | | x | |
| Gregg <i>et al.</i> (1999) | | x | |
| Heim (2006) | | | x |
| Leigh (2007) | | | x |
| Lydon and Walker (2005) | x | | |
| McKay (2003) | | x | x |
| Paull <i>et al.</i> (2002) | | | x |

Some of the reviewed interventions have multiple, and often complex, goals. This review is concerned only with their effectiveness in reducing in-work poverty for working couple families with dependent children. While a number of proxy outcomes for poverty (either income- or employment-related) were considered to be relevant for answering the review question, other outcomes that the studies may have reported (for example, take-up of benefits or childcare use) were not data-extracted for inclusion in the synthesis.

5.3.4 Evaluation design

All studies included for synthesis conducted retrospective quantitative evaluations. Four studies conducted evaluations using data from before the introduction of the interventions (Blundell *et al.*, 2000; Blundell *et al.*, 2004; Gregg *et al.*, 1999; Paull *et al.*, 2002).

Seventeen studies analysed data using statistical modelling of varying degrees of complexity. These studies used a research method that measured the statistical association between an 'input' variable (for example, a change in the amount or proportion of tax paid or tax credit awarded) and an 'outcome' variable that was a proxy for poverty (such as employment participation). Ten studies compared the outcomes for the same group of individuals after two different inputs (i.e. a single group pre-post intervention design) (Blundell *et al.*, 2000; Blundell *et*

al., 2004; Bonin *et al.*, 2003; Brewer *et al.*, 2006; Creedy *et al.*, 2003; Eissa and Hoynes 2006; Flood *et al.*, 2004; Gregg *et al.*, 1999; Heim, 2006; Paull *et al.*, 2002). Seven studies compared the outcomes of two different groups of people who had different inputs (i.e. a comparison group design) (Azmat, 2006; Blundell *et al.*, 2005; Chzhen and Middleton, 2007; Ellwood, 2000; Francesconi *et al.*, 2007; Leigh, 2007; Lydon and Walker, 2005). The study by McKay (2003) used a single-group post-test only design; measurements were taken over time, but there was no counterfactual comparative group.

5.3.5 Summaries of the characteristics of each study

The studies addressed a range of different questions about the intervention they evaluated. The summaries presented here focus on the aspects of the studies that are directly relevant to the question addressed by the synthesis.

For each study, reviewers identified the 'new' intervention (often referred to as 'a change' or 'reform' in the studies) and the control condition with which it was compared. In labelling the intervention and comparison, we have focused on the main financial element. The clarity of reporting varied between the studies. Therefore, in some studies, ascertaining which intervention or interventions were being evaluated, and/or to what the intervention was being compared, and/or for whom exactly the outcomes were measured, was not always straightforward. In some cases, this was partly because the main financial intervention of interest was one of several reforms introduced around the same time. There were instances of key information not being reported (e.g. whether changes in weekly or annual hours had been measured). Occasionally, we had to infer crucial information; where this was the case, it has been noted in the summary.

Our approach to summarising the results of the studies was to report the direction of effects reported in the papers, regardless of their size, even where the authors have concluded that the effect was so close to zero that there was no impact. The approach to summarising the methods used in each study has focused on the overall study design, rather than the statistical analysis. Unless otherwise stated, the studies used a single-group pre-post test design in which individuals acted as their own 'controls'.

The following summaries are necessarily brief. More detailed information on the studies, including details of other publications in which studies have been reported, are presented in Appendix E of the Technical Report.

Study 1

The study by **Azmat (2006)** focused on the indirect consequences of replacing the UK's main in-work benefit, FC, with WFTC in October 1999. The study sample was drawn from the LFS, 1997-2003. After pooling all 15 quarters, and excluding people in full-time education, those who were sick, disabled, or on a government training programme, the sample size was 366,317. Both one- and two-parent families were sampled.

The study found that married male WFTC claimants had a 17 per cent fall in gross wages, relative to matched, non-eligible workers who were similarly skilled and had the same predicted wage. As the rate of WFTC increased for the married male eligible worker, the study found that the gross wage fell by 24 per cent, relative to matched workers. In contrast, married mothers claiming WFTC were associated with a slight rise in gross wages of six per cent, relative to matched, non-eligible workers. For this group, as the rate of WFTC increased, wages were still found to have risen, compared with matched, non-eligible workers, but only by two per cent. The study also found evidence that suggested that, as the fraction of women or men eligible for WFTC increased in an industry (in an education group) there was a wage fall for all similar workers. This negative spill over effect applied to both married fathers and mothers. Finally, by including a measure which controlled for the change in generosity from FC to WFTC, the study concluded that the effects on gross wages were the result of a change in payment method (i.e. through the wage packet). The majority of the effects were reported as being statistically significant.

Study 2

The study reported in **Blundell et al. (2000)** estimated the likely labour market effects of replacing FC, with WFTC using pre-reform data drawn from the Family Resource Survey, 1994/95 and 1995/96. Initial predictions were based on 100 per cent take-up of FC, WFTC and the childcare component of WFTC. The study population included both lone parents and couple families with children. A total of 4,694 couple families were used in the analysis. The authors reported that the greatest income gain following the immediate replacement of FC with WFTC (assuming no change in labour market behaviour) would fall to mothers not in employment, whose partners were in-work.

The study found a net fall in the employment participation of mothers with employed partners, of around 0.57 per cent (i.e. approximately 20,000 individuals), following the replacement of FC with WFTC. This overall reduction was the result of 0.8 per cent moving from work to non-participation and around 0.2 per cent moving into the labour market. The authors report that 0.3 per cent of the male sample moved out of work (all had employed partners).⁶

The authors also experimented with a number of alternative scenarios, one of which was said to give a feel for the work incentive effects of basic WFTC (i.e. without the childcare credit component) relative to FC. For women with employed partners, an analysis based on 50 per cent take-up of the childcare credit found a net fall (0.73 per cent) in their participation, while zero per cent take-up was associated with a higher reduction in employment (0.81 per cent).

For employed mothers with working partners (i.e. dual-earning families), the study predicted that the increased generosity of WFTC would lead to a very

⁶ Some inconsistency in the various reports of this study.

small average reduction of 0.03 hours in their working hours (assumed weekly). Working mothers in single-earning families were predicted to increase their hours by an average of 0.14 hours.

Study 3

The study by **Blundell *et al.* (2004)** estimated the labour market impact of the UK personal tax and benefit system (dominated by WTC and CTC introduced in April 2003), compared with the previous personal tax and benefit system (as of April 2000). As no publicly available data-set existed (at that time) that would support an ex-post evaluation of the new tax credits, the authors estimated their likely labour market impact using FRS data from before their introduction (1995/96 to 2001/02). Sample sizes were not reported.

The study found that the tax and benefit changes between April 2000 and April 2003 (dominated by the new tax credits) produced a slight increase (0.22 per cent) in the employment participation rate of fathers with employed partners. This gain was counteracted by a considerably larger decrease (0.75 per cent) in the participation rate of mothers with employed partners (leading to an overall reduction of 0.29 per cent in dual-earning families). The study found that the younger the age of the youngest child in the family, the more positive the employment response. This pattern held for fathers with employed partners and to a lesser extent mothers with employed partners. The only families (as defined by ages of children) that increased the proportion of dual-earner couples were those in which the youngest child in the family was aged zero to two years. The study also estimated that the fewer the number of dependent children, the more positive the labour participation response. Again, this pattern held for both mothers and fathers with employed partners. Furthermore, only those families with one child were predicted to contribute positively to the proportion of dual-earning families. The study also found that the intervention led parents to reduce their hours of work. The effect was stronger for mothers (-0.7 hours) than for fathers (-0.23 hours).

Study 4

The study by **Blundell *et al.* (2005)** evaluated the effect of the WFTC and other contemporaneous reforms on parents, compared with the tax and welfare system previously in place. The main estimates are based on a very large sample of parents and childless adults (the comparison group) drawn from the LFS (more than 500,000 couple family observations). The model used in the analysis contains a variable that controls for eligibility to the intervention. For men with working partners, average treatment effects were found to be negative (-1 per cent). For women with employed partners, there was a positive treatment effect (0.1 per cent). Using FRS data, however, the study found that the estimated effect of WFTC and related reforms on mothers in couples whose partners were working was negative (-0.6 per cent). For fathers, there was no such discrepancy (i.e. treatment effect was still negative). None of these effects were statistically significant.

The analysis also explored other factors that might influence treatment responses. It was reported that labour supply responses for both mothers and fathers with working partners (again, using LFS data) tended to become more positive as the age of the youngest child increased. The study also suggested that the negative impact of WFTC and related reforms on mothers with working partners was concentrated among families with three or more children, although all the treatment effects were close to zero. For fathers with working partners, however, a different picture emerged: the study found the effects became more positive as the number of children in the family increased.

Study 5

The study by **Bonin et al. (2003)** assessed the impact on German couple families of three different proposals to support low-wage employment through subsidising social insurance contributions. One of the proposals (the Mainzer model) offered an additional subsidy to parents with dependent children. The study utilised data from the 2000 wave of the SOEP. A sample of 3,702 couple families, both with and without children, was used in the analysis (number of parents not reported). A comparison was made with estimated labour under the tax/benefit system that prevailed previously.

The study found that the Mainzer model, in the variant including the extra payment for children, reduced the proportion of dual-earning households (from 50.03 per cent to 49.27 per cent), primarily in favour of single-earner two-parent families with females working.

A further finding was that this variation of the Mainzer model reduced the average working hours of the sampled males by almost 1.3 per cent and produced no change in the number of hours worked by employed females.

Study 6

Brewer et al. (2006) investigated the effect of introducing WFTC using data from the FRS from 1995/96 to 2002/03 (i.e. the full period that WFTC was in effect: October 1999 to October 2003). The final sample was 12,729 lone parents and 31,403 couples with children. A number of slightly different comparisons were made. Here we report the results of only one.

The study found that replacing the FC programme (as of April 1999) with WFTC (as of April 2002) resulted in a statistically significant reduction of 0.48 percentage points (ppt) in the proportion of dual-earner families. The shift to one-earner families was mainly driven by mothers with employed partners leaving work (0.64 ppt), as this was larger than the positive effect on the participation of fathers with employed partners (0.19 ppt). In terms of those families that became dual-earner families, the study found that roughly similar proportions of male and female secondary workers were encouraged to enter employment as the result of the replacement of FC with WFTC (0.9 ppt/0.7 ppt).

Overall, reductions in the proportion of dual-earning families following the replacement of FC (April 1999) with WFTC (April 2002) was found to be greatest among families with three or more children (0.68 ppt), gradually becoming less negative as the number of children in the family decreased. Reductions in the proportion of families who were dual-earning were also more likely in families in which the youngest child was under the age of five years (zero to two years, 0.57 ppt; three to four years, 0.64 ppt); and here again there was a tendency for the effect to become gradually less negative as the age of the youngest child in the family increased. This pattern was similar for mothers with employed partners. However, the effect of replacing WFTC with FC on the labour supply of fathers with employed partners displayed the opposite pattern: the treatment effect was most positive for those with young children (0.42 ppt, 0.37 ppt) and those with large families (0.75 ppt), and gradually became less positive as the age of the youngest child increased and as the number of children decreased. For one sub-group of families, those in which the youngest child was aged 11+ years, there was a negative effect on the participation of fathers (0.02 ppt).

The study also found that the replacement of the FC programme (as of April 1999) with WFTC (as of April 2002) had little impact on the weekly working hours of members of couple families already in the labour force. For employed fathers with dependent children, the average change in weekly hours was a reduction of 0.03 hours, and for employed mothers a weekly reduction of 0.02 hours.

Study 7

The study by **Chzhen and Middleton (2007)** reported the impact of WTC compared with non-receipt of WTC, on the average working hours of mothers with employed partners from couple families. Data for the study was drawn from the three waves of the FACS, covering the periods 2002/03 to 2004/05. The main sample contained 5,320 mothers, including 3,607 with working partners. This particular analysis compared sub-samples totalling 353 recipients of WTC and 334 eligible non-recipients.

Coupled mothers from dual-earning families in receipt of WTC in 2004/05 were found to work an average of almost two and a half hours less than comparable non-recipients in the matched sample. This difference was reported as not being statistically significant.

Study 8

The study by **Creedy et al. (2003)** estimated the effects of a hypothetical change to the Australian tax and transfer system operating in March 1998. This hypothetical reform involved reducing the rate at which means-tested benefits were withdrawn as income increased. The study considered taper rate reductions to 30 per cent (from 50 per cent and 70 per cent), while leaving all basic benefit levels unchanged. A range of benefits were considered by the study (but not separately evaluated). The study used the Melbourne Institute Tax and Transfer Simulator (MITTS) and

data from the 1997/98 SIHC. The total sample (which included childless individuals) was 13,382 weighted observations. Effects of the policy change are presented for a range of sub-groups of the study population, including couple families with dependent children (sub-sample sizes not reported).

The study found that, as a result of the intervention, couples with dependent children would reduce their working hours by an average of 0.42 hours. In total, 33 per cent of the sample decreased their hours and less than half that proportion (15 per cent) increased their hours.

Study 9

The study by **Eissa and Hoynes (2006)** examined the labour market responses of married couples to expansions to the USA's EITC between 1984 and 1996. The study used the US Department of Labor's Current Population Survey data from 1984 to 1996, focusing solely on couple families containing at least one working member, and who had less than 12 years of schooling. Estimates were based on observations of over 17,000 married couples with dependent children. Some key findings are reported in an earlier journal article published in 2004.

The study found that the EITC expansions had modest effects on the participation rates of both members of two-parent families. Mothers were, on average, less likely to work under the 1996 EITC schedule, than if the 1984 schedule had been applied (1.1 per cent less likely to be in employment). Fathers were very slightly more likely to be working (0.2 per cent increase).

Further results (disaggregated by deciles of the husband's gross hourly wage and across different 'ranges' of the 1996 EITC schedule) highlighted substantial heterogeneity across the sampled population. Across the wage distribution, the response of fathers remained minimal but positive, gradually becoming less positive as wages increased. For mothers (all understood to have employed partners), treatment effect went in the opposite direction: the higher their partner's predicted wage, the less negative their labour supply response. Mothers married to the lowest waged men were found to have the largest responses (a reduced likelihood of working of 1.7 per cent). For only one sub-group of mothers, employment was more likely after the 1996 expansion; mothers in the small group of couples in the phase-in range (see discussion of EITC in Section 5.3.2) were 1.1 per cent more likely to work.

The study also calculated the impact of expansions to the EITC on annual working hours. Decreased annual hours of work for both partnered mothers and fathers receiving credits under the 1996 EITC schedule (relative to the hours worked under both the 1984 and 1993 schedules) were reported. Overall, mothers in the labour force were estimated to decrease their working hours by 0.7 per cent

to 3.9 per cent (between 11 and 81 hours).⁷ This negative effect held for both mothers with employed husbands and for the very small group of women with non-working husbands. Fathers decreased their working hours by 0.7 per cent to 2.6 per cent (between 15 and 54 hours). (The employment status of their partners was not reported.) The authors again reported responses in further subsets of the population. Across all wage deciles, mothers decreased their working hours. Those partnered with low-wage men reduced their work hours considerably more than women partnered with higher-waged men. Men in the middle of the wage distribution faced the strongest disincentive to increase their hours. Three sub-groups were found to work more hours under the 1996 EITC schedule, relative to hours worked under the 1984 programme. Men in the phase-in range (i.e. the phase during which, for the lowest earning workers, the credit rises as earnings rise) worked more hours (29 to 46 hours), as did both women in the phase-in and flat range of the EITC (26 to 172 hours and 5 to 24 hours respectively).

Study 10

The study by **Ellwood (2000)** estimated the impact of social policy reform in the USA, between March 1986 and March 1999, on the employment participation of married mothers. The main focus was on changes to the EITC programme during that period. Women were grouped by predicted wage quartiles. To construct the four groups, the study utilised the Current Population Survey to estimate wages for over 350,000 women aged 18-44 (both single and married), using characteristics highly correlated with pay, such as education and age, as independent variables. Those with calculated earnings of less than \$1 per hour or greater than \$75 per hour were excluded from the analysis. All the married women in the sample are understood to have husbands who were employed.

Individuals, within each wage group, were tracked over time, with comparisons made between the groups. A further comparison was made between the lowest waged mothers and low wage women without children. The author, having indicated that this latter comparison was highly questionable given the high rates of work among childless women, reported that the finding that he was 'most inclined to accept' was that the impact of social policy between 1986 and 1999 (understood to be dominated by the 1993 expansion to the EITC) reduced the employment participation of married mothers in the lowest quartile by three to five per cent, relative to married mothers in the third and second highest wage quartiles respectively.

⁷ All the findings reported in this section take into account the three expansions over the period 1984-1996. The study also simulated the effect of the 1993-1996 expansions (findings were similar). In the study (2006 publication), the figures reported in the text do not match those in tables labelled 5.7 and 5.8. We have used the information in the tables.

Study 11

Flood et al. (2004) reported the effects of a hypothetical reform to the tax and benefit system on the annual working hours and disposable incomes of Swedish two-parent families. The proposed modified tax and benefit system reduced income taxes for all workers and maximum benefit levels for both social and housing allowances (payable to those on low incomes). These opposing strategies (designed either to pull or push people not in paid employment into greater involvement in the labour market) operated simultaneously, although only the poorest families were likely to be affected by the financial 'push' element. While the income tax cut, by itself, would lead to working families being better off financially, the intervention 'package' was not necessarily more generous overall, compared with the system that was being replaced. Data used in the study was collected from the 1993 and 1999 HEK, the Official Statistics on Wages, government registers, telephone interviews and income tax returns. The total sample was 3,297.

Working parents in the poorest ten per cent of the sample were found to increase their annual working hours by an average change of four hours for fathers, and six hours for mothers. The authors also found that the effect on the poorest ten per cent of working couples in the sample was to increase net household income from an average of SEK 214,495 (about £17,000⁸) to SEK 249,270 (about £20,000) after the change.⁹

Study 12

Francesconi et al. (2007) assessed the impact that the introduction of WFTC had on two-parent households. Panel data was drawn from the first twelve waves (1991-2002) of the BHPS. The study purposively sampled lower income families, excluding couples where the man worked more than 16 hours and earned in the top quartile of the earnings distribution, as these earnings far exceeded the income eligibility thresholds for WFTC receipt. The sample that remained was further divided into low-earning and higher earning couples, based on male earnings. Low-earning was defined as having earnings in the bottom third of the male earnings distribution. The final sample included 3,235 couples, of whom 1,430 had dependent children. After pooling all 12 years of data, the sample size was 22,146 person-year observations, roughly evenly split between parents and childless couples. The model used in the analysis contains a variable that controls for eligibility to the intervention.

For both fathers and mothers coupled with partners working 16+ hours a week on low earnings, the authors reported that replacing FC with the more generous WFTC increased their probability of working (at both 16-30 hours and full-time hours): positive responses ranged between 0.5 and 2.3 percentage points. Limiting the sample to low-education mothers produced results that were only very slightly larger. It was also reported that, for employed mothers partnered with low-earning

⁸ Exchange rate as of July 2008.

⁹ These results were provided by the lead author.

men (i.e. dual-earning families), WFTC led to an increased probability of remaining in-work at 16+ hours per week by 0.8 per cent, and remaining in-work at 30+ hours per week by 1.3 per cent. For unemployed/inactive mothers partnered with low-earning men (i.e. single-earner families), the study again found a slight positive effect: here, the probability of entering part-time work increased by 0.3 per cent and by 0.6 per cent for full-time. It was also reported that, for mothers partnered with low-earning men, little variation was found in their labour supply responses by child's age and number of children. In each case, the effect on participation was positive, with 'one child' and 'youngest child aged zero to four years' producing the most positive effects for mothers working 16-30 hours. It was reported that none of the above treatment effects were statistically significant.

Study 13

The study by **Gregg *et al.* (1999)** investigated the likely impact of introducing WFTC on the probability of non-working people moving into work. Data was drawn from the LFS, the FRS, and the FES, for the period 1994/95 (10,453 adults, LFS; 8,267 adults, FRS). It is not known how many parents were in the sample. The authors reported that the greatest income gain following the immediate replacement of FC with WFTC (assuming no change in employment behaviour) would fall to inactive/unemployed mothers, whose partners were in-work.

The study found that, for men with employed partners, the replacement of FC with WFTC would increase the probability of their entering work by 0.18 ppt. For women with employed partners, WFTC decreased the probability of participation in employment by 0.41 ppt.

Study 14

Heim (2006) reported the effects of the USA's EITC on the average annual hours of both mothers and fathers from two-parent working families, for the period 1985 to 2003. Based on findings for the individual years, the study also assessed the impact of the changes resulting from the expansion to the EITC following the OBRA93. Data for the study was drawn from the 1985-2003 waves of the Current Population Survey. Low-educated married couples with dependent children were purposively selected for analysis (a total of 3,961 observations). For each year in the sample, the study estimated the labour supply of both parents, both in the presence and in the absence of the EITC. The difference between these two was interpreted as the effect of the EITC in that particular year.

Overall (i.e. taking into consideration each of the individual years 1985-2003), the study found that the EITC (relative to no EITC) had a negative effect on the average number of hours worked by both fathers and mothers in the sample. Average decreases for fathers were about four hours annually and about nine hours annually for wives. The results of the study also indicated that, prior to the 1993 expansion the EITC had a minor negative effect on annual hours, whereas after the expansion the negative effect on annual hours increased. The average decrease in the period pre-dating the expansion was 0.26 hours for fathers, 2.16

hours for mothers, while the post-1993 effects were five hours for fathers and 12 hours for mothers. The study also estimated that in both the pre- and post-1993 periods the majority of individuals (75 to 82 per cent) decreased their hours in response to the EITC; this result held for all subgroups. Decreases tended to be larger for younger individuals, although by a small amount. For the post-1993 period, the decrease in hours was larger for families with two or more children, particularly for mothers. The author noted that this was to be expected, given the more generous EITC awards that were made available to larger families as part of the 1993 expansion to the policy.

The author also estimated the effects of hypothetical changes to the EITC programme: two reforms that increased the generosity of awards and two that decreased payments. Increasing its generosity (by increasing payments for those with three or more children or increasing payments for all eligible families by ten per cent) had a negative effect on the average annual hours worked by both fathers (0.76, 0.49 hours) and mothers (3.59, 1.16 hours). In contrast, the two proposals to decrease EITC generosity (rollback to 1992 levels of payment and a wage-based system) increased average working hours of both fathers (3.99, 4.89 hours) and mothers (9.86, 11.39 hours).

Study 15

The study by **Leigh (2007)** estimated the effect of WFTC compared with the previous FC regime on the working hours of two-parent families who were already in paid employment. A sample of adults aged 25-59 years, who were not self-employed, was drawn from the LFS (3,863 individuals with children; 4,086 childless individuals). The two quarterly surveys before the introduction of WFTC (March to August 1999) provided the pre-reform data, with post-reform data coming from the two quarterly surveys after the change (December 1999 to May 2000).

It was reported that the replacement of FC with WFTC led to a rise in the average weekly working hours of non-single mothers with children, relative to the comparison group of childless women from couple families (1.746 hours; $p < 0.01$). Fathers from couple families with children also increased their average weekly hours, but by a lesser amount (0.718 hours) when compared with fathers from couple families. This effect was reported as not being statistically significant.

Study 16

The study by **Lydon and Walker (2005)** investigated the impact of FC and/or WFTC on individual wage growth, based on data from the LFS (1997 to 2003). Wage growth was defined with reference to changes in hourly pay. Both lone mothers and two-parent families were examined. In all, 40,546 couples were included in the final sample, of which 20,155 had dependent children.

The analysis of the summary statistics was largely focused on comparisons of wage growth between those who received tax credits (FC or WFTC) and those

who did not, by observable characteristics, including wage levels, job tenure, and qualifications. The authors reported that these comparisons did not provide clear-cut results, with some comparisons suggesting that FC/WFTC recipients experienced higher wage growth than non-recipients, and other comparisons suggesting the reverse. Based on these mixed findings, the authors concluded that 'at worst, recipients of FC/WFTC have similar wage growth, on average, to that of non-recipients' (2005, p.368).

However, for certain comparisons, the study found that some individuals receiving the tax credits had higher wage growth on average. Focusing on the results that were statistically significant, the authors reported that, for example, married working mothers with qualifications of NVQ level 2 and below had significantly higher wage growth, on average, than those not receiving the credit.

Study 17

McKay (2003) evaluated how WFTC was associated with changes in employment over time among families with dependent children. This study used data from the first three waves of the FACS, 1999-2001. One part of the study focused on analysing FACS interview data as a longitudinal dataset, and involved looking specifically at transitions from single to dual-earner status (and vice versa) among couple families over the period 2000 to 2001. While the main sample for data collection numbered several thousands of families, the analysis of WFTC's impact on single- and dual-earning families was based on a non-random sub-sample of the main study population: that is, all couples who had received WFTC in 2000 and remained in a couple in 2001 (less than 450 couples in total). A single-group design was used. Although there were measurements over time, they were all taken after the intervention was introduced. There was no comparative group.

The study measured the 2001 work status of couple families receiving WFTC in 2000, analysed by their work status in 2000 (i.e. either single- or dual-earning). It found that 16 per cent of single-earning families had become dual-earning one year later, while among those who began as dual-earning families, 17 per cent had become single-earning by 2001. Proportionally, the balance between dual-earner couples becoming single-earners, and single-earner couples becoming dual-earners, was therefore fairly even (a point made by the author). However, given that relatively few dual-earning couples received WFTC (and that this sub-group was greatly out-numbered by single-earning couples in the study sample), the study findings appear to be suggesting that WFTC was associated with an overall increase in the numbers of dual-earning couples.

The study also investigated the working hours of two-parent families, following up parents who were recorded as receiving WFTC in 2000. By 2001, regardless of whether they still received WFTC, a higher proportion of couple mothers (five per cent) worked additional hours than worked fewer hours (four per cent). This was also true for couple fathers (14 per cent of the sample worked more hours and ten per cent worked fewer hours in 2001).

Study 18

The UK study by **Paull *et al.* (2002)** estimated the impact of replacing FC with WFTC on mothers' employment, using pre-reform FRS (1994/95 to 1998/99). The sample contained mothers of pre-school aged children only; including 12,011 observations of married mothers.

It was estimated that, as the result of introducing WFTC, a larger proportion of mothers (0.5 per cent) moved from part-time work to full-time, than the proportion (0.3 per cent) that moved from full-time work to part-time employment.

5.4 Weight of Evidence

During the data-extraction process, a judgement about the quality of reporting in each study was made using the EPPI-Centre WoE framework. Details of the methods are given in Section 4.4 and Appendix C of the Technical Report. The WoE judgements for all the studies are shown in Table 5.2.

Four studies were given a 'medium' overall WoE and the remainder were given 'low/medium'. The lower the overall WoE score, the less confident reviewers were about the internal and/or external validity of the study results.

Table 5.2 WoE judgements

| Study | WoE A Internal Coherence | WoE B Appropriate design/analysis | WoE C Relevance | Woe D Overall |
|----------------------------------|---|--|----------------------------|--------------------------|
| Azmat (2006) | Medium | Low | Medium | Low/medium |
| Blundell <i>et al.</i> (2000) | Medium | Low | High | Low/medium |
| Blundell <i>et al.</i> (2004) | Medium | Low | High | Low/medium |
| Blundell <i>et al.</i> (2005) | Medium | Low | High | Low/medium |
| Bonin <i>et al.</i> (2003) | Medium | Low | Medium | Low/medium |
| Brewer <i>et al.</i> (2006) | High | Low | High | Medium |
| Chzhen & Middleton (2007) | Medium | Low | Medium | Low/medium |
| Creedy <i>et al.</i> (2003) | Medium | Low | Medium | Low/medium |
| Eissa & Hoynes (2006) | High | Low | Medium | Medium |
| Ellwood (2000) | Medium | Low | Medium | Low/medium |
| Flood <i>et al.</i> (2004) | Medium | Low | Medium | Low/medium |
| Francesconi <i>et al.</i> (2007) | Medium | Low | High | Low/medium |
| Gregg <i>et al.</i> (1999) | Medium | Low | Medium | Low/medium |
| Heim (2006) | High | Low | Medium | Medium |
| Leigh (2007) | High | Low | Medium | Medium |
| Lydon & Walker (2005) | Medium | Low | Medium | Low/medium |
| McKay (2003) | Medium | Low | Medium | Low/medium |
| Paull <i>et al.</i> (2002) | Medium | Low | Medium | Low/medium |

5.5 Synthesis of evidence

5.5.1 Introduction

The aim of the synthesis was to answer the following question:

What is the effectiveness of interventions with the potential to reduce in-work poverty in couple families with dependent children?

While acknowledging that there are differences between the interventions in the 18 studies included for synthesis (see Section 5.3.2), our interpretation is that they all involved the modification of tax/benefit arrangements, to make them more generous for those in-work (than the scheme to which they are compared). Although the intervention 'package' evaluated in the study by Flood *et al.* (2004) was not necessarily more generous overall for everyone who would be affected by it, our understanding is that the main element was the cut in income tax, which by itself would have been more generous. Taking the view that the 18 studies form a coherent subset of studies, for the purpose of the synthesis they have been grouped together.

The study outcomes were divided into three categories: wages/income (see Section 5.5.2), employment participation (see Section 5.5.3), and hours worked (see Section 5.5.4). These outcome categories provided the framework for comparing studies in the synthesis.

Summary tables showing the direction of effect as either increasing (positive +) or decreasing (negative -) wages, income, participation or hours are presented at the end of each outcome category, unless that study has not provided sufficient data for that outcome to be able to summarise the direction of effect. These summaries (of 'effect') are based on the overall main results (not sub-group analyses) of each of the studies. Where appropriate, interesting findings based on sub-group analyses are discussed in the text.

5.5.2 Synthesis of results: wages and household income

Three studies reported the effects of the intervention on a number of different income-related outcomes (Azmat, 2006; Flood *et al.*, 2004; Lydon and Walker, 2005). One study calculated whether or not poor two-parent families were better off financially as a result of the intervention, taking into account behavioural responses to the reform (Flood *et al.*, 2004). In contrast, the sole focus of the studies by Azmat (2006) and by Lydon and Walker (2005) was potential indirect consequences of in-work benefits. These two studies measured the impact on wage growth (weekly wages and hourly pay, respectively).

Gains in household income assuming changes in-work patterns

One study measured changes in household income following the introduction of a hypothetical financial intervention, taking into consideration any changes in employment behaviour (working hours) associated with its introduction. Table 5.3 shows the direction of effect found.

Table 5.3 Summary direction of effect: changes in household income

| Study | Direction of effect |
|----------------------------|---------------------|
| Flood <i>et al.</i> (2004) | + |

The study by Flood *et al.* (2004) found that net household incomes for the poorest ten per cent of the working sample increased as a result of the replacement of the actual tax and benefit system with the modified one.

Wages

One argument against the sorts of income subsidy programmes that have been evaluated by the studies in this review is that they trap recipients into low-skilled, low-waged work. Two studies in the review focused solely on answering whether this was the case (Azmat, 2006; Lydon and Walker, 2005).

Table 5.4 Summary directions of effect: changes in wages

| Study | Direction of effect |
|-----------------------|---------------------|
| Azmat (2006) | - |
| Lydon & Walker (2005) | No impact |

The study by Azmat (2006) found a relatively large negative effect on fathers from two-parent families and a much smaller positive effect on partnered mothers, compared with the comparison groups of matched, non-eligible workers. This has been interpreted as a negative net effect. By including a measure which controlled for the change in the tax credit generosity from FC to WFTC, the author concluded that the effect on gross wages was the result of a change in payment method (i.e. through the wage packet) which altered the amount of information available to the employer and allowed them to cut the wages of claimant workers. In contrast, the emphasis in the study by Lydon and Walker (2005) was on whether the incentive to engage in training would be affected by the presence of the wage subsidy, thereby affecting wage growth. In this study an analysis of summary statistics did not produce definitive findings, and the authors concluded that '*at worst, recipients of FC/WFTC have similar wage growth, on average, to that of non-recipients*' (2005, p.368). However, when focus was placed on results that were statistically significant, some groups (for example, married working mothers receiving FC/WFTC with qualifications of NVQ level 2 and below) were found to be have significantly higher wage growth, on average, than those not receiving the credits.

No conclusive reasons were found as to why the effects found by these two studies should be different, as there are no systematic differences in terms of overall study quality, study design (both use non-equivalent comparison groups),

sample characteristics or context (both studies were conducted in the UK). Possible explanations are the use of slightly different outcome measures, or the methods of analysis. The results in the study by Lydon and Walker (2005) that were relevant for this review came from an analysis of summary statistics, which the authors acknowledge as being methodologically weak.

5.5.3 Synthesis of results: employment participation

Ten studies reported the effects of tax and benefit reforms on the employment participation of second earners in couple families with dependent children (Blundell *et al.*, 2000; Blundell *et al.*, 2004; Blundell *et al.*, 2005; Bonin *et al.*, 2003; Brewer *et al.*, 2006; Eissa and Hoynes, 2006; Ellwood, 2000; Francesconi *et al.*, 2007; Gregg *et al.*, 1999; McKay, 2003).

Three of these studies reported changes to the proportions of workless, single-earner and dual-earning couple families with dependent children in the study population (Blundell *et al.*, 2004; Bonin *et al.*, 2003; Brewer *et al.*, 2006). A further four studies only reported changes to the participation of sub-groups of mothers and fathers with employed partners. As findings for these sub-groups were reported separately, we then estimated whether or not second earners overall were more or less likely to participate in the labour market; that is, we have estimated the net effect (Blundell *et al.*, 2000; Blundell *et al.*, 2005; Eissa and Hoynes, 2006; Francesconi *et al.*, 2007). We also interpreted the reports reported by McKay (2003), thereby allowing a comparison with other studies included in Table 5.5. Since our interpretations were sometimes based on estimated sub-sample sizes, these estimated net effects are clearly open to challenge.

The study by Ellwood (2000) focused solely on women. Consequently, it was not possible to estimate the net effect of the intervention on the participation of second earners overall. This study is therefore not included for comparison with others in Table 5.5, but has been discussed later on in this section in a comparison with other studies that report the effect on mothers with employed partners.

The study by Gregg *et al.* (1999) considered movements into work only, among inactive/unemployed individuals with employed partners. It did not model exits from work (i.e. did not estimate the extra movements out of work among working individuals with employed partners that would have resulted from the intervention) and will therefore not be discussed further in this part of the synthesis.

Table 5.5 Summary directions of effect: changes in employment participation

| Study | Second earners overall (men and women) | Mothers with employed partners | Fathers with employed partners |
|----------------------------------|--|--------------------------------|--------------------------------|
| | Direction of effect | | |
| Blundell <i>et al.</i> (2000) | –* | – | –* |
| Blundell <i>et al.</i> (2004) | – | – | + |
| Blundell <i>et al.</i> (2005)** | –* | –* | – |
| Bonin <i>et al.</i> (2003) | – | + | – |
| Brewer <i>et al.</i> (2006) | – | – | + |
| Eissa & Hoynes (2006) | –* | – | + |
| Francesconi <i>et al.</i> (2007) | +* | + | + |
| McKay (2003) | +* | | |

* Indicates reviewers' interpretation.

** Based on a combination of LFS and FRS analyses.

Table 5.5 reports the directions of effects for the eight studies that provide sufficient data on the employment participation of second earners overall (i.e. both men and women). Six studies found that the interventions reduced the participation of second earners in the labour market (Blundell *et al.*, 2000; Blundell *et al.*, 2004; Blundell *et al.*, 2005; Bonin *et al.*, 2003; Brewer *et al.*, 2006; Eissa and Hoynes, 2006). Two studies found that the intervention increased the employment participation of second earners (Francesconi *et al.*, 2007; McKay, 2003).

The study design used by McKay (2003) is the most likely reason for the positive effect that appears to be suggested by this study. There were no counterfactual outcomes (i.e. no estimate of the circumstances that would have prevailed had a new policy, or policy change, not been introduced). It is not clear why the results found by Francesconi *et al.* (2007) should be different from the other studies. There are no systematic differences in terms of study design, study quality, outcome measure, or sample characteristics that correspond to the different direction of effect. It was not, for example, the only study to focus on poor parents. A possible explanation might be that the study by Francesconi and colleagues used a definition of 16 hours per week as a cut-off point for defining in or out of work. The majority of other studies do not provide this information; but, as this particular number of hours determines eligibility for many of the UK employment benefits and tax credits, it might be assumed that at least some of the UK studies also used this indicator.

Six studies reported a negative effect on the employment participation of mothers with employed partners (Blundell *et al.*, 2000; Blundell *et al.*, 2004; Blundell *et al.*, 2005; Brewer *et al.*, 2006; Eissa and Hoynes, 2006; Ellwood, 2000). Two studies reported a positive effect on this sub-group (Bonin *et al.*, 2003; Francesconi *et al.*, 2007).

Four studies reported a positive impact on the employment participation of fathers with employed partners (Blundell *et al.*, 2004; Brewer *et al.*, 2006; Eissa and Hoynes, 2006; Francesconi *et al.*, 2007). Three studies reported a negative effect on the employment participation of fathers with employed partners (Blundell *et al.*, 2000; Blundell *et al.*, 2005; Bonin *et al.*, 2003).

Four studies found differential labour supply responses according to gender. Of these, the majority suggest that the reduction in the proportion of dual-earner families favours the formation of male-breadwinner model families (Blundell *et al.*, 2004; Brewer *et al.*, 2006; Eissa and Hoynes, 2006), with only one study suggesting that the less common female-breadwinner model family is favoured by the introduction of the intervention in question (Bonin *et al.*, 2003).

5.5.4 Synthesis of results: working hours

Twelve studies reported the effects of changes in family income induced by tax/benefit reforms on the working hours of one or more members of couple families with dependent children (Blundell *et al.*, 2000; Blundell *et al.*, 2004; Bonin *et al.*, 2003; Brewer *et al.*, 2006; Chzhen and Middleton, 2007; Creedy *et al.*, 2007; Eissa and Hoynes, 2006; Flood *et al.*, 2004; Heim, 2006; Leigh, 2007; McKay, 2003; Paull *et al.*, 2002).

One study reported the impact on the working hours of two-parent families as a whole; no findings were reported for individual members of these families (Creedy *et al.*, 2003). Eight studies reported the impact on the working hours of both parents separately. In seven of these studies the direction of effect for the sub-groups of mothers and fathers in the study sample was such that it was possible to interpret an overall direction of effect for two-parent families (Blundell *et al.*, 2004; Bonin *et al.*, 2003; Brewer *et al.*, 2006; Eissa and Hoynes, 2006; Flood *et al.*, 2004; Heim, 2006; Leigh, 2007).

The information in the study by McKay (2003) does not detail whether there was an overall increase or decrease in hours, and this cannot be calculated from the data provided. As a result, this study cannot be considered any further in this outcome category.

Three studies measured changes to working hours for mothers only. As an overall effect on couple families could not be calculated, these studies have also been excluded from further consideration in this part of the synthesis (Blundell *et al.*, 2000; Chzhen and Middleton, 2007; Paull *et al.*, 2002).

Table 5.6 Summary directions of effect: changes in working hours

| Study | Direction of effect | | |
|-------------------------------|-------------------------|-------------------|-------------------|
| | Couple families overall | Partnered mothers | Partnered fathers |
| Blundell <i>et al.</i> (2004) | –* | – | – |
| Bonin <i>et al.</i> (2003) | –* | No impact | – |
| Brewer <i>et al.</i> (2006) | –* | – | – |
| Creedy <i>et al.</i> (2003) | – | | |
| Eissa and Hoynes (2006) | –* | – | – |
| Flood <i>et al.</i> (2004) | +* | + | + |
| Heim (2006) | –* | – | – |
| Leigh (2007) | +* | + | + |

* Indicates reviewers' interpretation.

Table 5.6 shows the direction of effect for changes in working hours. Six studies found an overall negative effect on the working hours of couples with dependent children (Blundell *et al.*, 2004; Bonin *et al.*, 2003; Brewer *et al.*, 2006; Creedy *et al.*, 2007; Eissa and Hoynes, 2006; Heim, 2006). Two studies found an overall positive effect on couples working hours (Flood *et al.*, 2004; Leigh, 2007). These results should not be over-interpreted as the effect sizes in some studies were very close to zero in either direction.

The study by Heim (2006) also conducted alternative analyses that evaluated hypothetical redesigns of the EITC programme that decreased the generosity of payments (a wage-based EITC and a roll-back to the 1992 programme). It found that both mothers and fathers increased their working hours in response to these changed designs. (These findings are not presented in Table 5.6.)

Across the group of eight studies presented in Table 5.6, there do not appear to be any systematic differences between the studies in terms of study quality, method, sample, or outcome measure that correspond to the differences in direction of effect. In the case of the study by Flood *et al.* (2004) the arguments made earlier may also apply. It is possible that the intervention in this study was more generous overall, but that the difference in the exact mechanism – the tax 'pull' plus the benefits cut 'push' – played a role in the different directions of effect found. However, this argument is less convincing, given that one 'pull only' study (Leigh, 2007) also found positive effects.

5.6 Summary of synthesis results

- Eighteen studies were identified that answered the synthesis question.
- Twelve studies were conducted on UK populations, with three studies from the USA, and one each carried out in Germany, Sweden and Australia.
- All 18 studies evaluated financial interventions, the main focus of which modified tax/benefit arrangements in some way, to make them more generous for those in-work (i.e. the interventions allowed the worker to keep more of their salary/income, than in the scheme to which they were being compared).
- No studies of other kinds of interventions for in-work poor couple families with dependent children were identified.
- One low/medium quality study measured the impact on household income (following behavioural responses to the policy) and found that income increased.
- Two low/medium quality studies measured the impact on wages and had different findings.
- Eight studies informed us about the effect that the interventions had on the employment participation of second earners overall (i.e. both men and women):
 - two studies found a positive effect (two low/medium quality);
 - six studies found a negative effect (two medium, four low/medium quality).
- Eight studies informed us about the effect that the interventions had on the overall working hours of couple families (i.e. both men and women):
 - two studies found a positive effect (two medium quality);
 - six studies found a negative effect (two medium, four low/medium quality).

6 Conclusions and implications

6.1 Discussion

At the beginning of the report, the following question was asked: *What is the nature and extent of the research that has been undertaken on the barriers to, and facilitators of, reducing in-work poverty in families with dependent children?* This broad question drove the initial stage of the review and provided the conceptual basis for the systematic map of research in the area (see Chapter 3). Responding to suggestions from the DWP, we went on to ask a narrower question which we aimed to answer with a synthesis of the relevant studies: *What is the effectiveness of interventions with the potential to reduce in-work poverty in couple families with dependent children?*

The mapping exercise (stage 1 of the review) located and described a total of 439 studies which were relevant for answering the broad review question. Of these, 285 studies evaluated an intervention. Eighteen evaluative studies were identified as relevant for answering the second review question and were included in the synthesis (stage 2 of the review). Due to the nature of the available research evidence, there are limitations to its capacity for fully answering the synthesis question, and this has implications for the conclusions about effective solutions to in-work poverty that can be drawn from the synthesis findings. This will be discussed further.

The relatively limited number of studies included in the second stage of the review reflects the more focused synthesis question and the inclusion/exclusion criteria that were designed to locate studies capable of answering it. The synthesis question was clearly narrow, in as much as it specifically asks 'what works', and in terms of its sole focus on a particular population group: working couple families with children. Whilst policy interest had some bearing on the direction of the synthesis question, it was also largely determined by the information contained in the map (where research activity was concentrated and where there were gaps).

This information led the Advisory Group in the direction of a synthesis question that was also relatively broad in some aspects. In particular, the question indicated an interest in all types of intervention, and it allowed the inclusion of studies with different research designs.

Nonetheless, the 18 interventions that feature in the synthesis were all of a financial nature, with the majority evaluating tax credits schemes (83 per cent). The review was unable to identify a number of other types of interventions that have been identified (Harker, 2006) as potential solutions to three main causes of in-work poverty among families with children: **(1) low pay, (2) families relying on one earner and (3) single/dual earners not working enough hours.**

These other types of intervention include the following:

- human capital development/career advancement initiatives to support parents to progress in-work so that low-paid workers do not become trapped on low pay;
- minimum wage, sector pay agreements or voluntary approaches to improve wage levels;
- employment support services for second earners, to help with preparing for and moving into work;
- initiatives (for example, extended or more flexible childcare provision) to enable parents to increase their hours and/or progress in-work.

As these types of interventions were not identified in the second stage of the review, they were not investigated in the synthesis.

Despite differences in terms of their design and operation, the 18 interventions that were identified and included in the synthesis were reasonably similar. Each of them aimed to address one or more of the aforementioned three main causes of working poverty. As measures that supplemented low pay and/or earnings, a key function of each was to address the problem of **low pay** as a cause of in-work poverty. All 18 interventions were also motivated by the desire to tackle a second problem: **earners not working enough hours.** For example, lowering the rate at which tax credits or other means-tested benefits are withdrawn as income rises, subsidising social insurance payments, and reducing the rate of income tax, all allow workers to keep more of any additional earnings that they make, reducing the disincentive to increase the number of hours worked.¹⁰ Whilst, in this review, there are no interventions per se that focused solely on the third cause of working poverty, **families relying on one earner**, two of the UK tax credits

¹⁰ For example, both the 1999 replacement of FC with WFTC and the 1993 expansion to the EITC involved changes to the 'withdrawal' or 'taper' rate. Since 1995, UK tax credits schemes have typically also included an extra payment for those working over 30 hours per week, to provide an additional incentive to increase working hours.

reforms have involved technical adjustments that have been explicitly concerned with encouraging dual-earning amongst couples. In both cases, childcare credits (payable alongside the main awards) were made conditional on both earners working more than 16 hours per week.¹¹

Although we set our selection criteria to identify studies that considered transitions into and out of income poverty, none of the studies included for synthesis measured changes in income using thresholds derived from percentages of median income for the whole population (i.e. poverty lines). All 18 studies measured alternative outcomes that, nonetheless, directly relate to the three main causes of poverty identified by Harker (2006). Sixteen studies were concerned solely with measuring behavioural outcomes: employment participation and/or working hours. Three studies measured financial outcomes: wages and household income (of these, one study also measured behavioural changes). In this review, these outcomes were considered relevant proxies for the key outcome of interest: changes to levels of in-work poverty.

There are drawbacks, nonetheless, to a review relying on these proxy measurements. The Government's targeting of a relative measure of child poverty means that whether a working household is defined as poor depends on how much income it has relative to the median household. As a result, simply looking at the effects of policies on the employment behaviour, or even income-related outcomes, of families cannot inform us as to whether they have successfully made the transition out of poverty. Clearly, therefore, the studies included for synthesis are limited in terms of how far they can illuminate us about effective solutions in-work poverty among couple families with dependent children. Whilst we acknowledge a slight tension between this limitation and the strong message that systematic reviews offer the best method for assessing a body of evidence, the review's reliance on proxy outcomes for working poverty is a feature of the available research evidence, and does not reflect any decision on our part to focus exclusively on these particular outcomes.

Conducting the review was demanding. Studies varied in the manner and extent to which they reported their methods and findings. Particular difficulties were caused by the current under-development of methods for reviewing evaluations based on econometric techniques. Interpreting the results of the review was also challenging. With few exceptions, the overall quality of the studies in the review was judged to be relatively low (medium or lower), largely because of WoE B which focused on study design. None of the included studies were prospective, randomised controlled experiments: 'gold standard' evaluations widely accepted as giving firm and reliable findings. In practice, this means that one can be less confident that the results of an individual study provide a valid indicator of the

¹¹ Consider also the within-year £25,000 income disregard in WTC which allows a family's income to increase considerably without affecting their tax credits until the following year (although this is not solely targeted at couple families).

impact of the intervention. There are no agreed approaches to interpreting findings from studies with such limitations. Opinions differ about whether having a number of studies with the same effect (i.e. positive or negative) actually overcomes the validity issue created by the lack of high quality studies. Similarly, there is no consensus about how to interpret a synthesis of lower quality studies when the results in the individual studies are not in complete agreement (as in this review).

Included in the synthesis were a single study estimating the effects of an intervention on household income and two studies measuring changes to wages (hourly rate and weekly wages) which had different findings. These results do not allow us to say with greater confidence that any of these effects are valid.

Sixteen studies in the synthesis measured behavioural outcomes: employment participation and/or working hours. Applying a simple vote counting approach to these studies highlights that a total of six studies found that the interventions decreased the overall employment participation of second earners (i.e. taking into consideration both male and female second earners), and two studies found an increase in their participation. The findings for changes to working hours were similar: six studies found that the interventions led to an overall decrease in the working hours of two-parent families and two studies found an increase (again, these studies took into consideration the changes affecting both male and female members of couple families). Our interpretation of these results is that the evaluated financial interventions do not appear, on average, to have resulted either in attracting more potential second earners into work or encouraging members of two-parent families to work more hours. The alternative analyses conducted in the study by Heim (2006), which found that participants responded to decreased financial payments by **increasing** working hours, adds weight to the suggestion that there is an inverse relationship between the level of in-work financial support and labour market activity. However, we are not entirely convinced that the interventions did have the opposite effect to that intended. In many studies the effect sizes were relatively close to zero in both directions, the review failed to show a consistent direction of effect, and the overall quality of the reviewed studies was judged to be relatively low (largely reflecting the design of the studies). The limitations of the available evidence suggest we should be cautious about causal attribution. Our interpretations are, therefore, to some degree speculative and should be treated with caution. What we can reliably say on the basis of the studies included for synthesis is that they do not provide a conclusive answer about effective routes out of in-work poverty for couple families with children. Albeit a somewhat disappointing finding, this is an important finding nonetheless.

There are some important issues that should also be taken into consideration. Firstly, while the interventions do not appear to be having the impacts that were intended, this does not mean that they are not having any positive impact on the population of interest to this review. Many of the results are an average; the interventions will usually be having both a positive and a negative effect for different individuals among a group of people. A few of the reviewed studies did conduct additional analyses for sub-groups of the main population, and for

whom a positive effect was found (i.e. the opposite to that found for the sampled population as a whole). From the evidence presented in most of the studies in the synthesis, however, these details are obscured. Secondly, the findings from this review do not mean that the interventions are ineffective per se. They may be having the impacts looked for in other population sub-groups that were also targeted by the intervention (e.g. workless parents or low paid workers, generally). We make no further comment on this, as it was outside the scope of this review. Thirdly, our findings do not mean that financial interventions per se cannot be effective for working couple families with dependent children.

Finally, some of the studies suggest that the increased generosity of the evaluated schemes led to at least some people reducing their hours of work, or refraining from participation in employment (typically the mother). Presumably, this will be without too significant a loss in overall family income, as the introduction or increased generosity of the income supplement will partly compensate the family for loss of earnings. While this might not be judged favourably in terms of targets for reducing working poverty, it might well be interpreted as a positive result by the families themselves (for instance, in terms of increased leisure or time spent with children).

6.1.1 Strengths and limitations of this systematic review

A major strength of the review lies in its systematic and comprehensive nature. The process of systematically identifying, screening and critically appraising the studies helps to ensure that the review process is transparent, replicable, updateable, and the findings are based on well-conducted studies, where these are available.

Another important strength is the involvement of the DWP in the review, especially at the point of moving from the map to the synthesis. This helped to ensure that the review was more policy-relevant.

The presentation of the results of the studies, allowing for direct comparison across studies with similar outcome measures (in terms of the direction of effect), facilitates ease of interpretation.

By focusing on recent studies conducted internationally, the review was able to look at a range of studies addressing the review question. This meant that the usefulness of the review was broader than one which might only have focused on evaluations of current initiatives in the UK.

The coding tool used to describe the included studies was developed iteratively and collaboratively with the review commissioners. This ensured that appropriate definitions and categories were used, to maximise the relevance of the review for users. Analysis was undertaken using custom-designed in-house software, EPPI-Reviewer, which allowed cross-checking and sophisticated analyses to be undertaken. The coding for each of the 18 studies included for synthesis was carried out independently by two members of the team, thereby minimising the risk of error and improving the quality of the data.

Modelling techniques are increasingly employed in evaluations of public policy. The technical nature of many of these evaluations makes them difficult for non-specialists to interpret. Consequently, the potential for them either to be disregarded altogether, or accepted at face value, may be heightened. One limitation that reviewers encountered was the lack of a standardised tool for assessing the quality of studies employing the techniques of econometric modelling. Two strengths of this review are that it has involved the adaptation of existing in-house tools for assessing this complex material (initiating some limited methodological developments in this area) and it has reported findings and provided explanations in non-technical and accessible terms.

The review attempted to identify all relevant literature through the use of systematic and comprehensive searches of a wide range of databases and websites, with the latter being a particularly fruitful source of studies. While we did not handsearch journals, and searches of grey literature were limited, citation checking was carried out on studies included in the synthesis, and contact with experts in the field was also only carried out at this stage. As with all systematic reviews, the timescale for conducting the review (with a cut-off date for retrieval of reports) can result in relevant studies being missed.

The application of both the selection criteria and descriptive codes for the map was carried out by two reviewers working independently on sub-samples of studies, as part of the review's quality-assurance process.

The design of studies included in the synthesis allowed for sophisticated analysis of the degree of association between tax/benefit changes and the dependent variables (e.g. employment participation). However, as with any quantitative evaluation of interventions utilising non-experimental retrospective designs, the role of subjective judgment and assumption (since theory plays a central role in the selection of control variables) has the potential to bias the findings of the studies. As none of the included studies used an experimental design, they have an increased likelihood of being affected by selection bias. This is reflected in the overall quality of the evidence seen in the WoE ratings: no studies were judged to provide higher than an overall medium weight of evidence. Therefore, conclusions about causality must be considered tentative.

The review extended only to a consideration of the effectiveness of modifying the tax/benefit system on in-work poverty (or relevant proxy outcomes). This was a limitation imposed by the agreed focus of the review question, but meant that little contribution was made to discovering the detailed mechanisms through which tax/benefit system changes might affect outcomes.

Studies employed different analytical models, different methods of analysis, and different methods of constructing both the dependent and independent variables. Making comparisons across studies was therefore difficult, even when they measured the same dependent variable. For these reasons, the review has not standardised the measures of 'effect' presented in the studies, or quantitatively

synthesised the results to identify an overall measure of effect. As a result, the review findings are not able to clearly quantify the size of effect or make any claims as to the statistical significance of such an effect.

6.2 Implications

6.2.1 Policy

This review makes an important contribution to the evidence base on working poverty. Key to this has been the specific focus on working two-parent families in the synthesis stage of the review. This group has previously been under-researched, but is a key target group in terms of reducing in-work poverty. Although the key findings of the review have focused on the gaps in this evidence base and the absence of conclusive evidence about effective routes out of in-work poverty for working couple families, these are nonetheless very important findings in themselves, as they inform strategies for future research to inform policy.

Overall, we would strongly recommend that resources are devoted to further quantitative work on the relationship between financial interventions and changes to household income that take full account of behavioural responses to reforms and the interaction between different means-tested benefits. This should be supplemented by rigorous qualitative analysis, to unpack some of the more subtle relationships that quantitative analysis cannot detect.

The absence of evaluations of anything other than financial interventions would suggest that, if non-financial interventions are in place for working couple families with children, there is a need to commission rigorous evaluations of them. This also applies to those financial interventions that were not identified (for example, minimum wage initiatives).

However, it is important that future research uses more rigorous designs. Simply conducting more evaluation studies with weak research designs will not add to, or strengthen, the evidence base in ways which will be helpful. The lack of available high quality evidence on financial interventions for couple families with dependent children has been noted several times throughout the report. One of the elements of this is the non-use of prospective study designs that effectively control for the wide variety of biases that can impact on intervention evaluations. Such designs provide greater confidence in attributing causal relationships to observed findings, with prospective, RCTs widely viewed as the 'gold standard' for proving the efficacy of interventions. While acknowledging that there are practical, methodological, and ethical issues that need to be addressed for such study designs to be used in this field, it is important that policymakers consider evaluation before rolling out widespread changes to systems. Ideally, new policies should have evaluation of their effectiveness built into them from the start (Cabinet Office, 1999).

6.2.2 Research

This review highlights that the evidence base for working couple families with dependent children is limited, despite the relatively large number of studies identified during the mapping exercise that included two-parent families in their sample. Researchers conducting studies that include couple families should undertake separate analyses by family type, including by employment status. Ideally, study populations should include both members of the couple, and, where couples are the unit of analysis, there should be increased focus on intra-family dynamics. Without both partners in the sample, it may be harder to gain an understanding of how couples make decisions about work (including the role of the partner's employment status).

In conducting this review, we found evidence of poor reporting practices in a number of studies, including a lack of consistency in the reporting of data between different publications of the same study. Researchers should improve the quality of their reporting, particularly in terms of study design, intervention/comparator characteristics, and numerical data. Studies should also report raw data and information about the precision of estimates, and not just the point estimate/correlation coefficient alone.

As many of the results are an average, the interventions will usually be having both a positive and a negative effect for different individuals among a group of people. However, from the evidence presented in most of the studies, these details are obscured. Research should report findings for each income range separately, particularly when the focus of the study is on poverty.

This review also highlights the growing use of econometric modelling for policy analysis. Whilst it is arguing that RCTs should be promoted and encouraged, this should not be construed as an argument that econometrics has no role to play in policy evaluation. Nonetheless, in conducting this review, the relatively unexplored issue of the credibility of this type of policy evaluation comes to the fore. Adequate, formal validation exercises were lacking in many of the studies included for synthesis (or, at least, they were not clearly reported). Researchers conducting this type of evaluation need to develop and put into practice rigorous validation practices. This will aid the improvement of the models themselves and help ensure greater credibility of the findings. As there do not appear to be any developed and tested quality appraisal tools (outside the health field) for assessing the validity of evaluations based on econometric methods, this urgently needs to be addressed.

As we can make no clear recommendations about effective solutions to the problem of in-work poverty on the basis of the studies included in the synthesis, it may be useful to look into some of the research on the employment of coupled women and men more generally (that is, including childless couples and the non-poor). This research could potentially make a useful contribution to the policy issue. During the literature search, we identified a body of literature stretching

back several decades on the effect of different tax regimes on the labour supply of men and women, with and without children, and including partnered couples. In recent years, several European countries have introduced tax credit schemes to help raise participation among lower paid workers. Although many of these have already undergone evaluation, we found that parental status was ignored and so the studies did not meet our inclusion criteria. In conducting the review, we also became aware of evaluations that measured the impacts on parents (especially mothers) of widely-available interventions, such as family allowances, but which placed no emphasis on low-income populations (again, leading to their exclusion during the first stage of the review). As a first step, systematic reviewing methods could be used to ascertain the full extent of this evidence base.

References

General references in the text

Abel-Smith, B. and Townsend, P. (1965) *The poor and the poorest*, London: Bell.

Boruch, R. and Wortman, P. (1979) 'Implications of education evaluation for education policy', in Berliner D.C. (ed.) *Review of research in education*, Washington, DC: American Educational Research Association, 309-361.

Brewer, M. and Shephard, A. (2004) *Has Labour made work pay?* York: Joseph Rowntree Foundation.

Cabinet Office (1999) *Professional policy making for the 21st Century: report by the Strategic Policy Making Team*, London: Stationery Office.

Cook, T.D. and Campbell, D.T. (1979) *Quasi-experimentation: design and analysis issues in field settings*, Chicago: Rand McNally.

Department for Work and Pensions (2008) *Households below average income: an analysis of the income distribution 1994/5-2006/7*, London: DWP.

Gough, D. (2007) 'Weight of evidence: a framework for the appraisal of the quality and relevance of evidence', in Furlong, J. and Oancea, A. (eds.) *Applied and Practice-based Research. Special Edition of Research Papers in Education*, 22, (2), 213-228.

Gradus, R. and Julsing, J. (2001) *Comparing different European income tax policies making work pay*, OCFEB Research Memorandum 0101, Rotterdam: Research Centre for Economic Policy.

Harker, L. (2006) *Delivering on child poverty: what would it take?* Norwich: HMSO.

Higgins, J. and Green, S. (eds.) (2008) *Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.0* [updated February 2008] The Cochrane Collaboration.

Kemp, P., Bradshaw, J., Dornan, P., Finch, N. and Mayhew, E. (2004) *Routes out of poverty: a research review*, York: Joseph Rowntree Foundation.

Millar J. and Ridge, T. (2001) *Families, poverty, work and care: a review of the literature on lone parents and low-income couple families*, DWP Research Report 153, Leeds: Corporate Document Services.

Soto, J. (2002) 'Health economic evaluations using decision analytic modelling: principles and practices – utilization of a checklist to their development and appraisal', *International Journal of Technology Assessment in Health Care*, 18, 1, 94–111.

Tate, R. (1982) 'Experimental design', in Mitzel, H. and Rabinowitz, W. (eds.), *Encyclopaedia of education research*, London: Collier Macmillan, 553-561.

Thomas, J. and Brunton, J. (2006) *EPPI-Reviewer 3.0: analysis and management of data for research synthesis (EPPI-Centre software)*, London: Social Science Research Unit, Institute of Education, University of London.

Torgerson, C. and Torgerson, D. (2001) 'The need for randomised controlled trials in educational research', *British Journal of Educational Studies*, 49, 316-318.

Weinstein, M.C., O'Brien, B., Hornberger, J., Jackson, J., Johannesson, M., McCabe, C. and Luce, B.R. (2003) 'Principles of good practice for decision analytic modelling in health-care evaluation: report of the ISPOR Task Force on good research practices – modelling studies', *Value Health*, 6, 1, 9-17.

References for studies included in the synthesis

Azmat, G. (2006) *The incidence of an Earned Income Tax Credit: evaluating the impact on wages in the UK*, CEP Discussion Paper 724, London, London School of Economics, Centre for Economic performance.

Blundell, R., Brewer, M. and Shephard, A. (2004) *The impact of tax and benefit changes between April 2000 and April 2003 on parents' labour supply*, London: Institute for Fiscal Studies.

Blundell, R., Brewer, M. and Shephard, A. (2005) *Evaluating the labour market impact of Working Families' Tax Credit using difference-in-differences*, London: HM Revenue and Customs.

Blundell, R., Duncan, A., McCrae, J. and Meghir, C. (2000) 'The labour market impact of the Working Families' Tax Credit', *Fiscal Studies*, 21, 1, 75-104.

Bonin, H., Kempe, W. and Schneider, H. (2003) 'Household labour supply effects of low-wage subsidies in Germany', *Journal of Applied Social Science Studies*, 123, 199-208.

- Brewer, M., Duncan, A., Shephard, A. and Suarez, M. (2006) 'Did the Working Families' Tax Credit work? The impact of in-work support on labour supply in Great Britain', *Labour Economics*, 13, 699-720.
- Chzhen, Y. and Middleton, S. (2007) *The impact of tax credits on mothers' employment*, York: Joseph Rowntree Foundation.
- Creedy, J., Kalb, G. and Kew, H. (2003) 'Flattening the effective marginal tax rate structure in Australia: policy simulations using the Melbourne Institute tax and transfer simulator', *Australian Economic Review*, 36, 2, 156-172.
- Eissa, N. and Hoynes, H. (2006) 'The hours of work response of married couples and the Earned Income Tax Credit', in Agell, J. and Sorensen, P (eds), *Tax policy and labor market performance*, Cambridge, MA: MIT Press.
- Ellwood, D. (2000) 'The impact of the Earned Income Tax Credit and social policy reforms on work, marriage, and living arrangements', *National Tax Journal*, 53, 4, 1063-1105.
- Flood, L., Hansen, J. and Wahlberg, R. (2004) 'Household labour supply and welfare participation in Sweden', *Journal of Human Resources*, 39, 4, 1008-1032.
- Francesconi, M., Rainer, H. and van der Klaauw, W. (2007) *The effects of in-work benefit reform in Britain on couples: theory and evidence*, Bonn: Institute for the Study of Labor.
- Gregg, P., Johnson, P. and Reed, H. (1999) *Entering work and the British tax and benefit system*, London: Institute for Fiscal Studies.
- Heim, B.T. (2006) *The impact of the earned income tax credit on the labor supply of married couples: structural estimation and business cycle interactions*, Working Paper, Washington, DC: US Department of the Treasury.
- Leigh, A. (2007) 'Earned Income Tax Credits and labor supply: new evidence from a British natural experiment', *National Tax Journal*, 60, 2, 205-224.
- Lydon, R. and Walker, I. (2005) 'Welfare to work, wages and wage growth', *Fiscal Studies*, 26, 335-370.
- McKay, S. (2003) *Working Families' Tax Credit 2001*, Department for Work and Pensions Research Report 181, Leeds: Corporate Document Services.
- Paull, G., Taylor, J. and Duncan, A. (2002) *Mothers' employment and childcare use in Great Britain*, London: Institute for Fiscal Studies.

This report presents the findings of a systematic review of the evidence base relating to working poor families with dependent children. The review aims to increase understanding of research in this area and produce findings that will help inform future policy and research. The systematic review was conducted in two stages.

The first stage described the research that has been undertaken on the barriers to, and facilitators of, reducing in-work poverty in families with dependent children. Stage two of the review involved a synthesis of a subset of these studies, focusing on the effectiveness of interventions with the potential to reduce working poverty in two-parent families.

If you would like to know more about DWP research, please contact:
Paul Noakes, Central Social Research Services,
Department for Work and Pensions,
2nd floor Floor Adelphi, 1–11 John Adam Street, London WC2N 6HT
www.dwp.gov.uk/asd/asd5/rrs-index.asp